

國立交通大學

電機與控制工程學系

碩 士 論 文

以內向式雙鏡頭取像之影像縫合技術

An Image Stitching Technique with Inward Image Acquisition



研 究 生：萬鴻政

指 導 教 授：董蘭榮 博士

中 華 民 國 九 十 九 年 九 月

以內向式雙鏡頭取像之影像縫合技術

An Image Stitching Technique with Inward Image Acquisition

研 究 生：萬鴻政

Student : Hong-Cheng Wan

指 導 教 授：董蘭榮 博士

Advisor : Lan-Rong Dung



A Thesis

Submitted to Department of Electrical and Control Engineering

September 2008

Hsinchu, Taiwan, Republic of China

中華民國九十九年九月

以內向式雙鏡頭取像之影像縫合技術

研究生：萬鴻政

指導教授：董蘭榮 博士

國立交通大學電機與控制工程學系

摘要

本篇論文主要探討在雙鏡頭時，內向式和外擴式的兩種擺法所造成的差異，近年來影像縫合越來越熱門，因為大家都希望能夠得到更大視角的影像，所以我們透過相機焦點距離、縫合後的像素差距、相機最大視角還有重疊區域，分別去討論和比較這兩種擺法所得到的結果，發現內向式的影像縫合可以大幅降低”鬼影”現象，而且內向式影像縫合比外擴式影像縫合有著較大的重疊區域，如此可以增加特徵點的匹配數，進而加強縫合的準確性。除此之外，之前的演算法在近距離縫合時會出現問題，主要是當近距離拍攝影像時，影像容易出現明暗不一，會造成原本的演算法在匹配特徵點時出現錯誤，所以我們使用不同的方法來解決影像明暗不一時無法縫合的問題，我們主要用相關性、像素值排列還有邊緣判斷三種不一樣的方法來試著改善此問題，也比較了這三種方法的效果和速度。最後決定使用邊緣判斷，因為他的速度比其他兩種方法都快了 2-3 倍，而且也能夠縫合大部分的圖。

An Image Stitching Technique with Inward image Acquisition

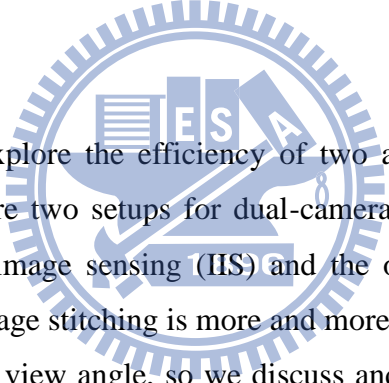
Student : Hong-Cheng Wan

Advisor : Lan-Rong Dung

Institute of Electrical and Control Engineering

National Chiao-Tung University

Abstract



This thesis aims to explore the efficiency of two approaches for dual-camera image acquisition. There are two setups for dual-camera system, considered in this thesis. One is the inward image sensing (IIS) and the other is the outward image sensing (OIS). Recently, image stitching is more and more popular. Everyone wants to get image which has wider view angle, so we discuss and compare the result of two arrangements by their distance of focal point, the disparity after stitching, the widest view angle, and overlapped area. We found that IIS can decrease the phenomenon “ghost-effect.” Otherwise, IIS has wider overlapped area, it can increase the matching point of feature points, and increasing the accuracy of image stitching. Besides, there is a problem when we stitch image at short range in previous algorithm. Because there is different brightness in two images when we take pictures at short range. It will make mistake when we are matching feature points in previous algorithm. So, we use some different method to solve the problem. We mainly use “correlation matching” ,”pixel ranking” and “edge enhancement” to improve the problem. And also, we compare the effect and speed between the three methods. Finally, we choose “edge enhancement”, because it has two or three times speed than other method and it can stitch most photos.

誌謝

這篇論文得以完成，首先要感謝的是我的指導教授，董蘭榮教授，在碩士班的兩年間，當我陷入瓶頸或遇到困難時，老師總是給我很多有用建議，讓我能研究上有所前進和突破，而且不至於失去自己的方向。老師不但讓我在研究上學習到很多，更是常常在meeting或聚餐時，講述很多人生的道理和作人處事的態度或是經驗談，有時候我覺得這些比研究還要重要，讓我在這兩年間獲益良多，真的很感謝老師。

也感謝實驗室的學長——邱哥、罩哥、handsome、智聖、華哥，巫穎毅和鄭銘輝，在我的求學過程中給予指點與幫助，而且帶給實驗室很多笑聲，讓我碩一過得很開心。還有陪伴兩年的同學們——阿原、建樺、barry，以及學弟妹——黃仁猷、小術、小水滴、介恩、楊耀銘，有了你們，使得實驗室不再苦悶，有出遊，有團購，還有烤肉，給了我一個充實又有趣的碩二。當然，還要感謝兩位辛苦的助理——唯茵和維婷，幫實驗室處理了很多事情。最後要感謝我的家人，有了你們的支持，使我能無後顧之憂，安心地完成碩士班學業。

謹將此論文獻給所有關心我的人，在此致上最深的謝意。

2010.10 萬鴻政

章節目錄

中文摘要	i
英文摘要	ii
誌謝	iv
章節目錄	v
圖目錄	vii
表目錄	x

第一章 簡介 1

1-1 影像縫合介紹	1
1-2 鬼影的產生及改善方式	1
1-3 特徵點匹配的準確性	2
1-4 章節規劃	3

第二章 背景 4

2-1 影像縫合流程	4
2-2 特徵點擷取	5
2-2.1 特徵點定義	5
2-2.2 一般 Harris Corner 取法	6
2-2.3 適應性非最大化壓縮法	7
2-3 特徵點匹配-- 次近鄰居法	11
2-4 剔除掉錯誤的匹配點	14
2-4.1 隨機取樣程序	14
2-4.2 極化型隨機取樣程序流程	17
2-5 歪斜改善	21
2-6 影像混合	22
2-7 環場影像縫合	24

第三章 內向式和外擴式影像縫合之探討 27

3-1 內向式和外擴式像素差距比較	27
3-1.1 像素差距	27
3-1.2 相機視角比較	34
3-1.3 相機焦點距離	36

3-2 內相式和外擴式相機影像重疊區域-----	38
3-3 內向式和外擴式相機縫合結果比較-----	42
第四章 強化特徵點匹配正確性之技術 -----	47
4-1 特徵點匹配的缺點 -----	47
4-2 改良作法 -----	48
4-2.1 相關性-----	48
4-2.2 像素值排列-----	49
4-2.3 邊緣判斷-----	50
4-3 模擬結果與匹配方法比較 -----	51
第五章 結論與未來發展 -----	55
參考文獻： -----	56



圖目錄

圖 1	鬼影範例圖	2
圖 2	明暗不一的兩張影像	2
圖 3	絕對座標縫合範例	4
圖 4	特徵點示意圖	6
圖 5	ANMS 流程圖	7
圖 6	適應性非最大化壓縮法 (ANMS) 範例	9
圖 7	Search Range 從無限大往內縮回時的特徵點擷取數量累積圖	10
圖 8	Bottom-Up Maxima Selection (BUMS) 流程圖	11
圖 9	特徵點擷取與特徵點匹配的範例	12
圖 10	次近鄰居法 (Second-Closest Neighbor) 的流程	13
圖 11	次近鄰居法 (Second-Closest Neighbor) 的成功機率圖表	13
圖 12	RANSAC 範例	16
圖 13	RANSAC (RANDOM SAMPLE CONSENSUS) 流程圖	16
圖 14	極化型隨機取樣程序 (PRSC) 流程圖	19
圖 15	PRSC 臨界值取高標的結果	20
圖 16	歪斜影像範例	22
圖 17	Weight Function 示意圖	23
圖 18	兩張圖的相對位置與權重函數關係	23
圖 19	經過 Blending Strategy 所產生的縫合影像	24
圖 20	環場影像縫和流程圖	25
圖 21	圓柱投影水平方向	25
圖 22	圓柱投影鉛直方向	25
圖 23	圓柱投影範例	26
圖 24	環場影像縫合範例	26
圖 25	鬼影範例圖	27
圖 26	外擴式相機擺法	28
圖 27	內向式相機擺法	28
圖 28	外擴式相機成像	29
圖 29	成像原理	30
圖 30	像素點匹配	30
圖 31	水平方向簡化圖	31

圖 32	垂直方向簡化圖-----	32
圖 33	水平垂直方向合併圖-----	33
圖 34	相機焦點距離和 disparity 關係圖-----	34
圖 35	內向式相機視角-----	35
圖 36	外擴式相機視角-----	35
圖 37	(a)內相式相機 -----	36
	(b)外擴式相機-----	36
圖 38	外擴式相機放大圖-----	37
圖 39	內向式相機重疊區域-----	38
圖 40	外擴式相機重疊區域-----	40
圖 41	重疊比例和距離的關係圖-----	42
圖 42	內向式相機拍攝結果及重疊區域 -----	43
圖 43	外擴式相機拍攝結果及重疊區域 -----	43
圖 44	內相式影像絕對座標縫合結果 -----	44
圖 45	外擴式影像絕對座標縫合結果 -----	44
圖 46	內相式影像圓柱投影縫合結果 -----	44
圖 47	外擴式影像圓柱投影縫合結果 -----	45
圖 48	外擴式影像 -----	45
圖 49	內向式影像 -----	45
圖 50	外擴式影像圓柱投影縫合結果 -----	46
圖 51	內向式影像圓柱投影縫合結果 -----	46
圖 52	影像縫合流程圖-----	47
圖 53	遠近拍攝影像明暗差異-----	48
圖 54	像素值排列範例一-----	49
圖 55	像素值排列範例二-----	50
圖 56	8*8 範圍的像素值示意圖-----	50
圖 57	縫合結果圖-----	51
圖 58	pattern 三組-----	52
圖 59	絕對座標縫合結果-----	53
圖 60	圓柱投影縫合結果-----	53

表目錄

表一	SR 從 $r=1$ 往外擴張時特徵點擷取數量與剔除數量統計表範例-----	10
表二	匹配點經過 5000 次 PRSC 投票結果之關係表-----	17
表三	執行 10 次 PRSC 低標和均標的成功機率比較-----	21
表四	內向式和外擴式影像重疊區域比較-----	42
表五	絕對座標縫合三種方法速度比較-----	54
表六	圓柱投影縫合三種方法速度比較-----	54



第一章 簡介

1-1 影像縫合

隨著相機校正和影像處理的發展，影像縫合(AutoStitch)近幾年在影像處理中已經佔有越來越重要的地位，其可應用的範圍也越來越廣，除了攝影愛好者原本受限於相機角度有限而無法拍攝廣角影像，因為 AutoStitch 的發展而一償宿願，其他還有如 Camera Motion[1]-[3]，Scene Motion[4]-[6]，Advanced Camera Modelling[7]，Photometric Modelling[8]-[10]，監視系統、軍事系統方面，AutoStitch 的應用也是越來越趨於廣泛。

在早期 AutoStitch 或 Camera Calibration 技術尚未成熟的時候，所有的影像要縫合或是校正，都需要人工徒手靠著繪圖軟體來接合或手動修正，這種人工方法相當麻煩，而且效果不盡理想。直到在相機校正領域，Chen and Ouhyoung，Bhat and Nayar 提出了用相機相對位置來做校正[40][41]，Szeliski, etc 提出了利用 direct method 來做影像縫合校正[11]-[14]，此技術才突破了人工徒手校正的瓶頸，Capel, etc 之後更進一步提出了 feature-based method[3], [15]-[16]，影像縫合與校正技術才慢慢趨近於成熟，feature-based method 比起 direct method，運算量更加節省，並且拒絕 outlier 的能力更強，其校正結果自然誤差更小更加可靠。

1-2 鬼影的產生及改善方式

現行的縫合演算法，只要是多鏡頭拍攝都不可避免的會出現鬼影，鬼影產生的原因可能是物體移動或是多鏡頭影像所產生的視差，本篇論文主要探討的是多鏡頭影像產生的視差，如圖 1 是雙鏡頭影像縫合時所產生的鬼影範例圖，圖中可以看出來紅色牛頭旁邊多了一些殘影，很多人會利用演算法來嘗試改善鬼影的產生，像是形態變形法(Morphological Interpolations) [45][46]或是等位函數變形法(Level Set Methods)[47]-[49]，但是本篇論文主要想探討的是外部相機的擺設方式，對於鬼影的產生也有很大的影響。當我們使用內向式和外擴式的兩種擺法，外擴式的雙鏡頭影像縫合出來的效果，會有很嚴重的鬼影，尤其是當影像中物體的分佈有遠有近時，鬼影會更明顯，而如果使用內向式雙鏡頭，鬼影的效果會大幅降低，這一方面我們可以用一些數學式子來證明為什麼會有這樣的差異。除了改善鬼影之外，內向式雙鏡頭還有一個優點，在同樣視角下所得到的

重疊區域會比較多，因此有利於縫合時的特徵點匹配，因為有較多的特徵點所找到的轉換模型也會比較準確。換句話說，在同樣的重疊區域底下，內向式雙鏡頭可以得到更大的視角。



圖 1 鬼影範例圖

1-3 特徵點匹配的準確性

在[38]中提到了兩種特徵點匹配的方式，一種是最簡單的最小平方差 SSD (sum of squared differences)，另一種是改良過的次近鄰居法(Second-Closest Neighbor)[28]，兩種方法都是為了把兩張圖的特徵點作匹配的動作，次近鄰居法雖然是改良過的，但他改善的地方在於當圖中有很多相似的像素點時，他能夠把錯誤的 match 給剔除，可是卻沒有考慮到當兩張影像的明暗不同時，單純做 SSD 的話，兩個原本是匹配的点，會因為像素值的差異而使得算出來的 SSD 很大，因而造成程式的誤判，如圖 2 中，左邊的影像明顯的比右邊的圖來的明亮，因此兩邊的向素質會差蠻多的。



圖 2 明暗不一的兩張影像

所以我們找了幾種方法來解決這個問題，第一種是相關性(correlation)[44]，相關性這方法很多人都在用，但雖然使用 correlation 可以得到不錯的效果，但是他的運算量實在是很大。第二種是像素值排列(rank order)[41]-[43]，他是利用相素質大小所排列出來的順序來作運算，他的速度比 correlation 快一點，但是效果並不是很好，因為像素值的順序並不是那麼的具有指標性。第三種是邊緣判斷(edge judgment)，首先將影像經過高頻濾波，我們使用的是 Laplacian of Gaussian filter(LOG)，經過高頻濾波的影像只會剩下高頻訊號，也就是角落(corner)或是邊緣(edge)，這些位置剛好很可能是我們特徵點的位置，因此經過高頻濾波之後特徵點附近的 edge 和 corner 都會留下來，再把他們拿去做 SSD 就可以得到正確的匹配，這個方法效果不錯而且速度快。

1-4 章節規畫

在本章節中介紹了一些縫合演算法的基本概念，還有本論文所要探討的縫合演算法的一些基本介紹。

第二章則是把縫合演算法做詳盡的描述，我們把縫合演算法中的每一個步驟都拆開來做完整的介紹。

第三章介紹的是雙鏡頭的擺設方式所帶來的差異，不同的擺設方式所照出來的影像，對於縫合的結果有很大的影響。

第四章是強化特徵點技術之探討，對於特徵點匹配的演算法提出幾種改善方法以及這些改善方法的比較和討論。

最後第五章把整篇論文做一個總結。

第二章 背景

2-1 影像縫合流程

影像縫合方法有很多種，本篇論文介紹兩種方法，一種是絕對座標的縫合 (Absolute Coordinated Image Stitching)[21][22]，另一種是環場影像縫合 (Panorama Image Stitching) [17]-[20]。絕對座標的縫合希望的是縫出來的圖能夠跟真實世界所看到的一樣，就是能夠完全符合人眼所看過去的角度，因此影像中物體的相對位置也是不能改變。圖2-1.1是絕對座標縫合的圖，我們可以看到為了要符合座標系，必須將影像做歪斜之後縫合，這是為了達到相對位置一樣的效果，但是這種縫合方式最多只能達到接近180度，因為當縫合的範圍越大，越旁邊的圖會因為線性放大而被拉得很開，圖3的右邊就已經有這種現象，所以絕對座標縫合不適合拿來縫太大範圍的影像。



圖 3 絕對座標縫合範例

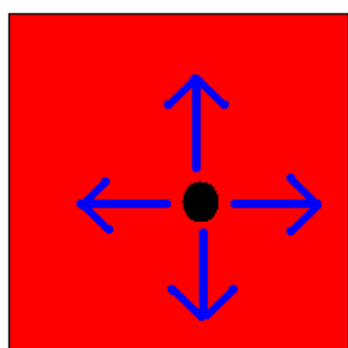
在影像縫合(AutoStitch)的過程中，大致上可分成幾個步驟，第一，特徵點擷取 (Feature Extraction)。為的是要讓在影像縫合的剛開始，並不需使用影像中的每個像素(pixel)來做運算，只需要找出幾個夠特殊的點(即特徵點)來做代表即可。第二，特徵點匹配 (Feature Matching)。在初步擷取特徵點過後，我們也需要做初步的特徵點匹配，以確認兩張圖當中的特徵點哪些是彼此對應到的。第三，剔除掉錯誤的匹配點。承接第二步驟，因為在初步特徵點匹配的過程結束之後，仍然有可能會存在錯誤的匹配點，因此需要建立一套機制使得錯誤的特徵點可以被剔除掉。第四，歪斜影像 (Warp Image)。不同需求的影像縫合在

效能上的判定準則也不一樣，我們需要把影像做歪斜使其可以跟另一張影像正確地縫合。第五，混合影像(Blending)。在把影像歪斜過後，最後一步就是要做適當地混合，不同的混合方法所呈現出來的最後縫合影像結果也不盡相同。接下來我們將針對上列影像縫合的基本步驟做一個完整地背景介紹。

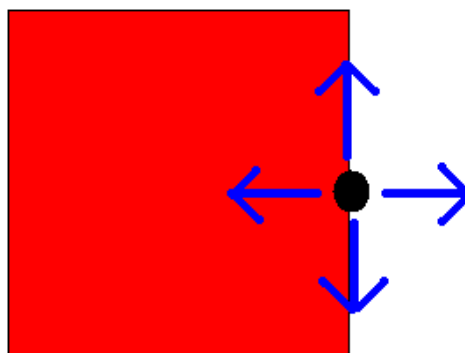
2-2 特徵點擷取

2-2.1 特徵點定義

在早期所有的影像要縫合或是修正，都需要人工徒手靠著繪圖軟體來接合或手動修正，這種人工方法相當麻煩，而且效果不盡理想。直到 Szeliski, etc 提出了利用 direct method 來做影像縫合與校正[11]-[14]，此技術才突破了人工徒手校正的瓶頸，之後 Capel, etc 更進一步提出了 feature-based method [3], [15]-[16]，即不需用整張影像的像素(pixel)來做運算，只需要找出幾個夠特殊的點(即特徵點)來做代表即可，Harris, etc 分別提出了許多不同的方法 [23]-[32]，其中 Derpanis, Harris Corner Detector 的作法[26]，是把影像當中的角落(Corner)擷取出來。首先我們需要先知道為什麼需要找角落來當特徵點，若我們使用影像中比較平坦(flat)的點來當特徵點，則往上下左右稍微移動一點點，所看到的東西仍會很像，所以這不是一個適合的特徵點。接著，若我們使用影像中邊界(edge)的點來當特徵點，則此點往某兩個方向稍微移動一點點，所看到的東西仍會很像，仍然不是一個適合的特徵點。最後，若我們使用影像中的角落(Corner)點來當特徵點，則此點往上下左右稍微移動一點點，所看到的東西仍會很不一樣，因此我們可以得知角落點(Corner)為我們所希望找到的特徵點，圖 4 即是一個最簡單的例子。



