

國立交通大學

資訊科學系

碩士論文

室內場景特徵點對應之改善

The Improvement of Extraction and Selection of Feature Points
of Indoor Scenes

研究生：黃耀輝

指導教授：莊仁輝 教授

中華民國九十三年六月

室內場景特徵點對應之改善

The Improvement of Extraction and Selection of Feature Points of Indoor Scenes

研究生：黃耀輝

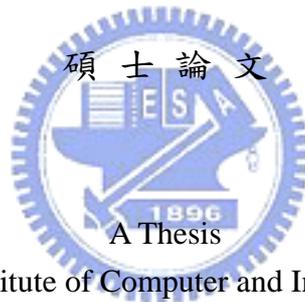
Student：Yao-Hui Huang

指導教授：莊仁輝

Advisor：Jen-Hui Chuang

國立交通大學

資訊科學系



Submitted to Institute of Computer and Information Science

College of Electrical Engineering and Computer Science

National Chiao Tung University

in partial Fulfillment of the Requirements

for the Degree of

Master

in

Computer and Information Science

June 2004

Hsinchu, Taiwan, Republic of China

中華民國九十三年六月

室內場景特徵點對應之改善

學生：黃耀輝

指導教授：莊仁輝 博士

國立交通大學

資訊科學研究所

摘要

為了兼顧機器人視覺系統建構室內場景所需的正確性和即時性，本篇論文提出一個演算法來改善特徵點對應的問題，並利用極線幾何的限制刪除錯誤的特徵點對應。我們使用 Harris 演算法擷取一對立體影像中的角點作為特徵點，接著分析在 Harris 計算過程所得到的影像梯度，將影像梯度方向量化為九個區域，分別以不同的顏色來代表這些區域，最後得到的彩色碼即可做為每個角點具有的特徵。若左右影像特徵點的彩色碼越接近，則表示此兩點對應到真實場景中同一點的機率越大，但相鄰的區域也可能有彩色碼完全相同的點，為了解決模稜兩可的問題，透過距離、角度限制和雙向對應一致性的要求，可以計算出一個評量對應關係的分數當作篩選標準，如此即可找出初步的對應，最後再加上極線限制，一組正確的對應應該要滿足兩個極線限制：也就是特徵點必須落在對應的極線上、所有的極線必須交於極點之上。觀察極線分佈的情形後，分別對極線距離以及極點到極線的距離設合理的門檻值，如此可去除錯誤的對應。同時以剩下的點計算基本矩陣，並重複以上步驟。我們可以找出最正確的特徵點對應。

The Improvement of Extraction and Selection of Feature Points of Indoor Scenes

Student : Yao-Hui Huang

Advisor : Dr. Jen-Hui Chuang

Institute of Computer and Information Science

National Chiao Tung University

ABSTRACT

To achieve the correctness of the vision system, this thesis proposes an algorithm to solve the problem of feature point correspondences which uses the epipolar constraints to delete false matches. We first extract corner points from a pair of stereo images as feature points by the Harris corner detector. The local gradient directions are represented as different colors for each feature point to form a color code. The point correspondences are then obtained by comparing the color codes as well as the spatial relationships between neighboring feature points. Finally, we add the epipolar constraints to find the best point matches. A good feature point must satisfy two epipolar constraints. One is that each feature point should lie on its epipolar line. The other is that epipolar lines should intersect at the epipole. For a pair of feature points, we compute the epipolar distance and the distance between the epipolar line to epipole. Then, we set two thresholds of distance to discard false matches and recompute the fundamental matrix in iteration until all matches satisfy the epipolar constraints.

致謝

本篇論文得以完成，中間經歷了相關研究蒐集、摸索程式、實驗、論文內容的撰寫與修正，深深體會到要完成一篇論文，是需要多少的心血。這兩年來在學習的過程中有很多的收穫，首先要感謝指導教授莊仁輝博士，他對研究的態度給我很大的啟發，也感謝口試委員李嘉晃教授、顏嗣鈞教授、雷欽隆教授對本篇論文不足之處所提出的建議，讓本篇論文更加完整。

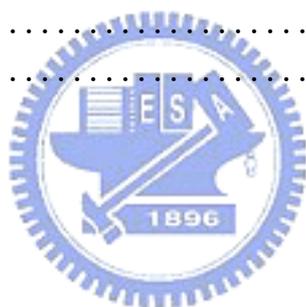
在兩年的碩士班生活中，感謝實驗室博士班學長高肇宏和林建州在課業上的指導，感謝哲寬、存厚、兆良，在那些一起努力寫論文的的日子，彼此互相鼓勵是進步的動力，最後我要感謝我的家人，因為有他們的支持我才能取得碩士學位，感謝所有曾幫助過我的人。



目錄

摘要.....	I
ABSTRACT.....	II
致謝.....	III
目錄.....	IV
圖目錄.....	VI
表目錄.....	VIII
第一章 簡介.....	1
第二章 相關文獻探討.....	5
2.1 RANSAC.....	6
2.2 M-estimators.....	8
2.3 Least Median of Squares (LMedS).....	10
2.4 前三種方法之比較.....	11
2.5 Case deletion diagnostics.....	12
第三章 基本矩陣與極線幾何.....	14
3.1 極線幾何.....	14
3.2 基本矩陣之特性.....	16
3.3 必要矩陣.....	18
3.3.1 正規化影像座標.....	18
3.3.2 必要矩陣之特性.....	19
3.3.3 透過必要矩陣計算相機矩陣.....	19
第四章 特徵點擷取與對應.....	21
4.1 特徵點的擷取.....	21
4.2 彩色碼.....	26
4.3 立體影像特徵點的對應.....	30
4.3.1 彩色碼之特性.....	30
4.3.2 雙向對應一致性.....	32
4.3.3 利用彩色碼進行特徵點的對應.....	33
第五章 計算基本矩陣.....	36
5.1 八點演算法.....	36
5.1.1 線性最小平方法.....	37
5.1.2 特徵分析(Eigen analysis).....	38
5.1.3 限制基本矩陣的秩為二.....	38
5.1.4 線性關係的幾何表示.....	40
5.2 極線距離之非線性最小化.....	40
5.2.1 反覆式線性法.....	41
5.2.2 參數空間之非線性最小化.....	42

5.3	梯度分析.....	44
5.4	正規化.....	45
第六章	利用極線幾何改善特徵點之對應.....	47
6.1	極線距離之限制.....	47
6.2	極點到極線的距離之限制.....	48
6.2.1	利用pseudo-inverse計算極點.....	49
6.2.2	利用基本矩陣計算極點.....	50
6.3	不同門檻值的影響.....	51
6.4	特徵點對應之改善.....	52
第七章	實驗結果.....	54
7.1	利用極線限制刪除錯誤的對應.....	54
7.1.1	僅對極線距離設限制.....	56
7.1.2	僅對極點到極線的距離設限制.....	58
7.1.3	同時對極線距離和極點到極線的距離設限制.....	61
7.2	重算基本矩陣和不重算基本矩陣的影響.....	65
7.3	極點位置的不確定性.....	67
第八章	結論.....	70
參考文獻	71