(a)



(b)

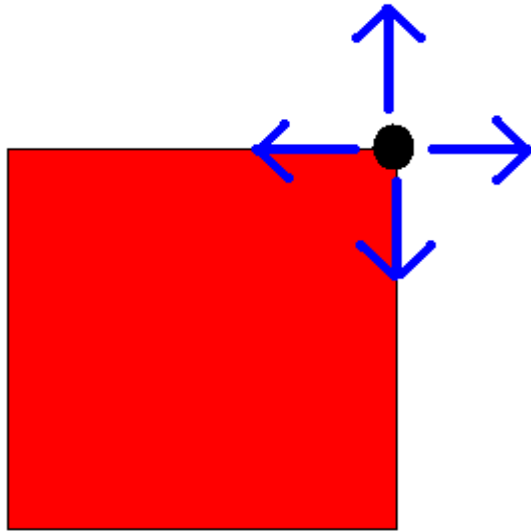


圖 4 特徵點示意圖

(a)平坦(Flat)

(b)邊界(Edge)

(c)角落點(Corner)

(c)

2-2.2 一般 Harris Corner 取法

接著我們要來討論如何找出影像的角落點以作為特徵點，Derpanis[26]首先把每個輸入影像的點拿來做計算， $I_x, I_y, I_x \times I_y, I_x^2, I_y^2$ 表示我們去計算每點在不同方向的微分，以及他們彼此相乘和各自的平方， w 為高斯函數，對於每個影像中的 Pixel，算出運算式 2-1 中矩陣的特徵值

$$A = \begin{bmatrix} (I_x^2) \otimes w & (I_x * I_y) \otimes w \\ (I_x * I_y) \otimes w & (I_y^2) \otimes w \end{bmatrix} \text{-----(1)}$$

當式(1)矩陣特徵值計算出來後，我們去觀察計算出來的特徵值(EigenValue):

- (1) 如果這兩個特徵值都很小，表示此點與附近的關係為平坦的(flat)，不是我們想要的情況(我們想要的為角落)。
- (2) 如果兩個特徵值其中一個很大，另一個很小，表示此點與附近的關係為邊界(Edge)，也不是我們想要的情況(我們想要的為角落)。
- (3) 如果兩個特徵值都很大，表示此點與附近的關係為角落(Corner)。

會有上述的特質是因為我們都知道矩陣的特徵值代表它梯度(Gradient)的大小，所以當特徵值越大表示梯度越大，但是當只有一個特徵值很大，另一個很小時，表示只有一個方向的梯度(Gradient)很大，另一個很小，很明顯地此種狀況為邊界(Edge)。相對地，當兩個特徵值都很大時，表示兩個方向的梯度(Gradient)都很大，這就是我們所想要的角落點(Corner)。在此因為求特徵值

需要花費較多的運算樣，Harris and Stephens 提供一個代替的方法[23]，假設 λ_1 ， λ_2 為矩陣 A 的兩個特徵值，則我們只要去計算式(2)的反應函數(Response Function)即可知道此點是 Corner 的可能性有多大：

$$R = \lambda_1 \lambda_2 - k(\lambda_1 + \lambda_2)^2 = \det(A) - k \times \text{trace}^2(A) \text{-----}(2)$$

其中根據文獻參考的經驗法則，k 在 0.04~0.15 之間都可以有彈性的空間，不同於式(1)的運算方法，在計算式(2)當中，只須計算加減法以及乘法，因此運算量將會比式(1)減少許多。

2-2.3 傳統適應性非最大化壓縮法

Brown, etc 提出了適應性非最大化壓縮法 (Adaptive Non-Maximal Suppression) 簡稱 ANMS[27]，ANMS 即是希望提出一套強健的策略，使得特徵點可以被壓縮在一定的數量，並且所擷取出來的特徵點也可以平均分散在整張影像當中。流程圖如圖 5。

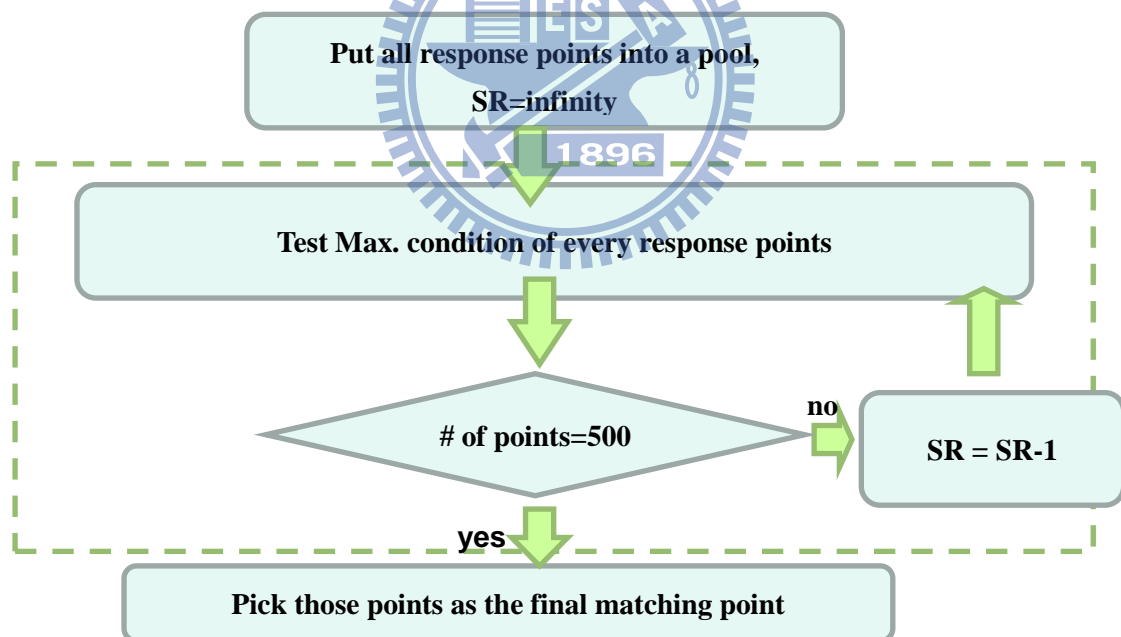


圖 5 ANMS 流程圖

讓我們先回顧一下在特徵點擷取 (Feature Extraction) 的步驟使用 ANMS 的優點，之所以會從 the highest corner strength 進展到 ANMS，主要的原因有兩個：第一、ANMS 所擷取的特徵點比起 the highest corner strength 可以平均分散在

整張影像當中，這是一個非常好的特性，因為在縫合兩張影像的過程當中，可能兩張影像的重疊部分(overlap)非常少，或是剛好沒有特徵點存在於兩張影像重疊的部份，則縫合的結果會大大的受到影響。圖 6(a)(b) 大部分的特徵點都集中在上半部份，下半部份特徵點非常少，因此若有另一張縫合影像的重疊部分(overlap)是與圖 6(a)的下半部分重疊，則這兩張圖的 Matching Points 會非常的少，因為下半部份存在的特徵點本來就很少，如此的情況非常不利於我們的影像縫合。但是同樣的圖若換成是圖 6(c)，因為所有的特徵點較為均勻分佈，因此若有另一張縫合影像的重疊部分(overlap)是與圖 6(c)的下半部分重疊也不會造成太大的影響，因為本身下半部份就有許多的特徵點可以提供之後作 Matching 和縫合的動作，因此能有像圖 6(c)(d)這樣均勻分佈而且特徵點數目又控制在所希望的大小是一個非常不錯的特性。第二、特徵點個數可以壓制在一定的數目以下。運算消耗量與特徵點的多寡是呈現超線性(superlinear)的關係，因此 ANMS 同樣也把特徵點個數壓制在 $n_i=500$ ，使得特徵點數目不至於過多進而拖垮整個演算法的運算效率。因為同時具備這兩個優點，ANMS 在擷取特徵點的使用相當普遍。

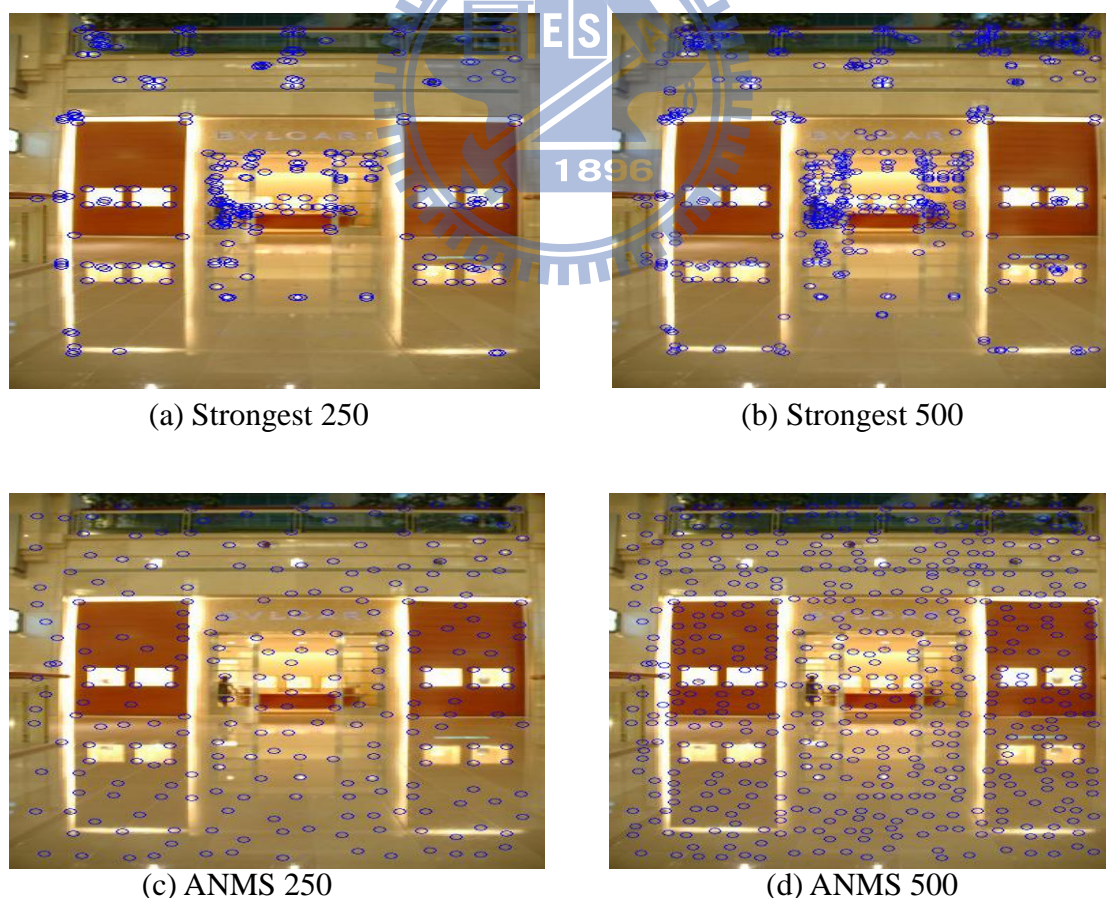


圖 6 適應性非最大化壓縮法(ANMS)範例

接著我們要從另一個角度來看適應性非最大化壓縮法 (ANMS) 的缺點，在前面我們有提到 ANMS 為了達到整張影像平均分布的效果，因此把反應函數 Response Function R 的搜尋範圍 r 從無限大往內縮回，當某個點為此 Search Range 的最大值時，就令此點為 ANMS 的特徵點。當 Search Range 從無限大往內縮回時，特徵點數目也會相對增加，當特徵點數目增加到我們想要的數目 n_i 時 ([27] 中令 $n_i=500$)，ANMS 才停止擷取特徵點。

但是在這裡因為 Search Range r 剛開始就從無限大搜尋，但是特徵點卻很少，之後慢慢往內縮，特徵點數目才慢慢增加，如圖 7。也就是說剛開始 Search Range r 很大時，運算量消耗非常多，但是能擷取出來的特徵點卻非常少 (因為 r 很大表示限制嚴格，能擷取進來的特徵點較少)，這不是一個好的現象，正常來說當演算法某部份消耗很大運算量時，則希望這部份能發揮較大的效果，但是 ANMS 這裡卻剛好相反，運算量消耗非常多，但是能擷取出來的特徵點卻非常少。並且根據統計，整個 AutoStitch 流程當中，特徵點擷取佔最大的運算量，也就是 ANMS 為了達到平均分佈的效果，使得 r 從無限大往內縮，但是在這裡卻因此佔了縫合影像流程當中最 Critical 的部份，拖垮了整個影像縫合演算法的效率。

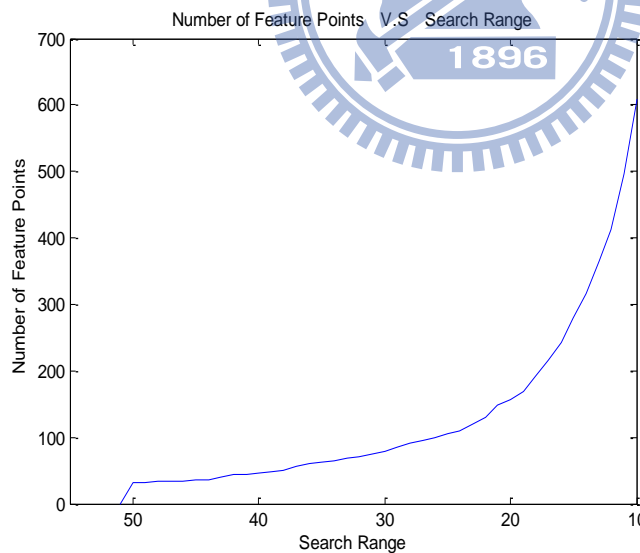


圖 7 Search Range 從無限大往內縮回時的特徵點擷取數量累積圖

SR	1	2	3	4	5	6	7	8	9	10
Points	4672	4450	3710	3035	2420	1795	1371	1098	873	716
delete	222	740	675	615	625	424	273	225	157	92
	11	12	13	14	15	16	17	18	19	20

	624	558	486	445	411	382	357	323	284	238
	66	72	41	34	29	25	34	39	46	38
	21	22	23	24	25	26	27	28	29	30
	200	174	156	140	126	114	109	105	98	92
	26	18	16	14	12	5	4	7	6	3
	31	32	33	34	35	36	37	38	39	40
	89	86	83	82	74	70	65	63	61	57
	3	3	1	8	4	5	2	2	4	4
	41	42								
	53	50	50							
	3	0								

表一 SR 從 $r=1$ 往外擴張時特徵點擷取數量與剔除數量統計表範例

在[38]中提出了 Bottom-Up Maxima Selection (BUMS)，做法是把 Search Range r 從 1 開始往外擴張，並且在擴張的過程當中只針對候選特徵點 (Candidate Points) 來做擴張，不是候選特徵點的就不重複做，另外我們也不使用 $n_i=500$ 做最後的特徵點，而是當某個 Search Range r 剔除掉的點小於 50，BUMS 即停止，會這樣做的原因是因為當特徵點剔除數量 < 50 時，整個特徵點已經趨近於收斂，如表一都可以清楚地觀察到此特性。例如當表一的 $r=13$ 時，特徵點剔除的數量只剩下 41 個，之後都越來越趨近飽和，因此若 r 繼續往外擴張，特徵點剔除的數量會越來越少，Search Range 變大但是能剔除的特徵點數量卻越來越少，不符合運算效率，因此我們定當特徵點剔除數量小於 50 時，BUMS 即停止，且如此的做法可以適應性地選擇特徵點的數量，若是特徵點數量小於 500 點，對之後的 Matching 也有很大的幫助，如圖 8。

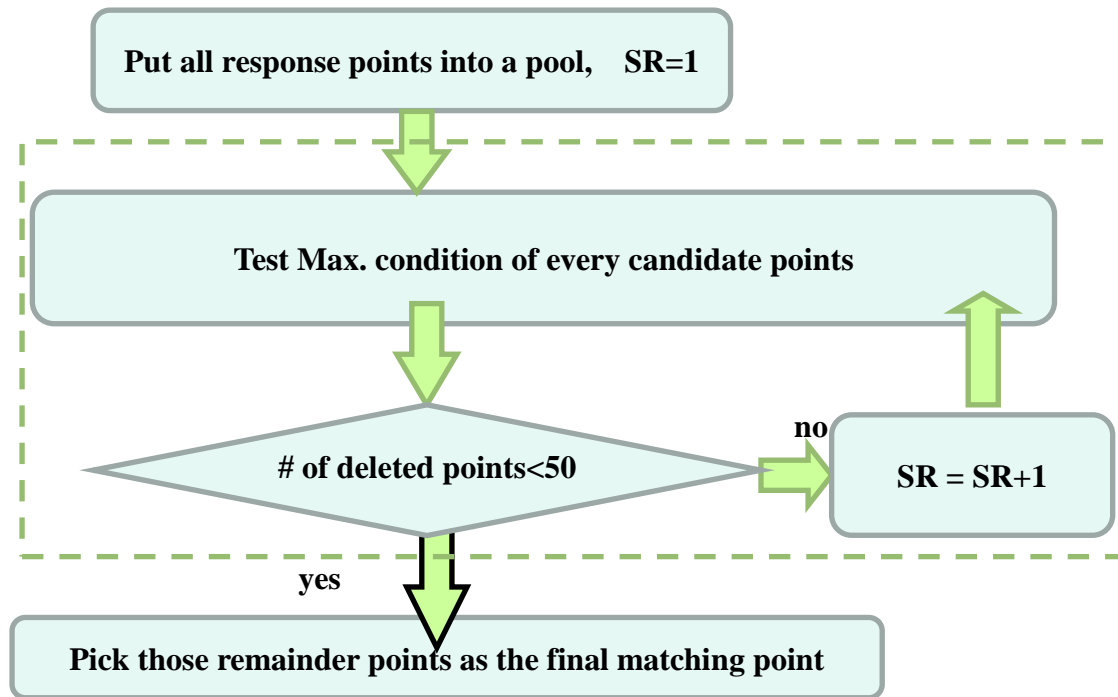


圖 8 Bottom-Up Maxima Selection(BUMS)流程圖

2-3 特徵點匹配-- 次近鄰居法

一般來說,SSD方法(或稱Closest Neighbor法)是一種最直觀最簡便的方法,但是在影像中的特徵點可能因為背景的模糊或是對應點根本不在所對應的影像當中,所以理論上來說應當不會產生好的Match在所對應的影像上,因此SSD並不是很好的方法,Lowe提出了次近鄰居法(Second-Closest Neighbor)來有效地剔除掉這種不該有好Match的特徵點[28]。