圖目錄

圖 3.1	(a)特徵點對應 (b)極線限制.....	14
圖 3.2	極點可視為任兩條極線的交點.....	15
圖 3.3	基本矩陣的產生過程.....	17
圖 3.4	透過必要矩陣做校正重建後的四種可能解.....	20
圖 4.1	方形遮罩在影像中移動 (a)遮罩在灰階變化平坦的影像區域中移動 (b)遮罩在邊或線的影像區域中移動 (c)遮罩在具有特徵點的影像 區域中移動.....	22
圖 4.2	各種不同特徵點擷取演算法之結果.....	23
圖 4.3	使用Harris演算法擷取一對立體影像之特徵點.....	25
圖 4.4	(a)原始影像 (b)做 $\partial I / \partial x$ 後的影像 (c)做 $\partial I / \partial y$ 後的影像 (d)Sobel operator G_x (e)Sobel operator G_y	26
圖 4.5	以不同顏色代表不同的梯度方向.....	27
圖 4.6	(a)原圖 (b)將影像梯度方向分類.....	28
圖 4.7	(a)特徵點四周的八個 3×3 小視窗 (b)以彩色碼代表每個視窗中的 梯度方向.....	28
圖 4.8	(a)左影像 (b)右影像 (c)左右影像紅色圓圈特徵點之彩色碼 (d)左影像橘色圓圈特徵點之彩色碼 (e)右影像橘色圓圈特徵點之 彩色碼.....	29
圖 4.9	彩色碼距離.....	31
圖 4.10	雙向對應一致性.....	32
圖 4.11	利用已知對應點解決模稜兩可的對應問題.....	33
圖 6.1	右影像和左影像的極線.....	47
圖 6.2	特徵點對應之改善.....	53
圖 7.1	(a)左影像特徵點和彩色碼 (b)右影像特徵點和彩色碼 (c)160 組 對應特徵點.....	55
圖 7.2	(a)160 組對應點左影像極線分佈圖 (b)160 組對應點右影像極線分 佈圖.....	56
圖 7.3	(a)150 組對應點左影像極線分佈圖 (b)150 組對應點右影像極線分 佈圖.....	57
圖 7.3	(續).....	58
圖 7.4	(a)130 組對應點左影像極線分佈圖 (b)130 組對應點右影像極線分 佈圖 (c)107 組對應點左影像極線分佈圖 (d)107 組對應點右影 像極線分佈圖 (e)51 組對應點左影像極線分佈圖 (f)51 組對 應點右影像極線分佈圖.....	60
圖 7.5	圖 7.4(b)去掉極線距離誤差最大的 5 組對應點並重算基本矩陣，可得 正確的結果.....	61

圖 7.6	(a)113 組對應點左影像極線分佈圖 (b)113 組對應點右影像極線分佈圖(c)103 組對應點左影像極線分佈圖 (d)103 組對應點右影像極線分佈圖.....	63
圖 7.7	(a)左影像特徵點和彩色碼 (b)右影像特徵點和彩色碼 (c)22 組對應特徵點.....	64
圖 7.8	未改變基本矩陣狀況下 (a)150 組對應點右影像極線分佈圖 (b)120 組對應點右影像極線分佈圖 (c)80 組對應點右影像極線分佈圖 (d)40 組對應點右影像極線分佈圖	66



表目錄

表 2.1	不同情況下所需的子集合數， $P=0.95$ 。	7
表 6.1	門檻值的大小對於特徵點對應組數的影響	52
表 7.1	僅對極線距離設限制	57
表 7.2	僅對極點到極線的距離設限制	59
表 7.3	同時對兩種距離設限制	62
表 7.4	同時對兩種距離設限制	62
表 7.5	不同場景的實驗	65
表 7.6	不改變基本矩陣	65
表 7.7	極點位置(僅對極線距離設限制)	67
表 7.8	極點位置(僅對極點到極線的距離設限制)	67
表 7.9	極點位置(同時對兩種距離設限制)	68
表 7.10	極點位置(同時對兩種距離設限制)	68



第一章 簡介

特徵點的擷取和對應是電腦視覺 (computer vision) 領域中的重要關鍵，過去三十年來已有不少人提出相關研究，受制於一些先天條件，至今仍無一完美的解決方案，特徵點對應理論的相關應用廣泛，如視覺導航 (navigation)、物體辨識(object recognition)、物體避碰(obstacle avoidance)···等。

在場景的三維重建 (3D reconstruction) 中，通常需要對所要重建的場景做不同角度的取像，並由這些從不同角度所拍攝的影像中找出對應的特徵點，再藉由投影幾何理論求出二維影像點在實際三維場景中的正確位置。由於目前特徵點對應演算法並無法解決所有的問題，本篇論文的目標是分析並改進特徵點的對應，希望藉由後續處理步驟找出正確的特徵點對應，以應用於室內場景的三維重建。

我們從兩個不同的角度拍攝場景，取得一對立體影像 (stereo images)，再利用特徵點擷取演算法分別找出兩張影像的特徵點，以利後續的特徵點對應，目前特徵點對應的演算法主要可歸納為以下四類：

- 利用極線限制 (**epipolar constraint**) 進行特徵點對應：極線限制是說若給一個左影像平面的點 m ，在右影像平面必對應一條極線 l' ，且 m 的對應點 m' 必落在 l' 上，利用極線限制可以減少特徵點對應的搜尋，我們也將利用極線限制驗證對應的正確性[1]。
- 利用灰階比對 (**block matching**) 進行特徵點的對應：若左右影像每一個像素皆能正確對應則可以得到影像中不同物體的位移距離，具有相同位移距離的像素表示在實際場景中深度相同。相關係數的計算方法以 SSD error (Sum of squared differences) 最為常見，如(1.1)式，其中

I_l 與 I_r 分別為左影像與右影像的相關係數視窗，計算時分別以不同大小的距離 d 代入，所得最小相關係數值的 d 值即為此像素的水平位移距離。

$$C(x, y, d) = \frac{\sum_{(\xi, \eta)} [I_l(x + \xi, y + \eta) - I_r(x + \xi + d, y + \eta)]^2}{\sqrt{\sum_{(\xi, \eta)} I_l(x + \xi, y + \eta)^2 \sum_{(\xi, \eta)} I_r(x + \xi + d, y + \eta)^2}} \quad (1.1)$$

此方法的缺點是：若相關係數視窗在影像中的位置恰位於景深(depth)不同之處，相關係數容易因為取像的角度不同而產生較大變化，此時對應的結果便易受到景深的影響。

- 利用特徵比對 (**feature matching**) 進行特徵點的對應：特徵比對的原理是希望尋找影像中一些較為明顯且具代表性的特徵，例如：線段、邊界、輪廓…等，利用這些已知特徵來做兩張影像裡特徵點的對應。特徵比對在下列三種情況發生時，會增加對應的難度：(a)若影像裡無足夠的特徵以供比對(b)影像裡重複或相似的特徵太多(c)因為相機本身的因素，使影像產生扭曲變形。
- 利用非參數法(**non-parametric measures**)進行特徵點的對應：第一個非參數法是 rank transform，首先定義一個 3×3 的小視窗，分別在左右影像裡移動，同時記錄像素周圍灰階值較像素本身灰階值小的數目，若兩個數字相同，則可能為對應的特徵點。第二個非參數法是 census transform，同樣的定義一個 3×3 的小視窗，分別在左右影像裡移動，將像素周圍的灰階值以二進位碼之形式串接，比較兩個字串的漢明距離(Hamming distance)，若兩字串之漢明距離很小，可能為對應的特徵點。其他常見的非參數法還有 median filter、mode filter、rank filter…等。

特徵點對應目前仍然沒有通解，主要是受限於以下限制。

- **取像失真 (photometric distortions)**: 在取像的過程中常因相機或光線等因素使得左右影像的顏色、灰階等特徵失真，或是相機使影像邊緣產生扭曲變形，導致增加對應的困難。
- **投影失真 (projective distortions)**: 由於相機成像是從三維場景投射至二維影像，在投射的過程中就可能使特徵在影像中造成形變，如此在三維場景中的特徵經由不同視角所取得的影像可能因為在二維影像產生形變而無法對應。又因為投影必然會損失深度資訊，故此困難點也很難避免。
- **模稜兩可的比對判斷 (ambiguity)**: 在二維影像比對的過程中經常會有一些特徵對應是以人的觀點也不一定能夠正確分辨出來的，可判斷為相似，也可判斷為不相似。此種先天限制造成了特徵點對應的不穩定。
- **遮蔽 (occlusions)**: 左右影像可能會有部分特徵點因相機的取像角度或位置而導致某些特徵點被遮蔽，使得特徵點只出現在其中一張，在另外一張並無出現，造成在尋找特徵點對應時產生錯誤對應的機率會提高。

本篇論文的主要觀念是由 Harris 演算法找出室內場景中的角點 (corner) 作為特徵點，並分析 Harris 過程裡產生的影像梯度，找出每一角點周圍灰階值的梯度方向 (gradient direction)，並以 9 種顏色代表 9 種不同方向，我們稱之為彩色碼，利用彩色碼做為角點具有之特徵，並和已知對應組合的相對關係相比，以建立特徵點對應，最後透過極線限制改善特徵點的對應關係，以利後續進行室內場景的三維重建。