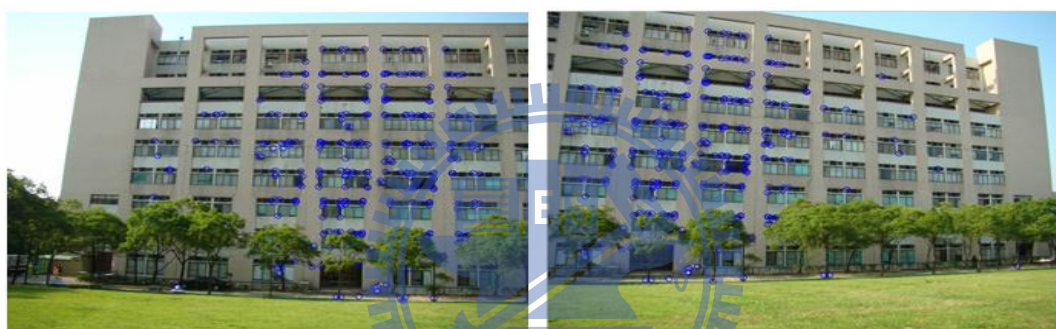
次近鄰居法(Second-Closest Neighbor)的構想是每一對好的匹配點(Matching Points)相對於其他對匹配點,都應該要夠獨特並且有區別性(discriminative),因此我們對於某個特徵點Matching時,除了計算最近的距離(Closest Neighbor) d_1 ,也去計算次近的距離(Second-Closest Neighbor) d_2 ,若是這兩個的比例 $D=d_1/d_2$ 大於 0.8,則表示此組的 Matching 太過類似(d_1/d_2 接近 1),則我們把這組 Match 剔除掉並稱他們為錯誤的 Matching Points。相反地若是 $D=d_1/d_2$ 小於 0.8,則表示此組的 Matching 夠獨特有區別性(discriminative),我們稱這兩個特徵點為一組好的 Matching Points。圖 10 為次近鄰居法(Second-Closest Neighbor)的流程圖。圖 11 為使用 40000 的特徵點(Feature points)來作次近鄰居法,水平軸為 d_1/d_2 的比例,垂直軸為 false match 的機率,由圖我們可以發現當 d_1/d_2 比例很低時,表示此組的 Matching 夠獨特有區別性,圖形也顯示 false match 的比例也非常低,但是當 d_1/d_2 比例很高時(大於 0.8),表示此組的 Matching 太過類似,圖形也顯示 false match

的比例也非常高，應該要剔除掉，[28]聲稱此種作法可替除掉 90%的 false match，但只會剔除掉 5%的 correct match，比起 SSD 作法，算是相當值得。圖 9 為次近鄰居法的範例，(a)-(b)為特徵點擷取，(c)-(d)為特徵點匹配，我們會發現(c)(d)的許多點都成功對應到了。



(a) 特徵點擷取

(b) 特徵點擷取



(c) 特徵點匹配

(d) 特徵點匹配

圖 9 特徵點擷取與特徵點匹配的範例

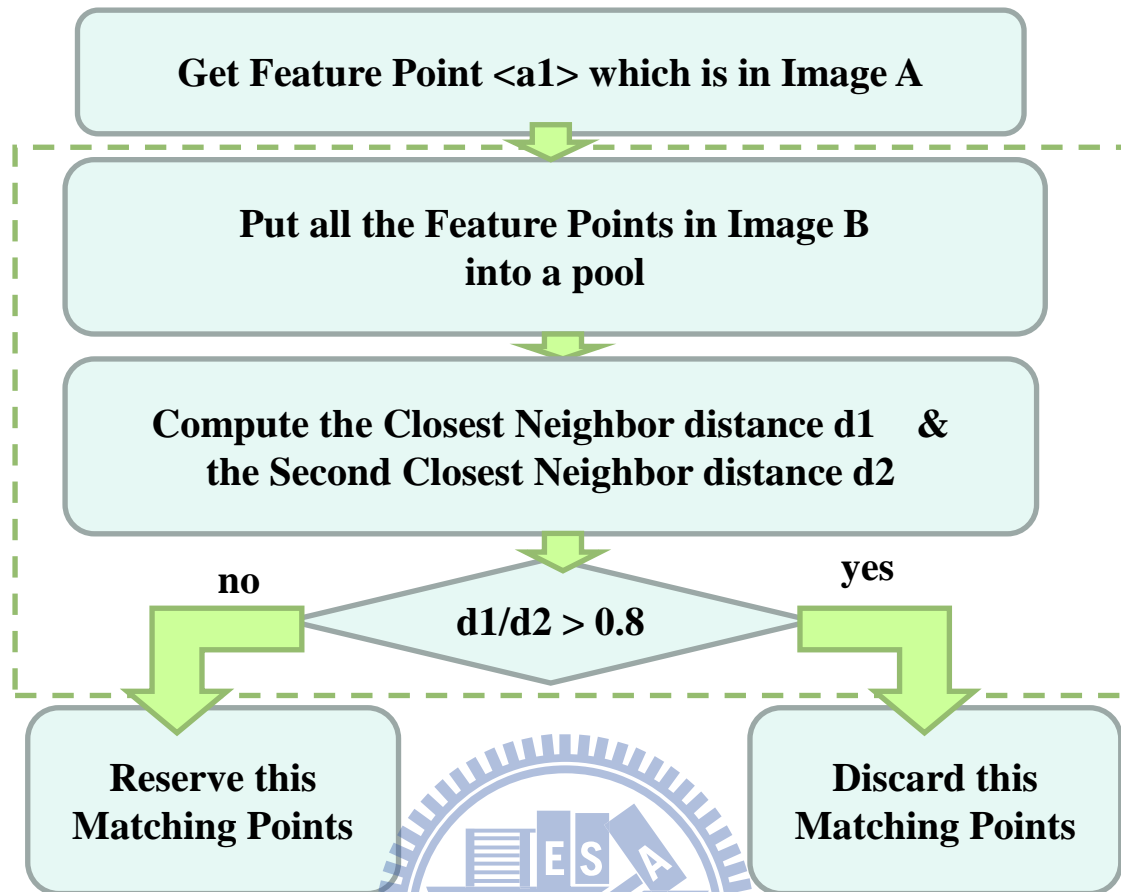


圖 10 次近鄰居法(Second-Closest Neighbor)的流程

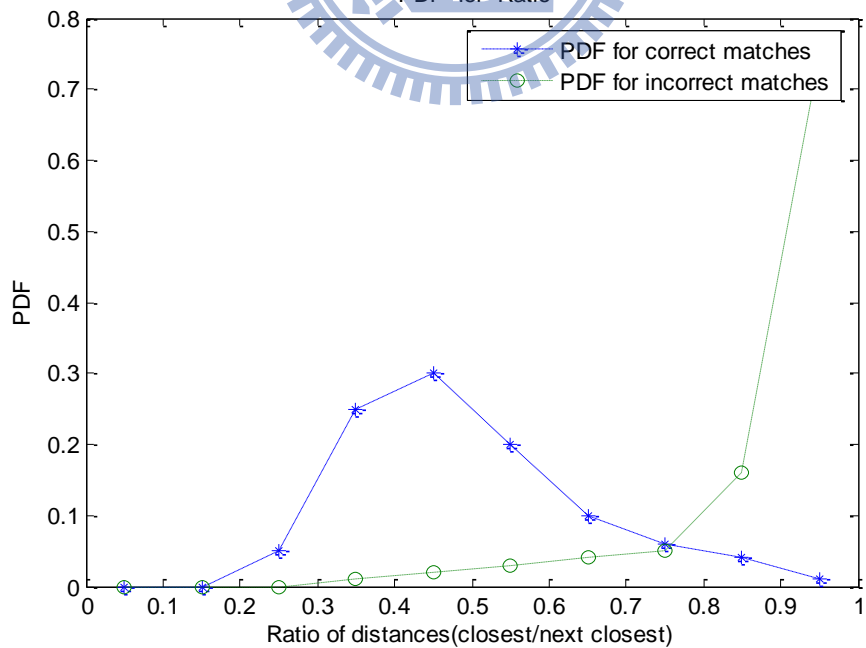


圖 11 次近鄰居法(Second-Closest Neighbor)的成功機率圖表

2-4 剔除掉錯誤的匹配點

2-4.1 隨機取樣程序

經過特徵點匹配(Feature Matching)後，我們會發現有許多的特徵點都彼此對應到了，而且大部分皆為正確的匹配(correct Match)，但是由圖 2-3.1 中我們也不難發現仍然有少數的特徵點彼此的對應是錯誤的(即 False Match)，所以在特徵點匹配之後的下一個動作，就是要把錯誤的匹配點剔除掉>Delete The False Match)，Fischler, etc 提出了幾種方法來建立一個正確的模型[33]-[34]，再利用此模型來剔除掉錯誤的資料(data)，這些方法不僅可以適用於影像上，也可以適用於許多統計資料的處理。而在[36]-[37]都使用[33]的作法來建立正確的模型以剔除掉錯誤的特徵點。Fischler and Bolles 使用隨機取樣程序(Random Sample Consensus)來建立模型剔除掉錯誤的資料，簡稱 RANSAC[33]。RANSAC(Random Sample Consensus)為一套使用隨機取樣來估計影像轉換參數的強健估計機制，在一般的例子裡，RANSAC 會隨機取 n 個 data 來建立一個模型，並且用剩餘的 $(N-n)$ 個 data 來測驗此模型是否正確，但是並非每次取的 n 個 data 都是好的 data，自然地經過測試過後會顯示不是一個好的模型，所以通常 RANSAC 都要執行一定的次數才能擷取出最後我們所要的正確 data。至於要重複多少次，則完全看使用者和問題的不同而定，當然，RANSAC 是在大部分的資料都是正確只有少部分資料是錯誤的情況下才能使用，但是在 2-3 我們已經經過初步的特徵點匹配(Feature Matching)來淘汰大部分錯誤的匹配點，因此剩下來的匹配點大部分都是正確的(由圖 2-3.1 我們也可以看到)，因此 RANSAC 在這個地方是絕對適用的，以下為 RANASC 步驟的流程：

- (1) 隨機地取 n 個樣本資料進來。
- (2) 使用這 n 個樣本資料來建立一個模型 Φ 。
- (3) 把剩下的 $(N - n)$ 個樣本資料代入 Φ 模型，測試此模型的建立是否正確。
- (4) 重複執行 K 次。
- (5) 把最大聯集的內部點 (inlier points) 取出來做最後的結果。

由上面的 RANSAC 流程我們並不難發現 RANSAC 流程當中有許多參數都是由使用者自己決定的，例如：如何定義樣本資料的數目 n 、如何定義重複執行次數、如何定義何為內部點 (inlier points)，這些都沒有一定的答案，而是根據我們遇到的問題來決定。[36]-[37] 當中使用樣本資料數目 $n=4$ ，重覆執行次數 $k=500$ ，定義內部點為 500 次當中最大對應成功次數的樣本資料，之所以會這樣定義的原

因，我們先看式(3)和式(4)。式(4)為兩個 2D 座標互相轉換的關係矩陣，式(4)為是式(3)的矩陣型式。

$$P' = HP \text{ ----- (3)}$$

$$\begin{bmatrix} wx' \\ wy' \\ w \end{bmatrix} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \text{ ----- (4)}$$

將式(4)做簡單的展開，如式(5)。

$$\begin{aligned} wx' &= ax + by + c \cdots (1) \\ wy' &= dx + ey + f \cdots (2) \\ w &= gx + hy + 1 \cdots (3) \end{aligned} \text{ --- (5)}$$

將式(5)的(1)-(3)×x' 和(2)-(3)×y' 得到式(6)

$$\begin{aligned} ax + by + c - gxx' - hyy' - x' &= 0 \\ dx + ey + f - gxy' - hyy' - y' &= 0 \end{aligned} \text{ ----- (6)}$$

再將式(6)展開成矩陣的型式得到式(7)，

$$\begin{bmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -x_1x_1' & -y_1x_1' & -x_1' \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -x_1y_1' & -y_1y_1' & -y_1' \\ x_2 & y_2 & 1 & 0 & 0 & 0 & -x_2x_2' & -y_2x_2' & -x_2' \\ 0 & 0 & 0 & x_2 & y_2 & 1 & -x_2y_2' & -y_2y_2' & -y_2' \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \\ g \\ h \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \text{ --- (7)}$$

注意式(7)是可以往下無限延伸的，由式(7)我們可以發現有 a~h 總共八個參數，每個樣本資料點可產生兩個方程式，所以至少需要 4 個樣本資料點以產生八個方程式，以提供我們解出這 8 個未知參數，這也就是為什麼 Brown, etc 使用的 RANSAC 每次取 4 個樣本資料點來建立模型 [36]-[37]，因為資料點取越多越容易取到錯

誤的點，以致於建立錯誤的模型，但是又必須取一定的點數來解答八個未知參數，所以使用 $n=4$ 。並且 Brown, etc 也用簡單的機率模型來粗略證明當重複執行次數 $k=500$ 時[36]-[37]，模型錯誤的機率已經小於 10^{-14} ，因此使用 $k=500$ ，最後在內部點(inlier points)的部分求出來之後，再把所有內部點代入式(7)並使用[35]direct linear transformation(DLT)以得到最後最精確的模型 H 。圖 12 為 RANSAC 的範例，藍色圈為經過初步 Matching 的結果，紅色又為經過 RANSAC 的結果，我們會發現所有 Match 都正確對應到了。圖 13 為 RANSAC 的流程圖。



圖 12 RANSAC 範例

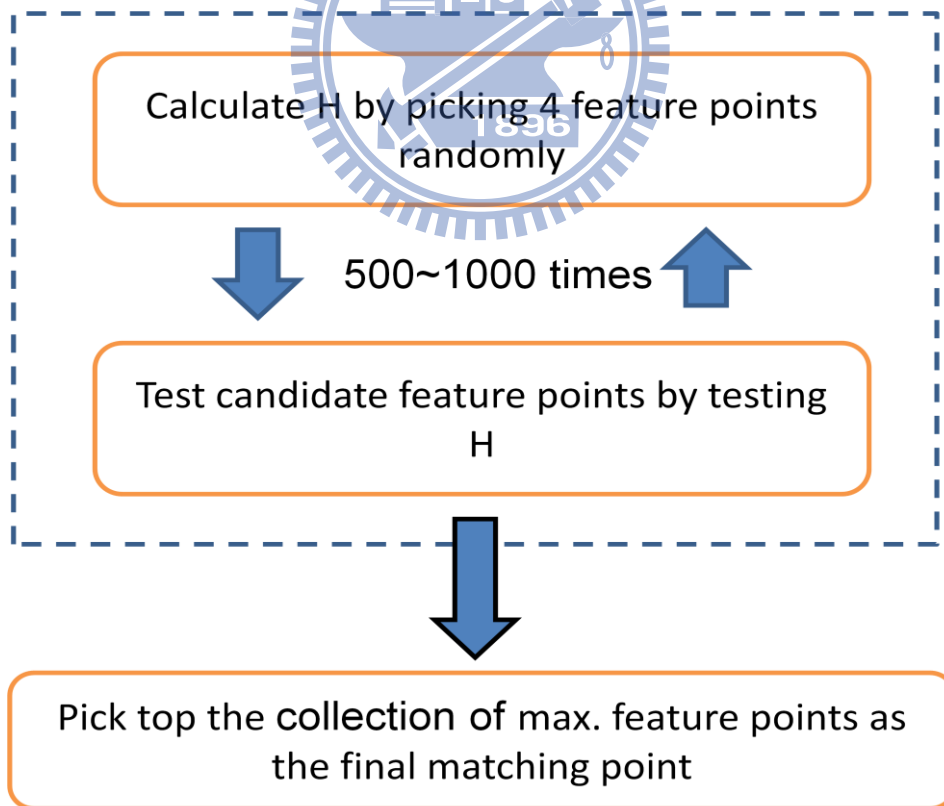


圖 13 RANSAC(RANdom SAMple Consensus)流程圖

2-4.2 極化型隨機取樣程序流程(The Flow of Polarized Random Sample Consensus)

[36]-[37] 當中使用樣本資料數目 $n=4$ ，重覆執行次數 $k=500$ ，定義內部點為500次當中最大對應成功次數的樣本資料。執行500次的原因最重要的是要讓最後擷取正確特徵點的機會提高，因此RANSAC次數 k 當然是越高越好，次數越低錯誤的機會越高。因此不同於[36]-[37]的做法，Chuang and Dung提出極化型隨機取樣程序(PRSC) [38]，PRSC首先使用投票機制(Voting)來做最後取內部點的判別，如表二，表二顯示直接做5000次RANSAC並使用投票機制的結果，如此做我們只是要先觀察使用Voting的好處與現象，表二左邊兩行Match X1與Match Y1為某組Pattern裡面其中一張影像的匹配點座標，Voting5000表示每個座標點經過5000次RANSAC之後每個點被Voting到的次數。從表二當中我們可以清楚地看到這24個匹配點經過5000次的RANSAC之後，有4個點被Voting到的次數非常少，有一個座標點甚至只被投票到22次，剩下20個座標點被Voting到的次數都非常高，最少都有1441次，也就是所有的匹配點呈現兩極化(Polarized)的現象。如果是正確的匹配點(Correct Match)，則經過一定次數的RANSAC之後被Voting到的次數會非常多，相對地，如果是錯誤的匹配點(False Match)，經過一定次數的RANSAC之後被Voting到的次數會非常少，這是一個非常自然也可以預知的現象，在式2-4中我們曾經說明歪斜影像的轉換關係，此運算式也是判斷投票與否的判定式子，當我們建立了一個正確的模式 H ，並把所有點套進 H 當中去測試此點是否為正確的匹配點，若測試成功就會在Voting上面加上1，若失敗則不加，因此正確的匹配點經過一定次數的RANSAC之後被Voting到的次數會非常多，錯誤的匹配點被Voting到的次數會非常少。之所以錯誤的匹配點仍有少數的Voting，原因在於若是我們建立的模型 H 是錯誤的，則錯誤

表二 匹配點經過 5000 次 PRSC 投票結果之關係表

Number	Match X1	Match_Y1	Voting5000
1	80	84	1638
2	81	84	1598
3	82	84	1590
4	238	101	254
5	106	134	1450
6	129	134	1443
7	130	134	1441
8	269	280	1878
9	269	281	2056
10	267	283	1861

11	228	289	1975
12	228	290	1991
13	336	301	22
14	200	311	2154
15	201	311	2183
16	302	317	412
17	301	318	438
18	193	336	2074
19	193	337	2053
20	193	338	2025
21	94	373	1549
22	91	379	1717
23	92	379	1733
24	94	380	1640

的匹配點也會被誤認會是正確的匹配點，因此仍然有少數的 Voting。當我們知道這個現象之後，我們就可以得到，當 RANSAC 次數到達一定程度，使得所有匹配點的 Voting 值呈現兩極化(Polarized)的現象(如表二的 Voting 兩極化)，RANSAC 即可停止，不需要執行 RANSAC 到 500 次才停止，因為若出現此情況，表示正確的匹配點已經幾乎知道，不需要再消耗運算量來做多餘的 RANSAC 了。因此我們在這裡定一個臨界值 0.3，當所有匹配點 Voting 的標準差已經大於最大 Voting 值的 0.3 倍，表示 Voting 值呈現兩極化的現象，RANSAC 即停止，不需要再消耗運算量來做多餘的 RANSAC 了。

除了投票機制之外，PRSC 還有許多修正的地方，傳統 RANSAC 演算法在每次隨機取樣的時候都從所有匹配點當中做取樣，但我們發現這樣不是一個好的做法，因為在很多時候已經知道哪些可能是正確的匹配點或錯誤的匹配點，因此應該可以訂出一套機制使得隨機取樣程序更精準。預設一個臨界值 0.3，當每做一次 RANSAC 結束後，若是對應到的匹配點數目達到所有匹配點數目的 0.3 倍，表示測試成功的樣本座標點很多，則我們稱這次的 RANSAC 為成功，而且下次的 RANSAC 從不再重新隨機取樣，而是從上一次的 RANSAC 當中測試成功的樣本座標點來做隨機取樣，如此的做法可以大大的提升擷取正確內部點的機率，比起單純做 Voting 標準差法還可以更節省 RANSAC 次數。圖 14 為極化型隨機取樣程序(PRSC)的流程圖。