本篇論文一共分為八章，第一章為簡介，敘述本篇論文的研究動機，並對前人所提出的特徵點對應演算法和存在已久的困難點做分類，第二章

是相關文獻探討，簡述目前改進特徵點對應的方法，第三章為基本矩陣與極線幾何，描述左右影像之間存在的幾何關係，第四章為特徵點的擷取與對應，介紹目前常見特徵點的擷取與對應的演算法，接著利用彩色碼、距離限制、雙向對應一致性與已知對應點的相對關係來建立特徵點對應，第五章為計算基本矩陣，在獲得特徵點的對應關係後，我們可以利用這些對應點求取基本矩陣，本章將介紹求取基本矩陣的方法，第六章為利用極線幾何改善特徵點之對應，我們提出兩種限制以改進特徵點的對應，以獲得正確之基本矩陣，第七章為實驗結果，分別對不同的場景加以分析，第八章為結論，說明此方法的特性以及未來可繼續深入探討的地方。



第二章 相關文獻探討

基本矩陣能完整的表達兩張相機投影影像的幾何關係，如何正確計算基本矩陣，是立體視覺的重要問題，在第三章裡將會詳細介紹基本矩陣的特性。

本章將簡介數種估計兩張立體影像間極線幾何之方法，假設已知許多組的特徵點對應，在這些對應中，我們可以發現兩種型態的雜點(outlier) [1]，分列如下：

1. 錯誤的定位誤差(**Bad location**)：特徵點擷取的方法有很多種，在不失一般性的前提下，特徵點位置的錯誤會呈現高斯常態分佈，我們假設特徵點擷取的過程中，大部分的錯誤在一至二個像素之間，少數一些點超過三個像素以上，後者的錯誤會降低基本矩陣計算上的精確性。
2. 錯誤的配對(**False match**)：若在建立特徵點對應的過程中，僅使用特徵點擷取演算法的結果，而未加入一些幾何或影像本身的限制(像是極線限制、灰階比對…等等)，所以許多左右影像特徵點的配對仍有可能錯誤，最常見的原因就是場景裡存在遮蔽、陰影…等等，這些錯誤的配對會影響基本矩陣計算的精確度。

以上兩種 outlier 在計算基本矩陣所用的最小平方法中，會影響基本矩陣的計算結果，一般最小平方法的估計(Least-squares estimators)假設資料中雜訊的平均值為零，這些資料集合將產生一個無偏差的參數估計，根據許多研究指出，此假設不成立時，最小平方法的估計容易產生錯誤。

目前最常被提到較強健的計算基本矩陣的方法有 M-estimators、RANSAC、和 Least-median-of-squares(LMedS)等等，我們將利用下面小節分別討論。

2.1 RANSAC

Fischler 和 Bolles[2]於 1981 年提出 RANSAC 的概念，RANSAC 的全名是 random sample consensus paradigm，是電腦視覺領域中一種有效的參數估計方法。

假設一開始擁有很多資料，其中包含數目不等的雜點和好點(inlier)，一般傳統的參數估計希望使用越多的資料以求得正確的解，並判斷何者為雜點，這和 RANSAC 大不相同，RANSAC 希望儘量用最少組子集合的資料估計參數，(例如：兩個點可構成一條線，七組以上對應點可計算基本矩陣)，取不同的子集合重複計算很多次，以確保 95%(可自訂標準)以上的機率至少有一組子集合只包含好的資料，刪除誤差大於自訂門檻值的資料後，可得到最佳的解，經過以上的計算步驟，將所有非雜點蒐集起來可得到去除所有雜點的資料集。

在理想情況下，每一組子集合都必須考慮，但是這樣會使得計算量過大，所以如何取適當組數的子集合是一個重要的問題，組數太少則資料不具代表性，組數太多則影響計算時間。針對此問題 Fischler 和 Bolles[2]和 Rousseeuw[3]分別提出不同的評估方式，但結果相去不遠，在此我們介紹後者的方法。