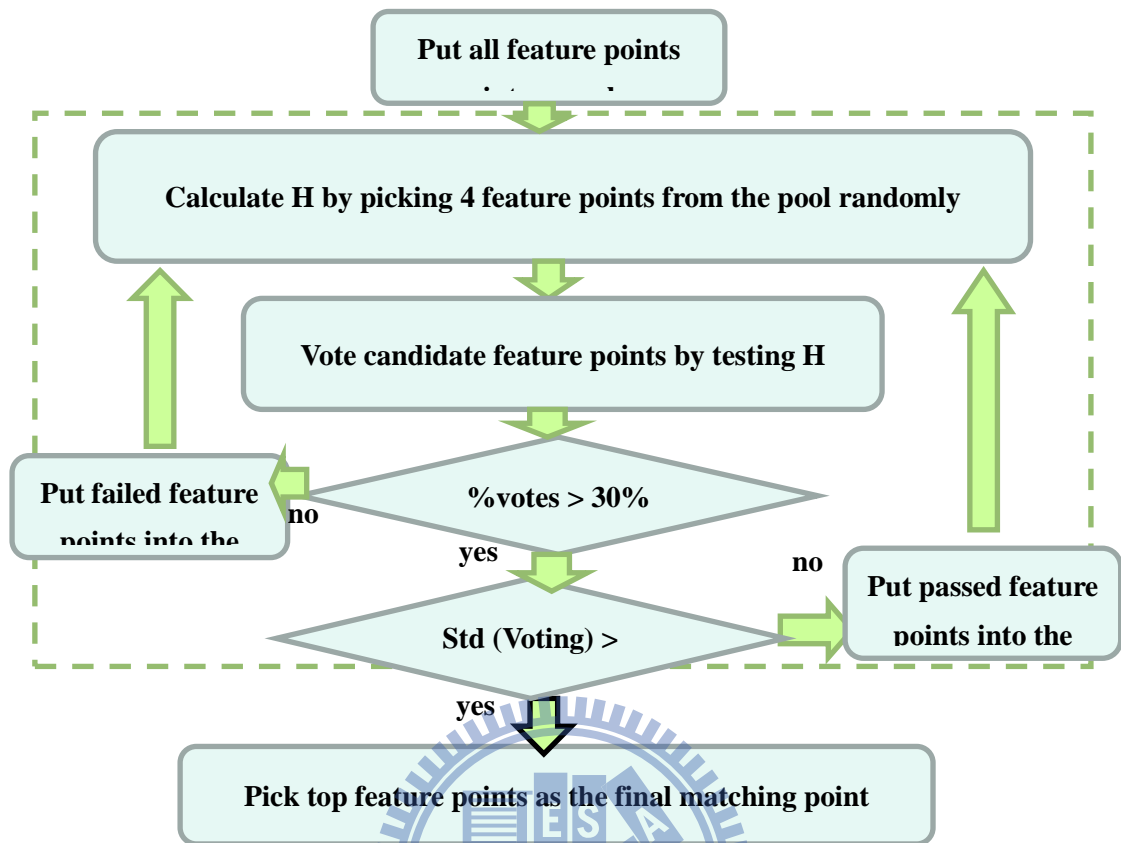


圖 14 極化型隨機取樣程序(PRSC)流程圖

PRSC 在最後階段擷取最後的正確匹配點或稱內部點(Inlier Points)時，我們是採取所有匹配點的低標來當臨界值，當 Voting 超過所有匹配點的低標值則稱為內部點(Inlier Points)，超過則視為外部點(Outlier Points)，傳統 RANSAC 定義內部點為 500 次當中最大對應成功次數的樣本點，因為它沒有做投票機制的動作所以無法取高低標或平均，但 PRSC 取低標當最後臨界值，原因是因為我們是採取適當的投票機制，幾次之後發現標準差拉開了，即 Voting 值呈現兩極化的現象，一組兩極化的資料，且大部分的資料值很大，少數的資料值很小，我們希望把大部分的資料值擷取出來，在統計上以低標最理想的臨界值。若是取高標當臨界值，雖然擷取出來的點都是 Voting 值很高的，表示非常可靠，但如此的取法卻容易使得特徵點數目不夠作之後的縫合，因為取高標會把多數正確的內部點也剔除掉，這不是好的現象，如圖 15，圖 15(a)(b)為臨界值取高標的匹配點與內部點結果，我們會發現取到的內部點非常少，大部分的正確內部點都因為臨界值取高標而被剔除掉了，因此圖 15(c)和圖 15(d)自然地也就因為內部點點數太少而失真了。若是我們取均標當臨界值，在大部分的情況仍然可行，但是在 PRSC 執行次數上比起低標當臨界值需要更多次的執行次數，



(a)(b)匹配點與內部點



(c) 歪斜影像

(d) 縫合影像

圖 15 PRSC 臨界值取高標的結果

原因與取高標當臨界值一樣，都是因為把部分正確的內部點剔除掉了，只是取均標比取高標強健，因為剔除掉的正確內部點沒這麼多，所以若是取均標仍然可行，但是相較於取後標當臨界值需要較多次的 PRSC 執行次數。若是取後標當臨界值，看似會比低標效果好，但是此種情況會有把外部點也誤認為內部點的機會，進而影響到了之後縫合的效果，因此取後標不是一個好的方法。表三證明在 PRSC 當中使用低標比均標成功的機率高，把 PRSC 大幅降低到只做 10 次並取觀察他們分

別的結果，由表三當中我們可以看到在執行 10 次 PRSC 時使用低標當臨界值比均標當臨界值成功機率高上許多，如此低的執行次數卻已經有 98% 的成功縫合機率，由此也可以間接證明傳統 RANSAC 演算法的執行次數事實上是不需要執行到 500 次的，因為若我們使用投票機制來做 RANSAC，當執行次數到了一定次數之後(即匹配點呈現兩極化時)，正確匹配點(Correct Match)或稱內部點(Inlier)的正確性就已經顯明出來，RANSAC 可以立即停止，但若沒有使用投票機制，正確匹配點(Correct Match)或稱內部點(Inlier)的正確性就無法顯明出來，只能大幅增加 RANSAC 次數以確保當次數增加到非常多，最後擷取到正確匹配點的機會才會萬無一失，但 PRSC 卻不需要如此做，因此從這裡我們終於可以一次清楚明白使用投票機制使得匹配點呈現兩極化現象與 PRSC 臨界值和取低標當最後門檻的原因。

	PRSC (average filtering)	PRSC (lower-average filtering)
#Correct trials	78	98
#False trials	22	2
Correctness	78%	98%

表三 執行 10 次 PRSC 低標和均標的成功機率比較

2-5 歪斜影像

在 2-4 經過 RANSAC 剔除掉錯誤的匹配點，同時也建立了正確的模型 H 之後，下一步就是要使影像做歪斜(Warp Image)，在前提是全自動絕對座標縫合影像(Automatically Absolute Coordinated Image Stithing)，希望是達到與真實世界(ground truth)內容一樣的結果，所以相對位置也必須一樣，假設我們在 Sensor 皆為水平移動的前提下，要縫合兩張影像一定要把其中一張影像做歪斜，因為他們彼此的關係不是水平移動就可縫合的，必定有某種投影的關係，我們稱為透視投影(Perspective Projection)，其關係是即為式 2-4，因此我們只要把需要做歪斜的影像座標經過式 2-4 即可得到新的歪斜影像，如圖 16、圖 2-5.1，左邊的影像經過 H 做歪斜即為右邊的影像。



圖 16 歪斜影像範例

2-6 混合影像

經過 2-5 歪斜影像的步驟後，兩張圖的相對位置已經吻合，最後一步就是要將兩張影像作混合(Blending)，理想情況若是在拍攝影像時，光源皆沒有任何改變，並且每一張圖的受光度都一樣(表示每張圖的明暗程度都一樣)，理論上應該是可以直接將兩張圖放在一起而不用作任何的動作，但事實上並非如此，在擷取影像的過程當中，即使是 multi-sensor 並且同步擷取影像，仍然會因為角度的不同而在明暗度上有些落差，若是沒有經過特別的混合策略就縫合，會造成影像的失真或是我們不希望看到的邊界層次的情況產生。我們可以發現因為影像受光度的不同，若是沒有聽過特殊的混合策略來縫合，所會造成的邊界層次失真情況產生，因此有一個好的混合策略也是非常重要的。

在此我們使用常見的線性混合(Linear Blending)來做縫合，[39]提出的線性混合(Linear Blending)理論使用距離的遠近為權重函數(Weight function)來解決邊界線不均勻的問題，使兩張影像在縫合時可以有平均的動作，圖 18 為其示意圖， H_l 為左邊影像的權重， H_r 為右邊影像的權重，越靠近本身影像的中心則權重越高，越靠近另一張影像則權重越低，圖 17 中間重疊的部份都介於 0~1 之間，是為方便之後作平均的動作，圖 18 為兩張圖分別對應的權重(Weighting)，我們會發現兩張權重圖是不一樣的，歪斜影像的權重圖也是歪斜的，原始影像的權重圖卻是長方形的，原因即是因為他們是分別由不同的圖產生的權重函數，當 Pixel 的距離離影像中心點越近則權重越大，距離越遠則權重越小，且所有權重皆介於 0~1 之間，因此產生了圖 18，我們也可發現兩張權重圖越往中心點越亮，此即表示他們的權重越大。

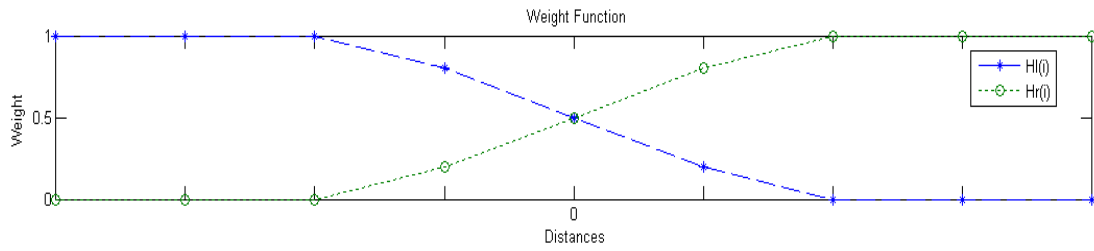


圖 17 Weight Function 示意圖

產生權重函數之後，最後只要利用簡單的平均關係即可將影像來作縫合，根據式(8)即可得到每個重疊地方的 Pixel 值，進而消除因為光線強度不同所造成不均勻失真的效果。

$$I_{linear} = \frac{\sum_{i=1}^n I^i * W^i}{\sum_{i=1}^n W^i} \quad (8)$$

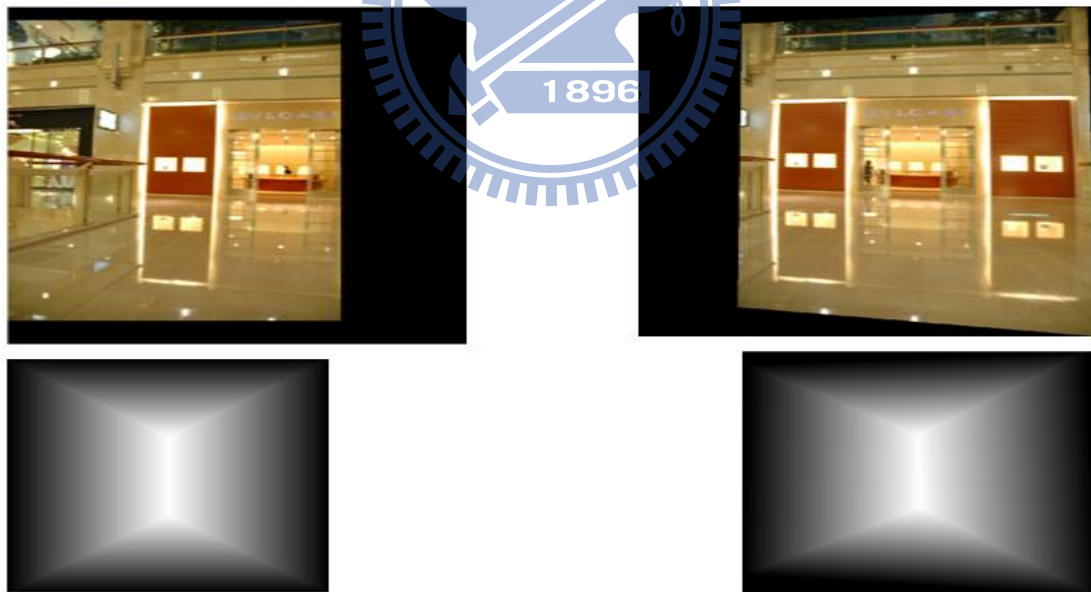


圖 18 兩張圖的相對位置與權重函數關係



圖 19 經過 Blending Strategy 所產生的縫合影像

圖 18 為兩張圖的相對位置與權重函數關係，由圖我們可以看到不僅兩張圖相對位置已經吻合，其權重函數(Weight Function)也已經確定，因此只要經過式 2-8 即可得到每個重疊地方的 Pixel 值，圖 19 為經過 Blending Strategy 所產生的最後縫合影像，我們可以明顯地看到它改善了失真效應。

2-7 環場影像縫合

如果要縫合 360 度的影像用的則是環場影像縫合，當然，要達到 360 度的視角，就必須扭曲物體來達到目的，原本的直線會被扭曲成曲線已達縫合的目的，在[36]-[40]中都有提到環場影像的縫合，因為不同角度所拍攝的照片，無法直接接合，因為兩支攝影機的視角不同，物體在圖像中大小或角度也勢必不同，然而跟絕對座標縫合不一樣的是，環場影像必須先將其投影至一個圓柱面（或球面上），投影後產生新影像的動作稱之為圓柱投影(Cylinder projection)。圖 20 是環場影像縫合的流程圖，他比絕對座標縫合多了一個動作，就是圓柱投影(Cylinder projection)。在此我們分別針對水平軸、垂直軸個別計算，從圖 21 的兩個式子(9)和(10)計算水平方向的座標投影關係，式(11)的式子是計算垂直方向的座標投影關係，圖 22，從相似三角形來看， y 比 y' 等於 $\sqrt{x^2 + f^2}$ 比 f ，

因此可以得到式(10)。

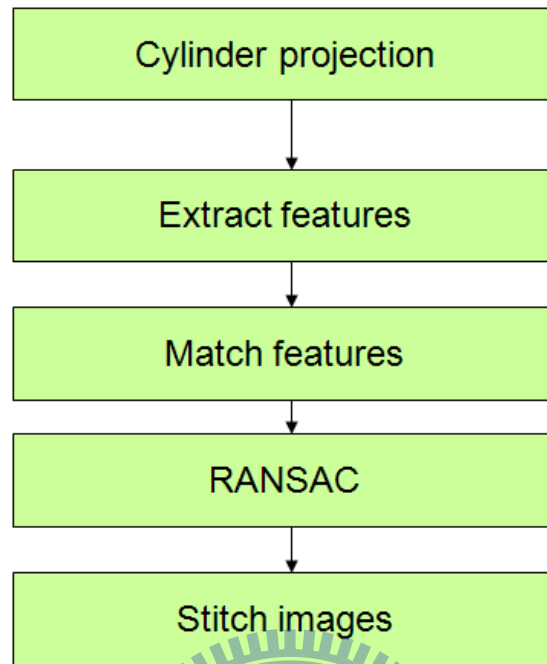
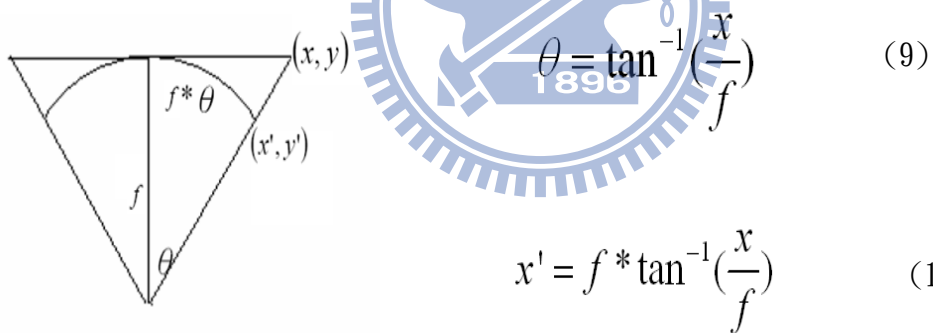


圖20 環場影像縫和流程圖



$$x' = f * \tan^{-1}\left(\frac{x}{f}\right) \quad (10)$$

圖21 圓柱投影水平方向

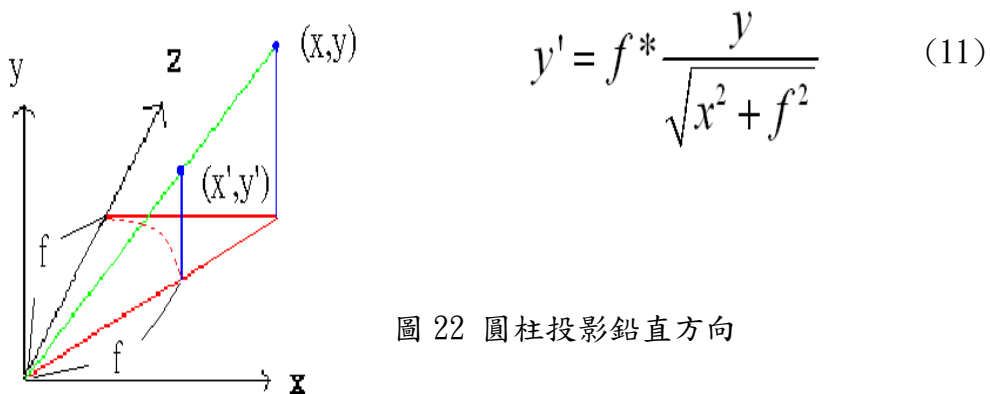


圖 22 圓柱投影鉛直方向

圖 23 中，(a)是原始圖，(b)是經過圓柱投影之後的圖，我們可以看得出來，圖形變得有點類似圓柱的形狀，這是為了要能夠把 360 度的景象都接在同一張圖，所以把 360 度的影像都投影到圓柱上之後，就能夠順利接出環場影像。

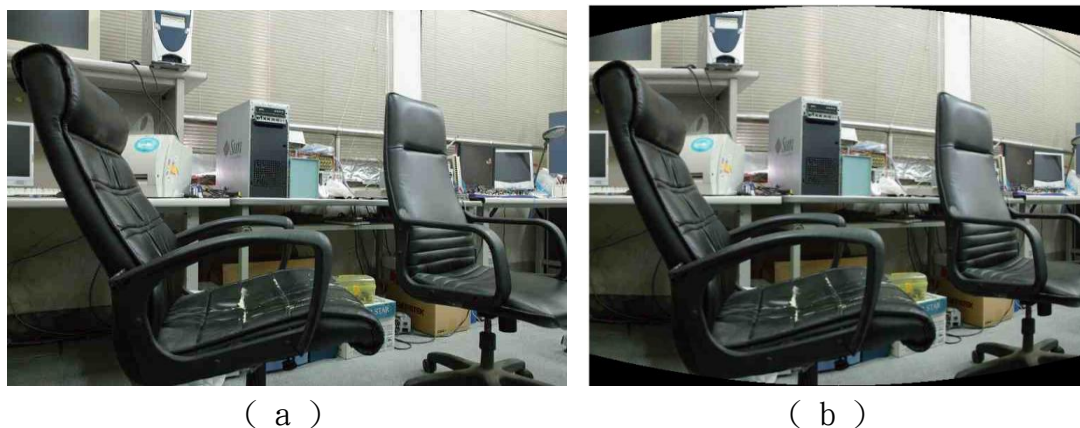


圖 23 圓柱投影範例

圖 24 是環場影像縫合的範例，可以明顯的看出桌子都變成彎的，和圖 3 中桌子還是直的完全不一樣，雖然是失真的影像，但是他卻可以比絕對座標縫合得到更大的視野，而且也不會有線性放大的問題。



圖 24 環場影像縫合範例

第三章 內向式和外擴式影像縫合之探討

在第二章介紹了一些關於影像縫合的基本理論之後，接下來本篇論文所要談的，是兩種雙鏡頭擺法，內向式和外擴式，所造成的差異以及優缺點。而這兩種擺設方式所照出來的影像都是使用第二章所介紹的演算法來實現，單純從外部相機所造成的影響來觀察。

3-1 內向式和外擴式像素差距(disparity)比較

3-1.1 像素差距

一般而言，只要是雙鏡頭以上所照出來的影像，在縫合時或多或少都會有鬼影產生，主要是因為遠近差異所造成的，當兩張影像重合的部分同時有遠近的影像時，你想要匹配(match)遠距離的景物的話，那麼近距離的景物就會有鬼影產生，相反地，當你想要匹配近距離的景物時，遠距離的影像就會有鬼影產生。所謂的鬼影，就是兩個匹配的像素點之間有差距(disparity)，如圖 25 中，因為在影像縫合時的 RANSAC，為了能符合大部份的像素點，所以近距離的紅色老鼠的像素點就無法匹配而產生差距，也就是鬼影。當像素值之間的差距(disparity)越大時，鬼影就越明顯。



圖 25 鬼影範例圖

接下來就是討論兩種擺設方式，對於鬼影所造成的影響。目前在做影像縫合

時，大多採取圖 26 的外擴式的相機擺法，而我們所要拿來比較的另一種擺試方式是圖 27 中的內向式相機擺試方式。

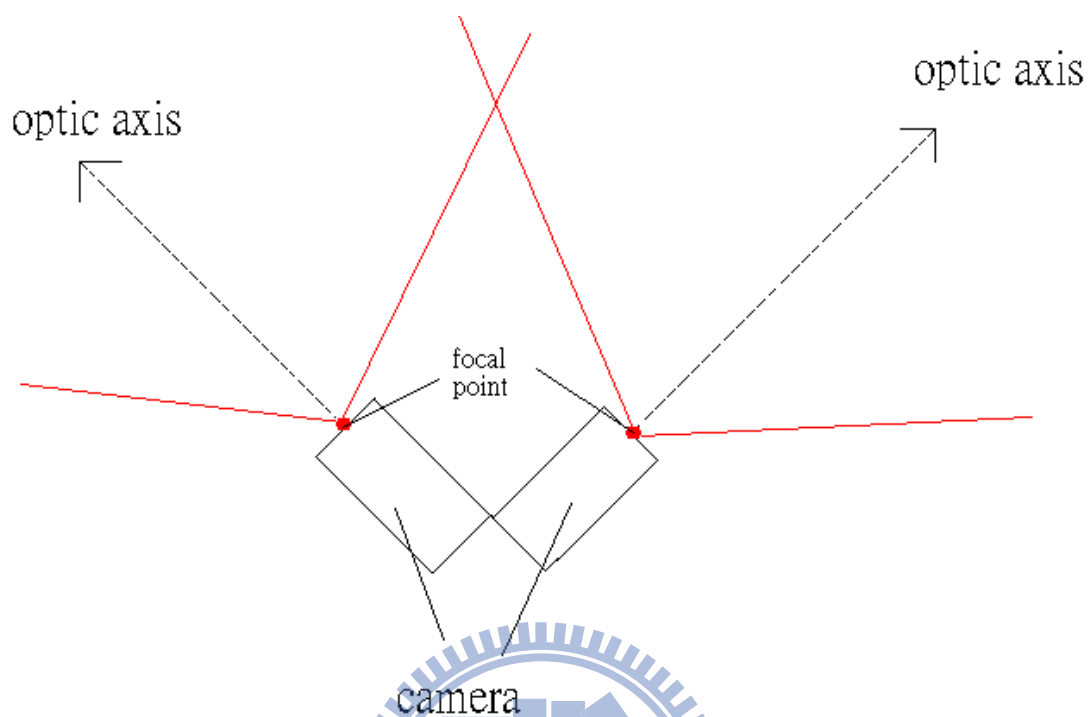


圖 26 外擴式相機擺法

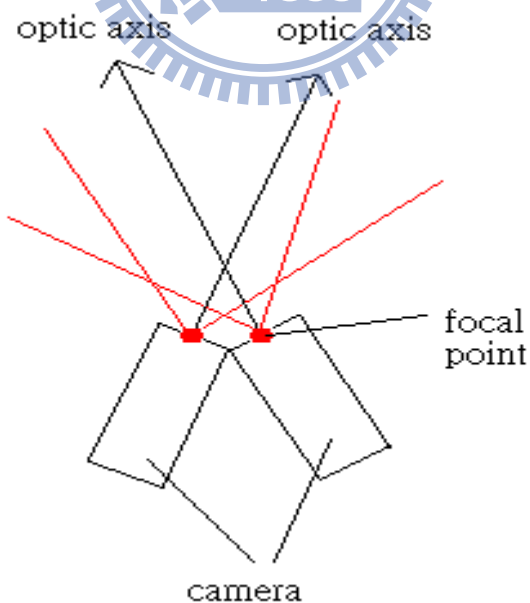


圖 27 內向式相機擺法

首先，我們從前面所談到的像素值的 disparity 來看看這兩種擺設方式所造

成的影響，從圖 26 和圖 27，我們把焦點(focal point)定義在相機前端，從兩台相機的焦點的水平距離來看看兩種擺法所造成影像上像素點差距的影響。圖 28 是外擴式相機成像， $P1$ 和 $P2$ 是現實世界中的任兩點，水平距離 $r1$ ，垂直距離 $r2$ ，而 $K1$ 、 $K2$ 、 $K3$ 和 $K4$ 分別是 $P1$ 、 $P2$ 在兩台相機上的呈像，原本應該成像在相機後方，但是根據圖 29 的成像原理，最右邊的紅線是物體，原本影像應該在焦點左邊，但是透過相似三角形，焦點左右兩邊的藍線是等長的，也就是我們可以把藍線移到等效的距離上。因此， $K1$ 、 $K2$ 、 $K3$ 和 $K4$ 就是等效的相機影像。 $d1$ 是物體和相機的距離。 f 是相機的焦距。 $O1$ 和 $O2$ 分別是兩台相機的焦點。 $\theta 1$ 和 $\theta 2$ 是物體對應到兩台相機的張角。 $x1$ 是相機焦點的距離。定義好這些參數就可以往下討論兩台相機的焦點的水平距離和影像上像素點 disparity 的關係。

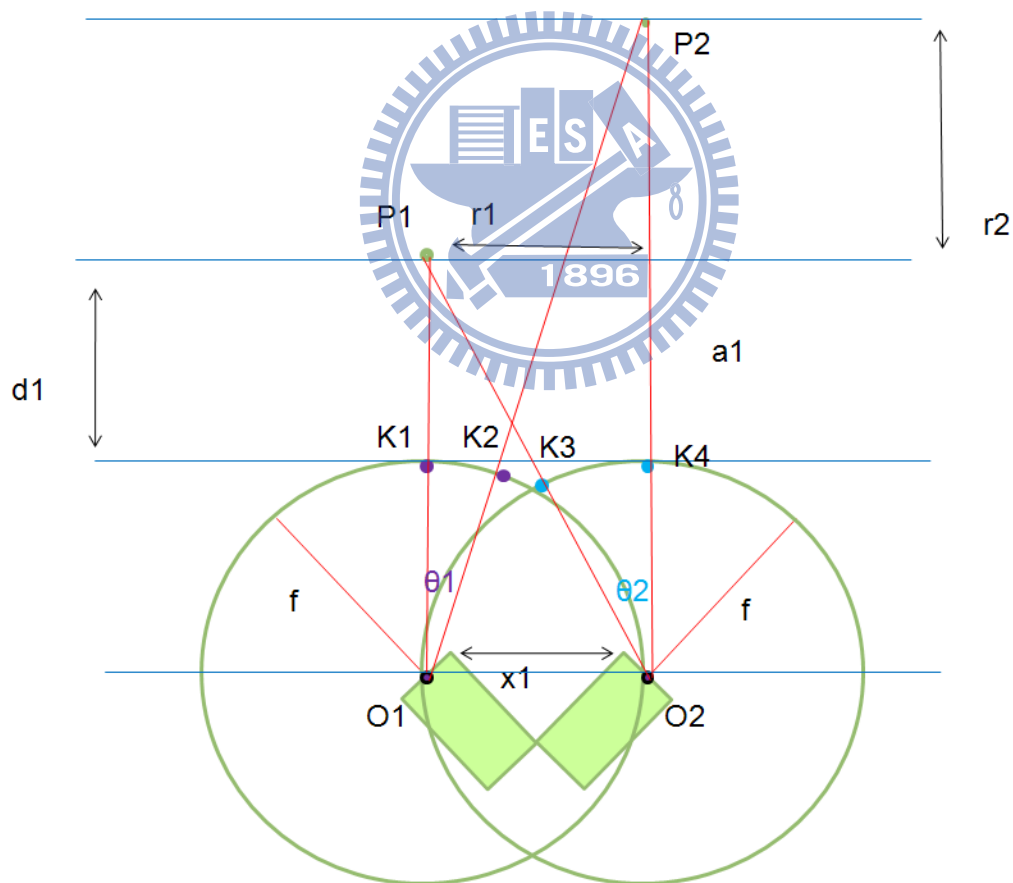


圖 28 外擴式相機成像

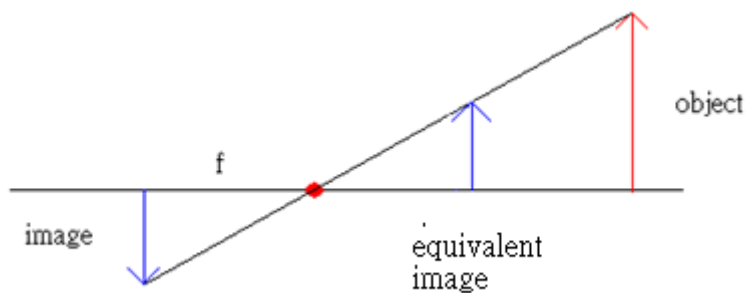


圖 29 成像原理

圖 30 是從圖 28 擷取出來，因為我們想要知道像素值的 disparity，知道相機的焦距 f 還有物體的夾角之後，我們就可以從圖 30 得到像素值的 disparity， $f * \theta$ 就是像素點在影像上的距離，在做縫合時，K1 和 K3 要匹配在一起，K2 和 K4 要匹配在一起，因此就是想要把 K1、K2 這一段 $f * \theta_1$ 和 K3、K4 這一段 $f * \theta_2$ 這兩段縫合在一起，但是兩段不一樣長，因此就產生 $disparity = f * \theta_1 - f * \theta_2$ 。所以如果我們想要把 disparity 縮小，就是要縮小 $f * \theta_1 - f * \theta_2$ ，但是 f 是固定的，因此能改善的就是 $\theta_1 - \theta_2$ ，也就是盡量讓 θ_1 和 θ_2 越接近，那麼 disparity 就會越小。那麼要如何讓 $\theta_1 - \theta_2$ 變小，主要影響的就是 X 的大小，相機焦點的距離。我們可以預想相機焦點越接近， θ_1 和 θ_2 會越接近，下面我們用數學式子來試著證明。

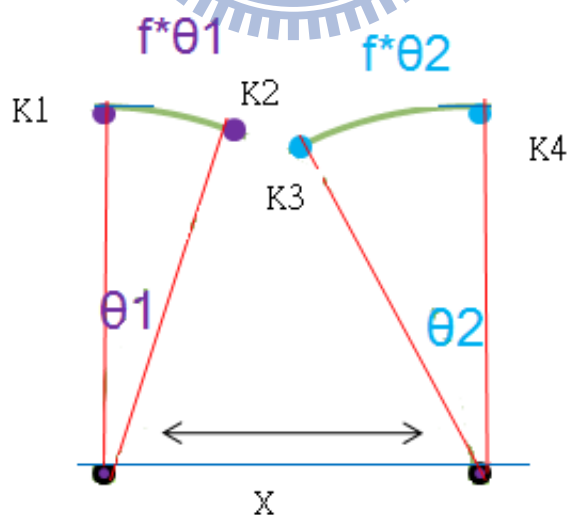


圖 30 像素點匹配

如果想從圖 28 直接得到 $\theta_1 - \theta_2$ 和 X 的關係式並不是那麼容易，所以我們把圖 28 簡化之後分成水平和垂直方向來討論，圖 31 是水平方向的示意圖， r_1

是水平方向的物體距離， c 是 2 號相機的焦點和右邊物體的距離， $O1$ 、 $O2$ 是相機焦點。原本我們想知道的是 $\theta 1$ 和 $\theta 2$ 的差距，但是這裡為了方便計算，改成探討 $\theta 3$ 和 $\theta 4$ 的差距，從三角形的原理可以知道 $\theta 1 + \theta 3 = \theta 2 + \theta 4$ ，所以 $\theta 1 - \theta 2 = \theta 4 - \theta 3$ ，因此我們可以從圖 31 推導得到 $\theta 3$ 、 $\theta 4$ 和 x 的關係。

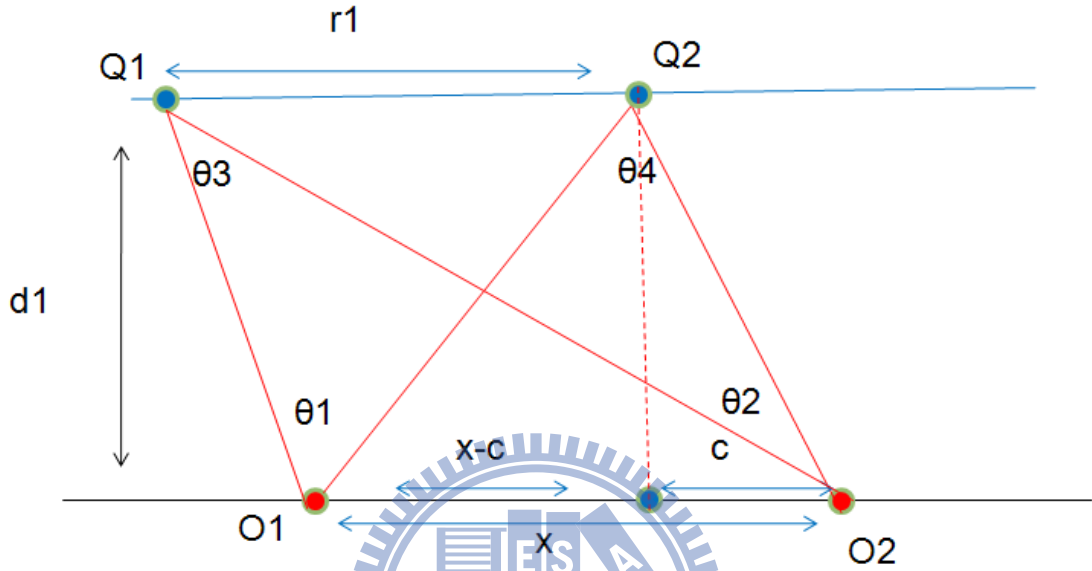


圖 31 水平方向簡化圖

從圖 31 中三角形 $Q1O1O2$ 可以得到式(12)，同樣的，從三角形 $Q2O1O2$ 可以得到式(13)， $\theta 3$ 、 $\theta 4$ 就是我們想要的兩個角度，它們的差值等於 $\theta 1$ 和 $\theta 2$ 的差值，因此把 $\theta 4 - \theta 3$ 定義成 α ，如式(14)。

$$\theta 3 = \tan^{-1} \frac{r1+c}{d1} - \tan^{-1} \frac{r1+c-x}{d1} \quad (12)$$

$$\theta 4 = \tan^{-1} \frac{c}{d1} + \tan^{-1} \frac{x-c}{d1} \quad (13)$$

$$\theta 3 - \theta 4 = \tan^{-1} \frac{r+c}{d1} - \tan^{-1} \frac{r+c-x}{d1} - \tan^{-1} \frac{c}{d1} - \tan^{-1} \frac{x-c}{d1} \quad (14)$$

有了水平方向的關係式之後，再來是垂直方向的，在圖 32 中， $r2$ 是物距， c 和水平方向的一樣是物體距離右邊相機焦點的距離， $d1$ 也一樣是物體距離相機的距離， $\theta 5 + \theta 6 - \theta 7 - \theta 8$ 就是我們想知道的 disparity。下面分別把它們求出來。從三角形 $R2M02$ 得到式(15)，三角形 $R2M01$ 得到式(16)，三角形 $R1M02$

得到式(17)，三角形 R1M01 得到式(18)。

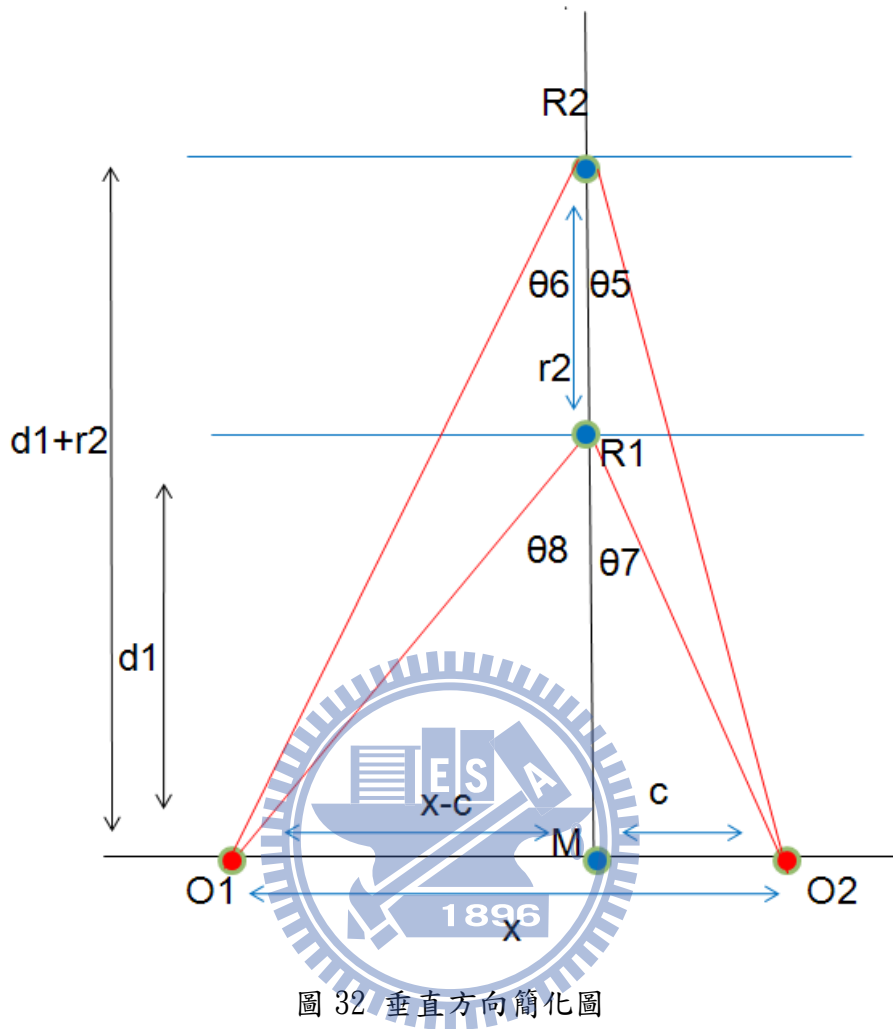


圖 32 垂直方向簡化圖

$$\theta_5 = \tan^{-1} \frac{c}{d_1 + r} \quad (15)$$

$$\theta_6 = \tan^{-1} \frac{x - c}{d_1 + r} \quad (16)$$

$$\theta_7 = \tan^{-1} \frac{c}{d_1} \quad (17)$$

$$\theta_8 = \tan^{-1} \frac{x - c}{d_1} \quad (18)$$

有了式(15)到(18)後，把他們給整理一下得到式(19)

$$\theta_7 + \theta_8 - \theta_5 - \theta_6 = \tan^{-1} \frac{c}{d_1} + \tan^{-1} \frac{x-c}{d_1} - \tan^{-1} \frac{c}{d_1+r_2} - \tan^{-1} \frac{x-c}{d_1+r_2} \quad (19)$$