若 m 代表子集合數目， p 為每個子集合所含資料數目， ε 是所有資料裡 outlier 的百分比， P 代表至少 P 的機率有一組子集合資料都不包含雜點，則 m 、 p 、 ε 和 P 的關係式可用(2.1)表示，舉例來說，若要求 $P=0.99$ ， $\varepsilon=40\%$ ，則需要取子集合數目 $m=163$ 個。

$$P=1-[1-(1-\varepsilon)^p]^m \quad (2.1)$$

表(2.1)為 $P=0.95$ 時，每組子集合資料數和原資料雜點比例不同的情況下所

需的子集合數。

表 2.1 不同情況下所需的子集合數， $P=0.95$ 。

每組子集 合資料數	所有資料含雜點比例						
	5%	10%	20%	25%	30%	40%	50%
2	2 個	2 個	3 個	4 個	5 個	7 個	11 個
3	2 個	3 個	5 個	6 個	8 個	13 個	23 個
4	2 個	3 個	6 個	8 個	11 個	22 個	47 個
5	3 個	4 個	8 個	12 個	17 個	38 個	95 個
6	3 個	4 個	10 個	16 個	24 個	63 個	191 個
7	3 個	5 個	13 個	21 個	35 個	106 個	382 個
8	3 個	6 個	17 個	29 個	51 個	177 個	766 個

使用者利用 RANSAC 進行資料估計時，必須先設誤差的門檻值，誤差的門檻值同時代表了可容忍的誤差和所需的精確性。

Torr 和 Murray[4]在 1993 年第一次將 RANSAC 用來估計基本矩陣，若左影像和右影像特徵點的齊次座標分別為 (x_i, y_i, ξ) 、 (x'_i, y'_i, ξ) ， n 為子集合裡對應的特徵點組數， Z 是 $n \times p$ 的資料矩陣，而 z 為 Z 中列方向的分量，其 $z = (x_i'x_i \quad x_i'y_i \quad x_i'\xi \quad y_i'x_i \quad y_i'y_i \quad y_i'\xi \quad x_i\xi \quad y_i\xi \quad \xi^2)^T$ ，且 W 為對角化矩陣，代表分配給每組對應的權重，則 $M = Z^T Z$ 為 $p \times p$ 矩陣。

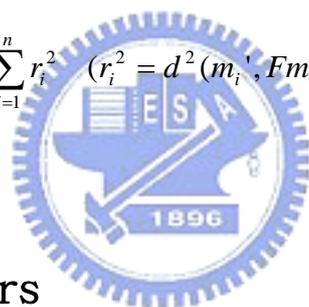
舉例來說，如果用七組對應點計算基本矩陣，則 $n=7$ 、 Z 可寫成 7×9 的矩陣、 W 為 7×7 矩陣、 M 為 9×9 矩陣，如(2.2)式。經由 M 可以定義一個符合這 n 組對應點特性的參數集合： $\alpha F_1 + (1-\alpha)F_2$ ，基本矩陣 F 可寫為 F_1 、 F_2 的線性組合，又 F 的行列式值必須為零，由 $\det|\alpha F_1 + (1-\alpha)F_2| = 0$ ，可得到一個 α 的三次方程式，所以 α 可解出一或三個實根，代入 α 可求得

一或三個 F 。

$$Z = W \begin{bmatrix} x_1'x_1 & x_1'y_1 & x_1'\xi & y_1'x_1 & y_1'y_1 & y_1'\xi & x_1\xi & y_1\xi & \xi^2 \\ \bullet & \bullet \\ \bullet & \bullet \\ \bullet & \bullet \\ x_7'x_7 & x_7'y_7 & x_7'\xi & y_7'x_7 & y_7'y_7 & y_7'\xi & x_7\xi & y_7\xi & \xi^2 \end{bmatrix} \quad (2.2)$$

其中 r_i 代表每一組對應點的誤差，則利用特徵分析，我們可找出矩陣 M 共有 9 個特徵值，由小到大依次為 $\lambda_1, \dots, \lambda_9$ ，其中最小的特徵值 λ_1 和誤差 r_i 的關係可寫成(2.3)式， λ_1 同時表示此子集合裡七組對應點的總誤差，最後比較 λ_1 和 RANSAC 所預設的門檻值，以決定這個子集合資料和基本矩陣的好壞。

$$\lambda_1 = \sum_{i=1}^n r_i^2 \quad (r_i^2 = d^2(m_i', Fm_i) + d^2(m_i, F^T m_i')) \quad (2.3)$$