最後，把圖 31 和圖 32 合併成圖 33，再把水平和垂直方向的 disparity 相加起來，式(14)和式(19)相加之後整理成為式(20)，就得到我們真正想要的角度差，進而得到 disparity，就是任意兩點在相機成像時造成的 disparity。圖 33 中， $\theta_{11} = \theta_7 + \theta_8$ ， $\theta_{22} = \theta_5 + \theta_6 = \theta_4$ ， $\theta_{33} = \theta_3$ 。

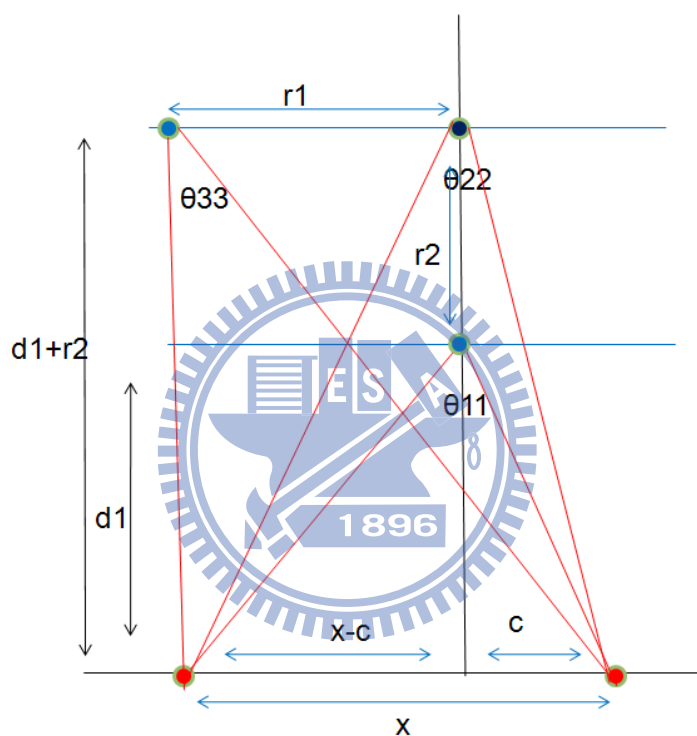


圖 33 水平垂直方向合併圖

$$\theta_3 - \theta_4 + \theta_7 + \theta_8 - \theta_5 - \theta_6 = \theta_{33} - \theta_{22} + \theta_{22} - \theta_{11} = \theta_{33} - \theta_{11}$$

$$|\theta_{33} - \theta_{11}| = \left| -\tan^{-1} \frac{r_1+c}{d_1} + \tan^{-1} \frac{r_1+c-x}{d_1} + \tan^{-1} \frac{c}{d_1+r_2} + \tan^{-1} \frac{x-c}{d_1+r_2} \right| \quad (20)$$

$$|\theta_{33} - \theta_{11}| = \left| -\tan^{-1} \frac{r_1+c}{d_1} + \tan^{-1} \frac{r_1+c-x}{d_1} + \tan^{-1} \frac{c}{d_1} + \tan^{-1} \frac{x-c}{d_1} \right| = A \quad (21)$$

得到了式(21)之後，從式子我們可能看不出什麼意思，但是我們可以設定一

些合理的參數帶入，藉此觀察式子的變化，我們將物距 $d_1=200\text{cm}$ 、水平方向的物體距離 $r_1=100\text{cm}$ 、2 號相機的焦點和右邊物體的距離 $c=-50\text{cm}$ ，這幾個變數給定值之後，使用 MATLAB 做圖，X 軸是相機焦點距離，Y 軸是像素點差距(disparity)，相機焦點距離通常不會太遠，因此讓他從 0 變化到 200，200 算是非常遠的距離了，但是我們只是想看看像素點差距的變化，畫出的圖形如從圖 34 可以看出來 disparity 隨著相機焦點距離越大而變大。所以相機焦點距離越大，所拍攝出來的影像經過縫合後，鬼影會越明顯，也有可能會到無法縫合的程度。知道了相機焦點距離和 disparity 的關係後，我們再來探討外擴式和內向式兩種擺法的相機焦點距離大小。

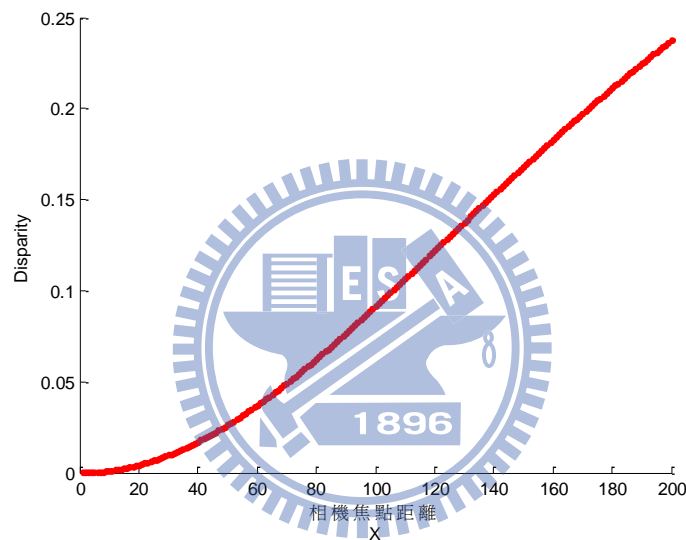


圖 34 相機焦點距離和 disparity 關係圖

3-1.2 相機視角比較

在討論相機焦點距離之前，我們必須先固定相機視角，在相同的相機視角下比較兩者的焦點距離。首先，圖 35 是內向式相機擺法，粗藍線是完整相機視角 w_1 ，紅線是單一相機視角 Φ ，兩台相機夾角 θ_1 ，相機長寬定為 a 和 b 。

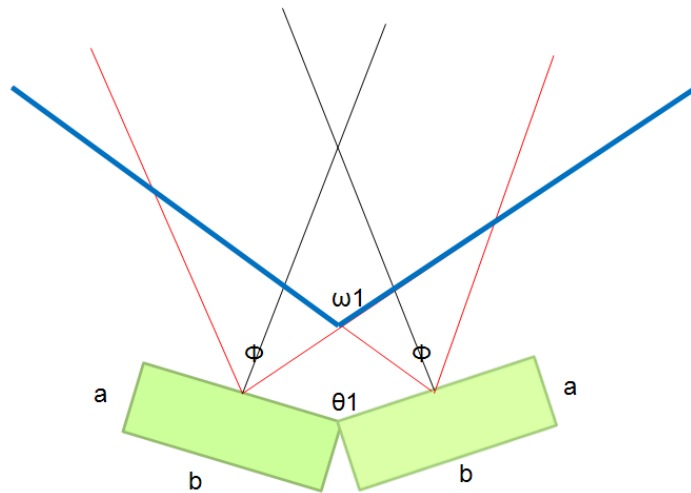


圖 35 內向式相機視角

固定好內向式的擺法之後，我們必須把外擴式的相機擺成跟內向式的視角一樣，圖 36 的(a)中假定外擴式相機夾角為 θ_2 ，視角為 w_2 ，圖 36(b)中，外角 $\omega_2 = \theta_2 + \frac{\phi}{2} + \frac{\phi}{2} = \theta_2 + \phi$ ，而圖 35 中 $\omega_1 = 180 + \phi - \theta_1$ ，我們希望這兩種的擺設方式的視角一樣，就是 w_1 要等於 w_2 ，從式(22)得到 θ_2 和 θ_1 的關係。因此當我們把內相式的相機夾角擺為 θ_1 ，而外擴式的相機夾角擺為 $\theta_2 = 180 - \theta_1$ ，它們會得到相同的視角，此時，我們再來比較它們的相機焦點距離。

$$\omega_1 = 180 + \phi - \theta_1 = \theta_2 + \frac{\phi}{2} + \frac{\phi}{2} = \theta_2 + \phi = w_2 \Rightarrow \theta_2 = 180 - \theta_1 \quad (22)$$

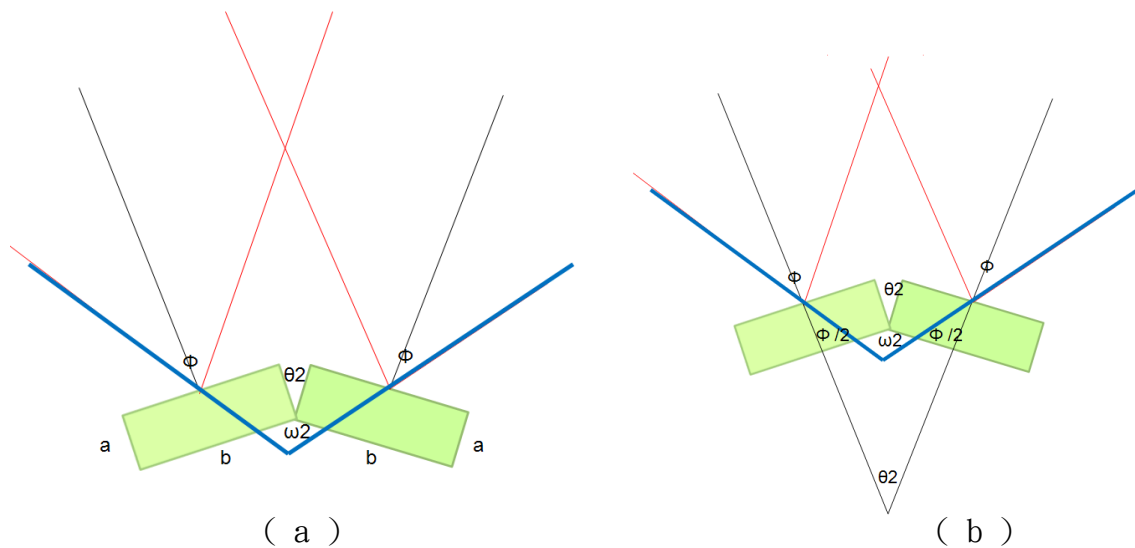


圖 36 外擴式相機視角

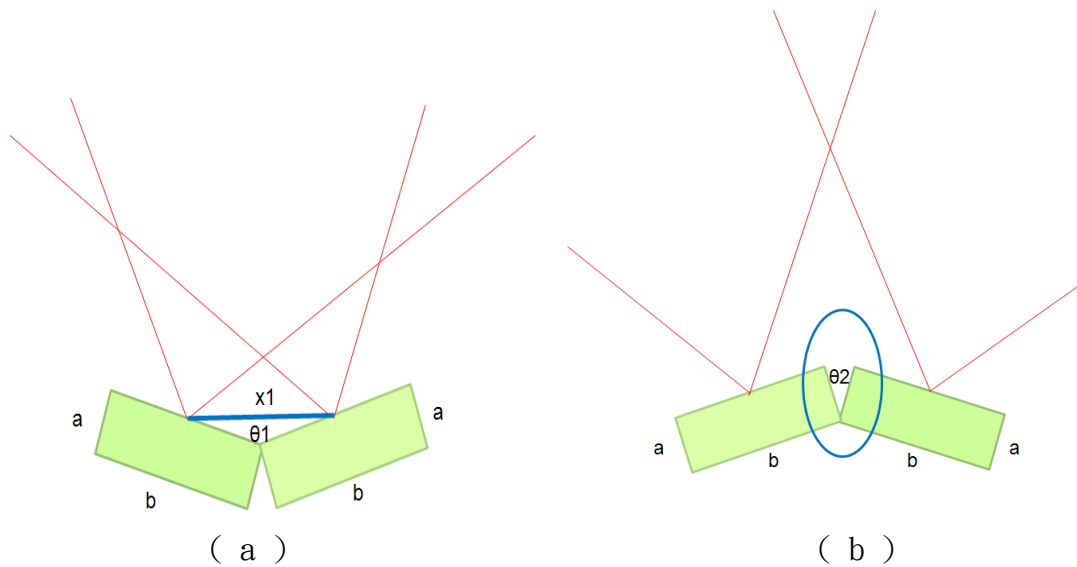


圖 37 (a)內相式相機

(b)外擴式相機

圖 37(a)中，相機夾角為 θ_1 ，相機長寬為 a 和 b ， X_1 是相機焦點距離，因為我們在先前有假設過相機焦點在相機的正前方中間位置，利用餘弦定理可以得到式(23)，進而求出 X_1 的大小。

3-1.3 相機焦點距離

式 3-12 中， X_1 只和相機寬度 b 以及相機夾角 θ_1 有關。圖 37(b)中，我們把外擴式相機的夾角擺為 $\theta_2=180-\theta_1$ ，使內向式和外擴式兩種的視角一致，然後來看看外擴式相機焦點距離，

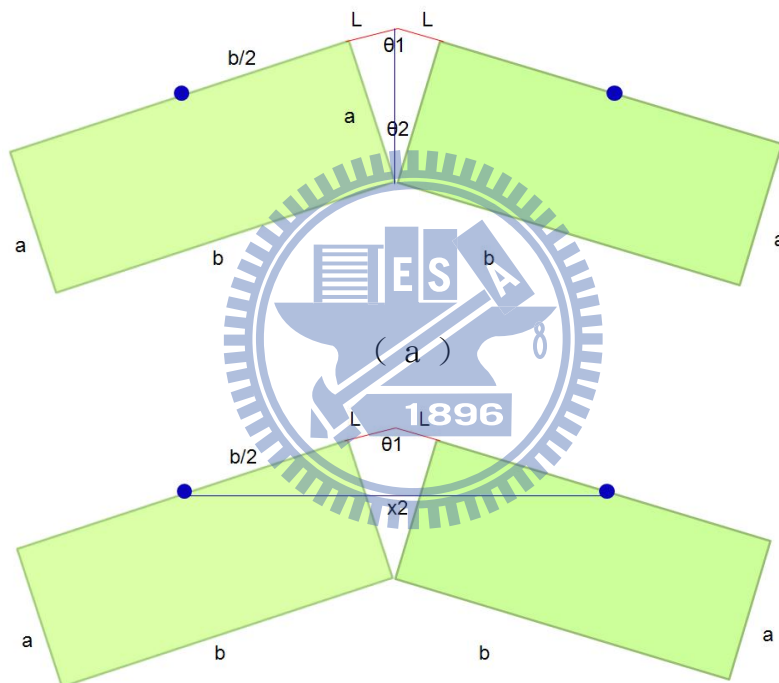
$$\cos \theta_1 = \frac{\frac{b^2}{4} + \frac{b^2}{4} - x_1^2}{2 * \frac{b}{2} * \frac{b}{2}} \Rightarrow x_1^2 = \frac{b^2}{2} (1 - \cos \theta_1) \quad (23)$$

圖 38(a)中我們先把 L 給算出來，因為外擴式相機夾角 θ_2 ，所以知道對角為 $\theta_1=180-\theta_2$ ，相基寬為 a ，因此得到式(24)。求出 L 之後，從圖 38(b)中把 X_2 給算出來，也是利用餘弦定理得到式(25)，然後求出 X_2 。

$$L = \frac{a}{\tan \frac{\theta 1}{2}} \quad (24)$$

$$\cos \theta 1 = \frac{2 * \left(\frac{b}{2} + L\right)^2 - x 2^2}{2 * \left(\frac{b}{2} + L\right)^2} \Rightarrow x 2^2 = 2 * \left(\frac{b}{2} + L\right)^2 (1 - \cos \theta 1) \quad (25)$$

$$x 2^2 = 2 * \left(\frac{b}{2} + L\right)^2 (1 - \cos \theta 1) = \left(\frac{b^2}{2} + 2bL + 2L^2\right) (1 - \cos \theta 1) \quad (26)$$



(b)

圖 38 外擴式相機放大圖

最後，我們把式(25)展開之後得到式(26)，把它拿來和式(23)比較，可以很明顯的發現 X_2 會大於 X_1 ，也就是外擴式相機焦點距離會比內向式焦點距離來的大，再配合前面推出的相機焦點距離和 disparity 的關係，當相機焦點距離越遠，disparity 就會越大。因為內相式相機焦點距離比較小，所以我們可以推測內向式相機拍出來會有比較小的 disparity，也就是鬼影效果會比較小，而理想的狀況就是當兩台相機焦點完全在一起時，這時兩張影像完全沒有 disparity，會是最好的。而外擴式相機焦點距離比較大，就比較會產生鬼影線像。因此，在使用

兩台相機拍攝時，建議使用內向式的拍攝方式，會有比較好的效果。

3-2 內相式和外擴式相機影像重疊區域

在 3-1 討論過相機焦點距離和 disparity 的關係後，接下來從重疊區域 (overlapped area) 來看看兩種相機擺法的差異。不過在討論重疊區域之前，我們也必須先做一些設定，要在內向式和外擴式能夠得到相同視角的情況下，相同視角的意思就是兩種擺法所能照到的範圍是一樣的，這樣才能公平的比較重疊區域。重疊區域的大小攸關到兩張影像是否能縫合，如果重疊區域太小影像可能會因為匹配的點不夠多而縫合失敗或是變形，但是為了要使相機取得更大的視角，因此視角和重疊區域必須要做取捨。後面我們會證明在同樣的視角下，內相式相機擺法會有較大的重疊區域。

重疊區域的比較方式是看影像中重疊部分佔完整影像的比例，圖 39 藍色斜線是重疊區塊，重疊區塊佔的比例隨著距離遠近而有所不同。X1 式相機焦點距離，在 3-1 中我們推導出的內向式相機夾角 θ_1 和外擴式相機夾角 $\theta_2=180-\theta_1$ 將之帶入， Φ 是相機視角， y_1 是一半影像， y_2 是一半重疊區域，影像距離為 D 。式 3-11 到式 3-14 分別是它們的關係式，我們想要的是 y_1 和 y_2 的比例。

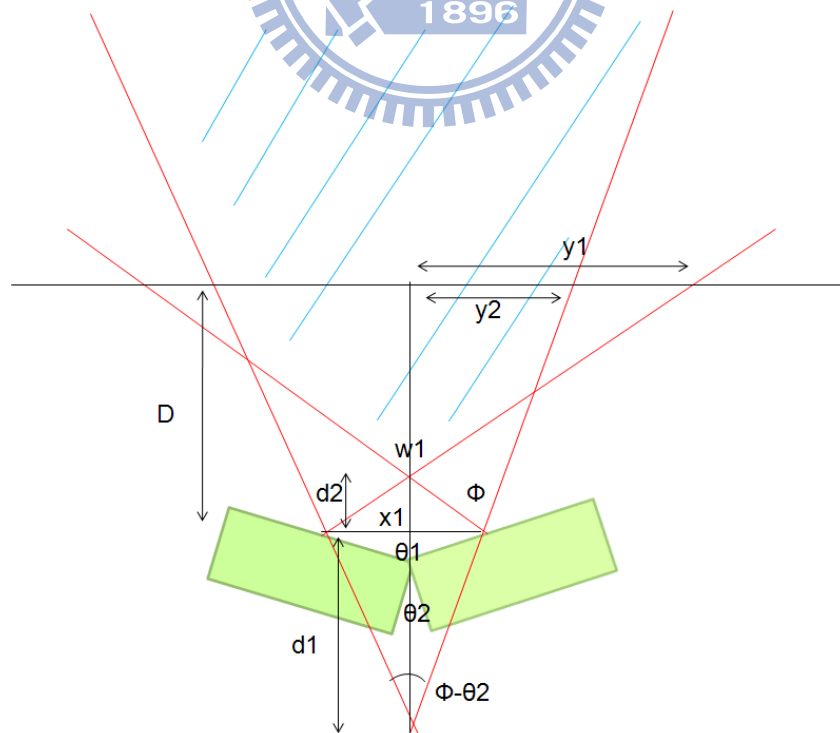


圖 39 內向式相機重疊區域

$$d1 = \frac{x1}{2 \tan \frac{\phi - \theta2}{2}} \quad (27)$$

$$d2 = \frac{x1}{2 \tan \frac{w1}{2}} \quad (28)$$

$$\tan \frac{\phi - \theta2}{2} = \frac{y2}{D + d1} \quad (29)$$

$$\tan \frac{w1}{2} = \frac{y1}{D - d2} \quad (30)$$

將式(28)代入式(30)得到式(31)，式(27)代入式(29)得到式(32)，最後，再把式(31)和式(32)結合成式(33)，就是我們想要的 $y1/y2$ 比，也就是重疊區域的比例。

$$\tan \frac{w1}{2} = \frac{y1}{D - d2} = \frac{y1}{D - \frac{x1}{2 \tan \frac{w1}{2}}} = \frac{2 * y1 * \tan \frac{w1}{2}}{2D \tan \frac{w1}{2} - x1} \quad (31)$$

$$\tan \frac{\phi - \theta2}{2} = \frac{y2}{D + d1} = \frac{y2}{D + \frac{x1}{2 \tan \frac{\phi - \theta2}{2}}} = \frac{2 * y2 * \tan \frac{\phi - \theta2}{2}}{2D \tan \frac{\phi - \theta2}{2} + x1} \quad (32)$$

$$\Rightarrow \frac{y2}{y1} = \frac{2D \tan \frac{\phi - \theta2}{2} + x1}{2D \tan \frac{w1}{2} - x1} \quad (33)$$

再來是外擴式的推導，跟內向式的方法一樣， $X2$ 為相機焦點距離， $\theta2$ 是上面推導的相機夾角， Φ 是相機視角， $y3$ 是一半影像， $y4$ 是一半重疊區域，影像距離為 D 。式(34)到式(37)分別是它們的關係式，我們想要的是 $y3$ 和 $y4$ 的比例。

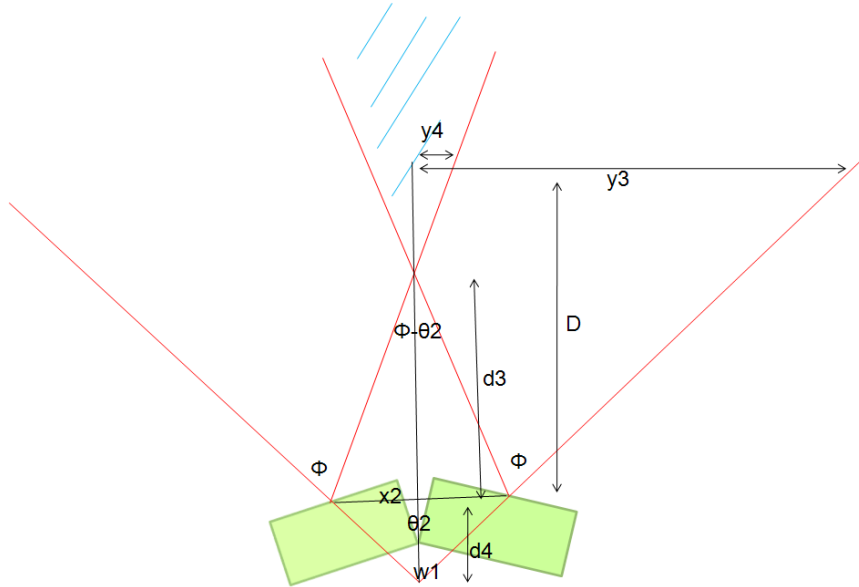


圖 40 外擴式相機重疊區域

$$d3 = \frac{x2}{2 \tan \frac{\phi - \theta2}{2}} \quad (34)$$

$$d4 = \frac{x2}{2 \tan \frac{w1}{2}} \quad (35)$$

$$\tan \frac{w1}{2} = \frac{y3}{D + d4} \quad (36)$$

$$\tan \frac{\phi - \theta2}{2} = \frac{y4}{D - d3} \quad (37)$$

將式(35)代入式(36)得到式(38)，式(34)代入式(37)得到式(39)，最後，再把式(38)和式(39)結合成式(40)，就是我們想要的 $y3/y4$ 比，也就是重疊區域的比例。

$$\tan \frac{w1}{2} = \frac{y3}{D + d4} = \frac{y3}{D + \frac{x2}{2 \tan \frac{w1}{2}}} = \frac{2 * y3 * \tan \frac{w1}{2}}{2D \tan \frac{w1}{2} + x2} \quad (38)$$

$$\tan \frac{\phi - \theta_2}{2} = \frac{y_4}{D - d_3} = \frac{y_4}{D - \frac{x_2}{2 \tan \frac{\phi - \theta_2}{2}}} = \frac{2 * y_4 * \tan \frac{\phi - \theta_2}{2}}{2D \tan \frac{\phi - \theta_2}{2} - x_2} \quad (39)$$

$$\Rightarrow \frac{y_4}{y_3} = \frac{2D \tan \frac{\phi - \theta_2}{2} - x_2}{2D \tan \frac{w_1}{2} + x_2} \quad (40)$$

有了內向式重疊區域比例式(33)和外擴式重疊區域比例式(40)，從式子我們就可以看得出來內相式的重疊區域比例一定比外擴式重疊區域比例來的大，因為式(33)的分子大於式(40)的分子而式(33)的分母小於式(40)的分母，或是給定數值作圖來看， $\Phi=60$ 度、 $\theta_2=25$ 度、 $w_1=80$ 度、 $x_1=10$ 、 $x_2=15$ ， D 從0到1000cm。表四可以看出來在近距離時，內相式的擺法能夠得到相當大的重疊區域，而遠距離時，兩種百發得到的重疊區域就差不多，將表四畫成圖41，藍線為內相式重疊比例，紅線為外擴式重疊比例，更可以明顯看出內相式的重疊區域永遠比外擴式的來的多，在縫合時更容易成功。

表四 內向式和向外擴式影像重疊區域比較

Distance (m)	0.5	1	3	5	10	100
y ₂ /y ₁ (inward)	56.21%	46.31%	40.37%	39.24%	38.41%	37.66%
y ₄ /y ₃ (outward)	16.71%	26.29%	33.60%	35.16%	36.36%	37.46%

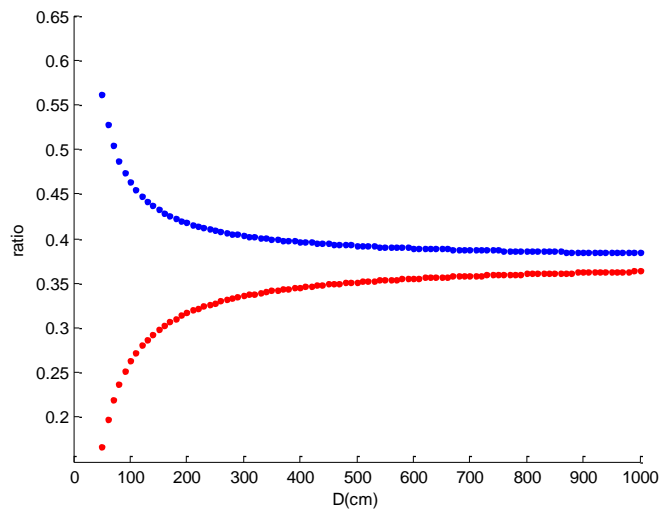


圖 41 重疊比例和距離的關係圖

3-3 內向式和外擴式相機縫合結果比較

在前面，我們了解了內向式和外擴式相機擺法的優缺點，所以本節就針對這兩種擺法所拍攝出來的影像去做縫合的動作，看看效果是否如預期的改善。我們有推導了關於內相式和外擴式影像所造成的 disparity，還有內相式和外擴式相機的重疊區域範圍的不同。圖 42 是內向式相機拍攝結果，圖 43 是外擴式相機拍攝結果。首先，我們先觀察這兩張圖的拍攝範圍，可以看得出來拍攝視角幾乎是一樣的，從最左邊的包包到最右邊的小椅子，再來是紅線標示的範圍，很明顯的，圖 42 的內向式相機比圖 43 的外擴式相機有著比較大的重疊區域。



(a)

(b)

圖 42 內向式相機拍攝結果及重疊區域



(a)

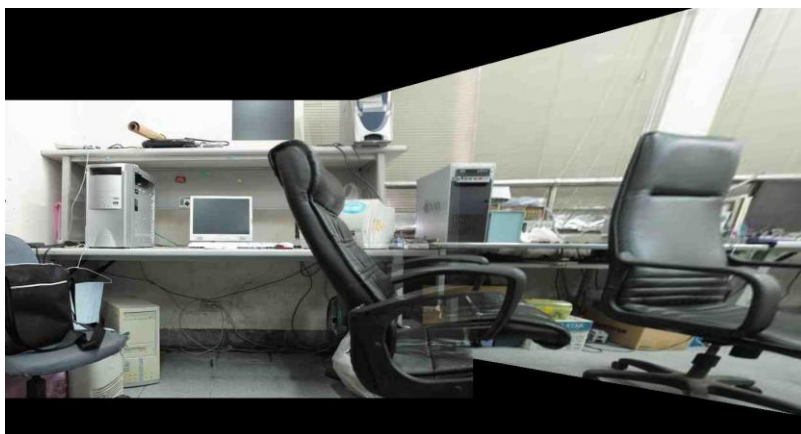
(b)

圖 43 外擴式相機拍攝結果及重疊區域

當我們要做絕對座標縫合時，圖 42 內向式的圖可以順利的逢合成圖 44，但是圖 43 外擴式的圖卻會縫成圖 45 失敗的縫合，這主要是重疊區域太小所造成的，當重疊區域太小時，我們抓到的特徵點數量可能就會不夠，使的在建立轉換模型時，容易取到錯誤的轉換模型，因此縫合的結果就會失敗。



圖 44 內相式影像絕對座標縫合結果



45 外擴式影像絕對座標縫合結果

圖 46 和圖 47 分別是內相式影像圓柱投影縫合結果和外擴式影像圓柱投影縫合結果，兩張圖乍看之下縫起來差不多，但是仔細看中間椅子的地方，外擴式的影像看起來就有較明顯的鬼影，而內向式看起來比較好一些。



圖 46 內相式影像圓柱投影縫合結果



圖 47 外擴式影像圓柱投影縫合結果

上面那組圖中圓柱縫合的影像其實看不太出來差異，式因為椅子都是黑色的比較不好分辨，下面是令外一組影像，圖 48 是外擴式影像，圖 49 是內向式影像，兩組圖的視角是一樣的，但是從縫合結果圖 50 和圖 51 可以很明顯的看出，圖 50 中的鬼影程度比圖 51 還要嚴重，也就是圖 50 中的 disparity 比較大所造成的。



圖 48 外擴式影像



圖 49 內向式影像

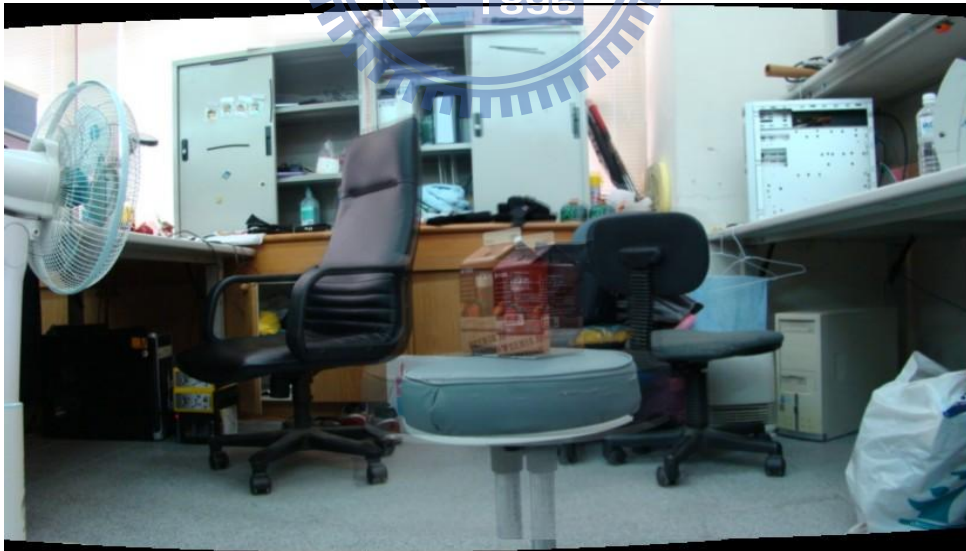
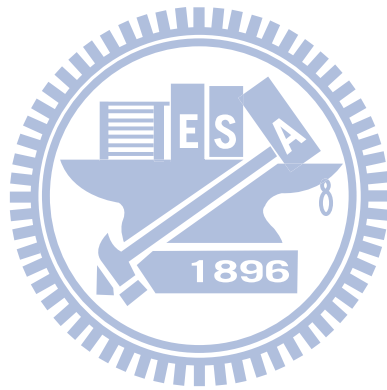


圖 50 外擴式影像圓柱投影縫合結果



圖 51 內向式影像圓柱投影縫合結果



第四章 強化特徵點匹配之正確性

在縫合時，我們發現先前的做法無法縫合明暗差距過大的兩張圖，尤其是在室內，由於燈光的影响常常會使的兩個區域有明暗差距，此時逢合就容易失敗，這裡我們也在演算法上做了一些改善，使它們能夠順利縫合。

4-1 特徵點匹配的缺點

我們知道縫合的流程如圖 52，但是我們把之前流程中的特徵點匹配做了一些修正，圖 52 中紅色標示部分。

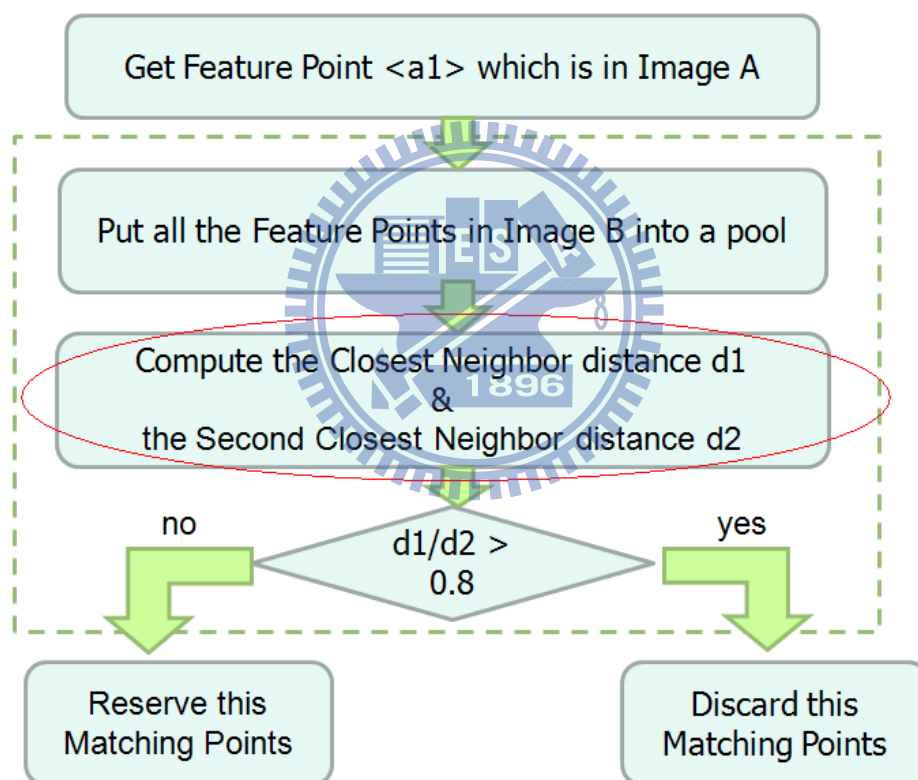


圖 52 影像縫合流程圖

原本的做法中，當擷取完特徵點(Feature points)之後，接下來需要把兩張圖的特徵點彼此之間做匹配(Matching)，先把要縫合的兩張圖分別用A圖和B圖代稱，所以看看A圖裡面的所有特徵點在B圖裡面是對應到哪些點，先前的方法是直接去計算兩個特徵點周圍的一個小範圍的SSD(Sum of Squared Differences)，並定下一個臨界值(Threshold)為門檻，因為當兩個點彼此是對

應點，則其附近的Pixel值會非常接近，所以當SSD越小，自然表示兩個A圖的 $(x1, y1)$ 點很有可能是對應到B圖的 $(x2, y2)$ 點。

先前的方法在遠距離影像的縫合時可以適用，因為遠距離縫合在旋轉相機視角時，相機不太會受到光線的多寡去調整整張影像的像素值，但是在近距離時，相機轉一個角度就很容易受到光線分佈的影響，因為我們沒辦法確保每個角度的光線都是充足的，如圖 53(a)(b)為遠距離，(c)(d)為近距離。可以很明顯的從(c)(d)中看到(d)的箱子受到光線的影響比(c)的箱子來的暗，因為圖 53(d)中日光燈照射量較少。如此一來，直接計算兩個特徵點周圍的一個小範圍的SSD的話，會造成對應點的pixel值並不相近，因而造成誤判。所以使用此做法來縫合圖 53 的(c)(d)是會失敗的。



圖 53 遠近拍攝影像明暗差異

4-2 改良做法

4-2.1 相關性

第一種做法是去計算對應點的相關性(correlation)[44]，公式為式(41)，因為兩張圖的像素值雖然差很多，但是對應點的像素值差距幾乎是一樣的，因此對應點的相關性會很高，藉由計算 A 圖特徵點旁邊 8*8 範圍的像素值和 B 圖特徵點旁邊 8*8 範圍的像素值兩者的相關性來判斷是否為對應點，會比計算他們的 SSD 來的準確。

$$r = \frac{\sum_m \sum_n (A_{mn} - \bar{A})(B_{mn} - \bar{B})}{\sqrt{(\sum_m \sum_n A_{mn} - \bar{A}^2)(\sum_m \sum_n B_{mn} - \bar{B}^2)}} \quad (41)$$