2.2 M-estimators

M-estimators 的全名是 maximum likelihood estimators，在理想狀況下，若資料本身的誤差呈現高斯分佈、且平均值為零時，利用最小平方法計算的效果等同於 maximum likelihood estimators。

M-estimators 是由 Maronna、Huber、Hampel 等人所提出，目的是希望在資料呈現非高斯分佈的情況下，遵循 maximum likelihood 式子推導，以獲得每個資料的最佳權重。

假設 r_i 代表第 i 組資料的誤差，也就是特徵點擷取演算法所得之特徵點位置和真實位置的差異。一般最小平方法的目標函數是 $\min \sum_i r_i^2$ ，如

果資料中有 outlier 存在，結果將變得不穩定，M-estimators 為降低每一組 outlier 的影響力，把誤差的目標函數換成(2.4)式。

$$\min \sum_i \rho(r_i) \quad (2.4)$$

$$\rho(r_i) = (\omega_i r_i)^2 \quad (2.5)$$

其中權重函數 ω_i 是經過計算後分配給每一組資料的權重， ρ 是一個對稱性、僅在正向有定義且最小值為零的函數，接下來我們可以用反覆式權重最小平方法解決這個問題。首先令 $P = [p_1, \dots, p_p]^T$ 為要估計的參數向量，根據函數 $\rho(r_i)$ ，我們利用 M-estimators 產生 p 個非線性方程式，如下

$$\sum_i \psi(r_i) \frac{\partial r_i}{\partial p_j} = 0 \quad \text{for } j = 1, \dots, p \quad (2.6)$$

在這裡導函數 $\psi(r_i) = d\rho(x)/dx$ 稱為影響函數(influence function)，而權重函數 $\omega_i = \psi(x)/x$ ，則方程式(2.6)式可改寫成(2.7)式，而(2.7)式可用反覆式權重最小平方法求解，其中 k 為迴圈數目，如(2.8)式。

$$\sum_i \omega(r_i) r_i \frac{\partial r_i}{\partial p_j} = 0 \quad \text{for } j = 1, \dots, p \quad (2.7)$$

$$\min \sum_i \omega(r_i^{(k-1)}) r_i^2 \quad (2.8)$$

常用的權重函數有很多種，像是 Tukey(1977)、Kumar and Hanson (1994)……等等，Olsen 和 Loung 於 1992 年的論文中提到，Huber 的權重函數(2.9)式常被用來估計極線幾何，其中 $\sigma = \text{median } r_i / 0.6745$ 。

$$\omega_i = \begin{cases} 1 & |r_i| \leq \sigma \\ \sigma / |r_i| & \sigma < |r_i| \leq 3\sigma \\ 0 & 3\sigma < |r_i| \end{cases} \quad (2.9)$$

影響函數 $\psi(x)$ 代表單一資料對整體參數估計的影響，舉例來說，最小平方法的 $\rho(x) = x^2/2$ ，則影響函數 $\psi(x) = x$ ，也就是說單一資料對整體估計

的影響隨著 error 大小而線性增加，這證明了最小平方法 $\min \sum_i r_i^2$ 的不強健，因為單一資料對整體的影響很大。

找出權重之後，我們可以判別哪些是好點哪些是雜點，並用好點來計算基本矩陣。M-estimators 能夠有效排除錯誤的定位誤差所產生的雜點，但對錯誤的配對產生的 outlier 則沒什麼效果。

2.3 Least Median of Squares (LMedS)

LMedS 的全名是 Least Median of Squares，希望找出誤差平方中位數的最小值，如(2.10)式，是一種非線性最小化的估計參數方法。LMedS 和 RANSAC 採用相同的方式，亂數選取不同子集合的資料並加以分析，且無直接以一個數學式的型態表示。

$$\min \text{median } r_i^2 \quad (2.10)$$

LMedS 和 M-estimators 比起來，對錯誤的定位誤差和錯誤的配對這兩種型態的 outlier 都很有效，而 M-estimators 僅能排除錯誤的位置所產生的 outlier。若給定 n 組對應點 $\{ (m_i, m'_i) \mid i = 1, \dots, n \}$ ，LMedS 的方法主要可分為五個步驟，分列如下。