4-2.2 像素值排序

但是計算相關性是非常耗時的作法，因此我們另外採用的方法是先將特徵點旁邊 8*8 範圍的像素值去作大小排列，[41]-[43]中都有類似的作法，如圖 54 為一個 3*3 的範例，左邊的矩陣式像素值，右邊的矩陣是排序大小，像 100 是最大值，因此排序出來為 1；而 10 是最小值，因此排序出來為 9。式(42)是把圖 54 轉成 vector 的方式來計算 SSD，把原本的矩陣，由左到右，由上到下，依序排序成為一 vector。

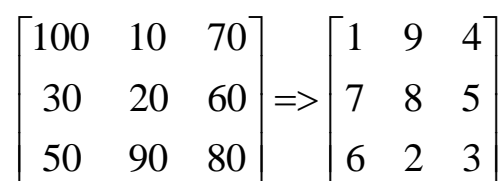

$$\begin{bmatrix} 100 & 10 & 70 \\ 30 & 20 & 60 \\ 50 & 90 & 80 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 9 & 4 \\ 7 & 8 & 5 \\ 6 & 2 & 3 \end{bmatrix}$$

圖 54 像素值排列範例一

$$[100 \ 10 \ 70 \ 30 \ 20 \ 60 \ 50 \ 90 \ 80] \Rightarrow [1 \ 9 \ 4 \ 7 \ 8 \ 5 \ 6 \ 3 \ 2] \quad (42)$$

就像前面說的，兩張圖的像素值雖然差很多，但是對應點的像素值差距幾乎是一樣的，因此兩張圖排列出來的順序(rank order)應該是差不多的，rank order 就是圖 54 中右邊的矩陣。舉例來說，假設 A 圖的像素值如圖 55(a)，那麼 B 圖的像素值可能就如圖 55(b)，每個像素值都比圖 55(a)中的像素值高 100，因此

經過排序後，兩邊都會得到相同的排序矩陣。所以，計算完兩張圖的 rank order，再套用原本的方法，計算兩個特徵點周圍的一個小範圍的 SSD，這裡跟原本方法不同的是，原本是用像素值計算 SSD，而現在是用排列順序(rank order)去計算 SSD，這樣我們不但就可以解決因為光線的影響而無法得到正確匹配特徵點的問題，而且比尋找相關性的方法耗時較少，只需要做排序，和計算 SSD，也不需要用到除法。

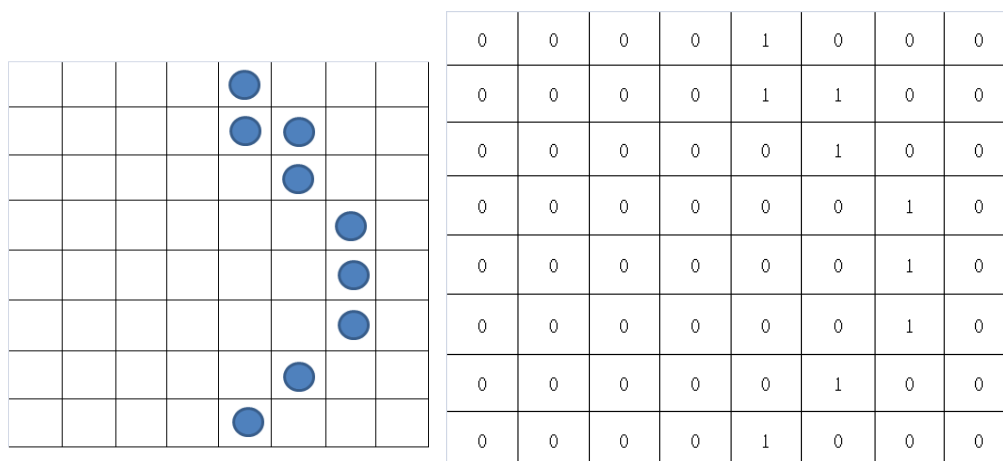
$$\begin{bmatrix} 100 & 10 & 70 \\ 30 & 20 & 60 \\ 50 & 90 & 80 \end{bmatrix} \qquad \begin{bmatrix} 200 & 110 & 170 \\ 130 & 120 & 160 \\ 150 & 190 & 180 \end{bmatrix}$$

(a) (b)

圖 55 像素值排列範例二

4-2.3 邊緣強化

在做特徵點擷取時，其實就常常會取到邊緣或是角落的點，而這些是比較殊的點所以可能拿來做匹配的動作，利用這個特性，我們把特徵點找到之後，可以想見附近應該有不少邊緣的點，把特徵點附近 8*8 的區域做高頻濾波(high pass filter)，然後把邊緣(edge)的點給取出來，方法是設定一個臨界值，高於臨界值設定為 1，低於臨界值則設成 0，如圖 56(a)是原本 8*8 大小的範圍的像素，藍點是 edge 的地方，我們設定臨界值之後可以把 edge 給取出來，變成圖 56(b)，1 的地方代表 edge，0 是平坦的區域。有了這 8*8 的數值之後，就可以拿去做匹配，這個方法比上面兩種方法來的快速很多，而且效果也很好，因為他取出來的 edge 正是我們特徵點想要的部分，剛好跟著特徵點的特性來走，當然，明暗度的問題也能夠順利改善。



(a)

(b)

圖 56 8*8 範圍的像素值示意圖



圖 57 縫合結果圖

圖 57 是利用 MATLAB 縫合完成之後，可以得到兩張影像的關係模型，然後完成縫合的動作。

4-3 模擬結果與匹配方法比較

4-2.2 中提到的三種方法，相關性(correlation matching)、像素值排列(pixel ranking)、邊緣強化(edge enhancement)，我們分別使用絕對座標和圓柱投影兩種縫合方式，來比較他們的速度和縫合效果。圖 58 有三組圖，從上到下分別是遠距離，中距離，近距離。使用上面三種方法都可以順利縫合，圖 59 是絕對座標縫合的結果，不論是近距離，中距離，遠距離都縫得不錯，中距離那張圖算是比較有挑戰性的，因為圖中他們的色彩顏色差不多，而且左邊比較亮，右邊比較暗一點，不過還是都可以縫的起來。圖 60 是圓柱投影的縫合結果，圓柱投影其實可以縫合 360 度，但是我們為了配合絕對座標縫合，所以只取兩張圖，比較方便比較它們的結果。



圖 58 pattern 三組



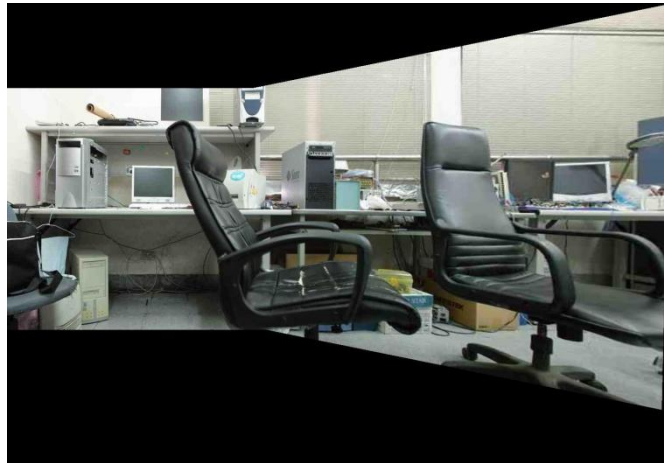


圖 59 絕對座標縫合結果



圖 60 圓柱投影縫合結果

我們把三組 pattern 都各跑兩次來比較它們的速度，從表五中，correlation

matching 的計算量大，所以平均需要 62 秒，pixel ranking 稍微快一點，耗時 53 秒，而 edge enhancement 平均只需要 21 秒，大幅提升原本的速度。表六是圓柱投影的速度比較，也可以很明顯的看到 edge enhancement 的速度比另外兩種方法快很多。所以，使用 edge enhancement 的方法可以改善明暗匹配的問題，而且也不會增加太多的計算量。

	1	1	2	2	3	3	Avg
Correlation Matching	55s	55s	59s	60s	70s	75s	62s
Pixel ranking	52s	52s	46s	47s	60s	59s	53s
Edge enhancement	15s	15s	26s	26s	36s	34s	21s

表五 絕對座標縫合三種方法速度比較

	1	1	2	2	3	3	Avg
Correlation matching	65s	65s	123s	124s	151s	157s	114s
Pixel ranking	50s	50s	104s	106s	131s	130	95s
Edge enhancement	17s	16s	45s	45s	59s	58s	40s

表六 圓柱投影縫合三種方法速度比較

第五章 結論與未來展望

縫合的演算法到目前已經算是很成熟了，大家不斷的追求快速且能夠準確縫合的方法，想盡辦法在演算法上面做改善，像是我們在第三章所提到的鬼影，其實也有演算法可以改善，但是如果能夠在外部就能做好的，不但減少演算法的運算複雜度，更能得到真實的影像。內向式(Inward)的相機擺設方式，是直接的從根本來改善會產生鬼影的原因，所以不論是絕對座標的縫合方式，或者是圓柱投影的縫合，兩種都可以大幅降低鬼影的效應。內相式相機不但可以改善鬼影，還能夠提升縫合的準確度，因為內向式相機的重疊區域(overlap area)比較大，所以我們能夠結取到更多的匹配的特徵點(feature points)，這樣我們才比較容易得到正確的轉換函數，使我們順利縫出漂亮的圖。

至於影像明暗不一造成縫合失敗的問題，我們也提出了幾種改善方式，可以使用相關性(correlation)、像素值排列(rank order)、邊緣判斷(edge judgment)，這三種方法都可以改善原本無法縫合的問題，但是在速度和成功率方面有著不同的表現，相關性這方法雖然速度很慢，但是他卻可以成功縫合大部分的影像，相較之下，雖然像素值排列比相關性的速度快一點，可是縫合的效果不是那麼好。邊緣判斷的速度是最快的，而且也能夠成功縫合大部分的圖片，縫合的能力只比相關性差一點點。

在未來雙鏡頭的縫合中，如果我們能夠使用內相式雙鏡頭盡量把焦點距離縮小，這樣不但鬼影的效果可以大幅改善，也代表著遠距離的影像和近距離的影像在匹配時，兩者的轉換模型幾乎一致，如此一來，當我們把雙鏡頭固定之後，只需要經過一次的校正得到轉換模型，之後就可以固定此轉換模型縫合任意雙鏡頭所拍攝的影像。如果是外擴式雙鏡頭就會發生遠距離和近距離無法匹配的情況，當你想匹配遠距離的影像那麼近距離的影像就會有鬼影，反之，如果想匹配近距離的影像那麼遠距離的影像就會出現鬼影，因此將無法固定一個轉換模型。有了固定的轉換模型之後，我們可以進行硬體實現，硬體部份只需要存入轉換模型，後面只要再作影像混合和縫合，前面的特徵點擷取、特徵點匹配還有 RANSAC 都可以省下來，達到即時(real time)縫合。

参考文献：

- [1] C. Rother and S. Carlsson. Linear multi view reconstruction and camera recovery using a reference plane. *International Journal of Computer Vision*, 49(2/3):117–141, 2002.
- [2] B. Bascle, A. Blake, and A. Zisserman. Motion deblurring and super-resolution from an image sequence. In *Proceedings of the 4th European Conference on Computer Vision (ECCV96)*, pages 312–320. Springer-Verlag, 1996.
- [3] D. Capel and A. Zisserman. Automated mosaicing with super-resolution zoom. In *Proceedings of the International Conference on Computer Vision and Pattern Recognition (CVPR98)*, pages 885–891, June 1998.
- [4] J. Davis. Mosaics of scenes with moving objects. In *Proceedings of the International Conference on Computer Vision and Pattern Recognition (CVPR98)*, pages 354–360, 1998.
- [5] M. Uyttendaele, A. Eden, and R. Szeliski. Eliminating ghosting and exposure artifacts in image mosaics. In *Proceedings of the International Conference on Computer Vision and Pattern Recognition (CVPR01)*, volume 2, pages 509–516, Kauai, Hawaii, December 2001.
- [6] A. Agarwala, M. Dontcheva, M. Agarwala, S. Drucker, A. Colburn, B. Curless, D. Salesin, and M. Cohen. Interactive digital photomontage. In *ACM Transactions on Graphics (SIGGRAPH'04)*, 2004.
- [7] D. Brown. Close-range camera calibration. *Photogrammetric Engineering*, 37(8):855–866, 1971.
- [8] D.B. Goldman and J.H. Chen. Vignette and exposure calibration and compensation. In *Proceedings of the 10th International Conference on Computer Vision (ICCV05)*, pages I: 899–906, 2005.
- [9] P. Debevec and J. Malik. Recovering high dynamic range radiance maps from photographs. *Computer Graphics*, 31:369–378, 1997.
- [10] H. Seetzen, W. Heidrich, W. Stuerzlinger, G. Ward, L. Whitehead, M. Trentacoste, A. Ghosh, and A. Vorozcovs. High dynamic range display systems. In *ACM Transactions on Graphics (SIGGRAPH'04)*, 2004.
- [11] R. Szeliski and S. Kang. Direct methods for visual scene reconstruction. In *IEEE Workshop on Representations of Visual Scenes*, pages 26–33, Cambridge, MA, 1995.

- [12] M. Irani and P. Anandan. About direct methods. In B. Triggs, A. Zisserman, and R. Szeliski, editors, *Vision Algorithms: Theory and Practice*, number 1883 in LNCS, pages 267–277. Springer-Verlag, Corfu, Greece, September 1999.
- [13] H. Sawhney and R. Kumar. True multi-image alignment and its application to mosaicing and lens distortion correction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(3):235–243, 1999.
- [14] H. Shum and R. Szeliski. Construction of panoramic mosaics with global and local alignment. *International Journal of Computer Vision*, 36(2):101–130, February 2000.
- [15] I. Zoghlami, O. Faugeras, and R. Deriche. Using geometric corners to build a 2D mosaic from a set of images. In *Proceedings of the International Conference on Computer Vision and Pattern Recognition, Puerto Rico*. IEEE, June 1997.
- [16] P. McLauchlan and A. Jaenicke. Image mosaicing using sequential bundle adjustment. *Image and Vision Computing*, 20(9-10):751–759, August 2002.
- [17] R. Szeliski. Image Alignment and Stitching: A Tutorial1 Preliminary draft, September 27, 2004 Technical Report. MSR-TR-2004-92.
- [18] Chen, S. E. QuickTime VR – an image-based approach to virtual environment navigation. *Computer Graphics (SIGGRAPH'95)*, , 29–38. 1995.
- [19] Szeliski, R. Image mosaicing for tele-reality applications. In *IEEE Workshop on Applications of Computer Vision (WACV'94)*, pages 44–53, IEEE Computer Society, Sarasota. 1994.
- [20] Szeliski, R. Creating Full View Panoramic Image Mosaics and Environment Maps. Microsoft Research.
- [21] irahu <http://www.irahul.com/workzone/stitch/index.php>
- [22] P. Heckbert. Projective Mappings for Image Warping Rendering, 15-862 .26 Sept 1995.
- [23] C. Harris and M.J. Stephens. A combined corner and edge detector. In *Alvey Vision Conference*, pages 147–152, 1988.
- [24] H. Moravec. Obstacle avoidance and navigation in the real world by a seeing robot rover. Technical Report CMU-RI-TR-3, Carnegie-Mellon University, Robotics Institute, 1980.
- [25] C. Schmid, R. Mohr, and C. Bauckhage. Evaluation of interest point detectors. *International Journal of Computer Vision*, 37(2):151–172, June 2000.

- [26] G. Derpanis. The Harris Corner Detector, October 27, 2004.
- [27] M. Brown, R. Szeliski, and S. Winder. Multiimage matching using multi-scale oriented patches. In *Proceedings of the International Conference on Computer Vision and Pattern Recognition (CVPR05)*, San Diego, June 2005.
- [28] D. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110, 2004.
- [29] J. Shi and C. Tomasi. Good features to track. In *Proceedings of the International Conference on Computer Vision and Pattern Recognition (CVPR94)*, Seattle, June 1994.
- [30] Y. Ke1, R. Sukthankar. PCA-SIFT: A More Distinctive Representation for Local Image Descriptors. Intel Research Pittsburgh In *Proceedings of the International Conference on Computer Vision and Pattern Recognition (CVPR04)*.
- [31] D. Lowe, "Object recognition from local scale-invariant features," *International Conference on Computer Vision*, Corfu, Greece (September 1999), pp. 1150-1157.
- [32] D. Lowe, "Local feature view clustering for 3D object recognition," *IEEE Conference on Computer Vision and Pattern Recognition*, Kauai, Hawaii (December 2001), pp. 682-688.
- [33] M. Fischler and R. Bolles. Random sample consensus: A paradigm for model fitting with application to image analysis and automated cartography. *Communications of the ACM*, 24:381–395, 1981.
- [34] P. Torr. Bayesian model estimation and selection for epipolar geometry and generic manifold fitting. *International Journal of Computer Vision*, 50(1):35–61, 2002.
- [35] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, ISBN: 0521540518, second edition, 2004.
- [36] M. Brown, D. Lowe. Automatic Panoramic Image Stitching using Invariant Features. *International Journal of Computer Vision(ijcv)*,2007.
- [37] M. Brown and D. Lowe. Recognising panoramas. In *Proceedings of the 9th International Conference on Computer Vision (ICCV03)*, volume 2, pages 1218–1225, Nice, October 2003.

- [38] 莊詠麟, 相機陣列影像縫合之快速校正技術, 國立交通大學碩士論文, 2008
- [39] P. Burt and E. Adelson. A multiresolution spline with application to image mosaics. *ACM Transactions on Graphics*, 2(4):217–236, 1983.
- [40] D.Y. Chen, M. Ouhyoung, “VideoVR: A real-timesystem for automatically constructing panoramic images from video clips”, Proc. of CAPTECH’98, Geneva,Switzerland, Nov. 1998, pp.140-143. (Lecture Notes in AI 1537, Springer)
- [41] Bhat, D. N. and S. K. Nayar (1998). "Ordinal measures for image correspondence." Pattern Analysis and Machine Intelligence, IEEE Transactions on **20**(4): 415-423.
- [42] Mittal, A. and V. Ramesh (2006). An Intensity-augmented Ordinal Measure for Visual Correspondence. Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on.
- [43] Posse, A., J. Torres, et al. (2009). Matching points in poor edge information images. Image Processing (ICIP), 2009 16th IEEE International Conference on.
- [44] M. Ehlers and D. Fogel "High-precision geometric correction of airborne remote sensing revisited: The multiquadric interpolation", *Proc. SPIE—Image Signal Processing Remote Sensing*, vol.2315, pp. 814 1994.
- [45] J. Serra, “Image Analysis and Mathematical Morphology” vol.1, Academic Press,1982.
- [46] J. Serra, “Image Analysis and Mathematical Morphology” vol.2, Academic Press,1988.
- [47] Stanley Osher and James A. Sethian, “Fronts propagating with curvaturedependent speed: algorithms based on Hamilton-Jacobi formulations”. *J.Comput. Phys.*, 79(1):12–49, 1988.
- [48] Richard Tsai and Stanley Osher, “Level Set Methods in Image Science”. IEEE Image Processing, 2003. Proceedings. 2003 International Conference on Volume 2,14-17
- [49] Stanley Osher and Ronald Fedkiw, “Level Set Methods and Dynamic Implicit Surfaces”, Springer; 1 edition, New York, November 1, 2002