1. 從 n 組對應點取出 m 個子集合，每一個子集合包含七組不同的對應點（最少只需要七點就能決定極線幾何）。
2. 利用 2.1 節提到的方法對這 m 個子集合計算其基本矩陣 F_j ，每一個子集合可找出一或三個基本矩陣。
3. 對每個基本矩陣 F_j ，找出左右影像特徵點與極線距離平方和的中位數，並用 M_j 代表，如(2.11)式。

$$M_J = \text{median } d^2(m_i', F_J m_i) + d^2(m_i, F_J^T m_i') \text{ for } i = 1, \dots, n \quad (2.11)$$

4. 找出所有 M_J 中最小的一個，令為 M_{JJ} ，並保留相對應的子集合和 F_J ，值得一提的是，我們同樣可用權重式最小平方法解決 LMedS 的問題。標準差 σ 和 M_{JJ} 可由(2.12)式計算出。

$$\sigma = 1.4826 \times [1 + 5/(n - p)] \times \sqrt{M_{JJ}} \quad (2.12)$$

有了標準差之後，根據(2.13)式，我們可以對每一組對應點都分配一個權重，而 outlier 的權重為零。

$$\omega_i = \begin{cases} 1 & \text{if } r_i^2 \leq (2.5\sigma)^2 \\ 0 & \text{otherwise} \end{cases} \quad (2.13)$$

$$r_i^2 = d^2(m_i', F m_i) + d^2(m_i, F^T m_i') \quad (2.14)$$

5. 利用(2.15)式的權重式最小平方法修正基本矩陣 F 。

$$\min \sum_i \omega_i r_i^2 \quad (2.15)$$



2.4 前三種方法之比較

下面舉出 RANSAC、M-estimators、LMedS 等三種方法的同異處。

- RANSAC 必須事先自訂門檻值(threshold)，而 LMedS 的門檻值(可經由標準差 σ 求得)是由演算法計算過程中產生。
- 在全部資料的 outlier 少於 50%時，LMedS 的效果比 RANSAC 好，反之則效果較差。
- 如果存在 n 組對應點，M-estimators 需要估算 n 個非線性方程式，即使用 Huber 建議的權重式最小平方法，門檻值(在此為資料的標準差 σ)可由資料誤差的中位數求得，在計算上仍然很複雜。

- 若單純使用最小平方法估計資料，由於最小平方法的總誤差易受單一資料的影響，所以總誤差較大，這同時也證明了最小平方法的缺點。
- 不管標準差知不知道的前提下，單純使用 LMedS 的平均效果最好。
- 在資料的 outlier 高達 90% 的例子中，RANSAC 是所有方法表現最好的。
- M-estimators 的結果好壞，和事先用最小平方法計算出的 σ 有關，如果我們先用 RANSAC 找出標準差 σ ，再利用 M-estimators 求解，最後的結果將會比 RANSAC 還要好一些，雖然效果並不明顯，但已足以影響極線幾何的精確性。

2.5 Case deletion diagnostics

Case deletion diagnostics[20]的概念源自於分析每一組資料對整體的影響，假設資料誤差的標準差為已知且精確，此方法效果不錯，但若不知道標準差，則效果不好。前述三種方法能夠忍受資料裡面有較多的錯誤，但所花的時間也較多，如果資料量不多，case deletion diagnostics 是不錯的選擇，演算法可分成下列幾個步驟，分列如下：

1. 對每組特徵點對應的權重 ω_i 設為 1，亦即假設所有資料都是正確的。
2. 進入(a)~(h)的迴圈，直到 $\sum_i d_i^2 < \sigma$ 為止 (σ 為預設的門檻值)， d_i 的計算方式如(h)式。

- (a) 首先對每一組資料乘上權重 ω_i 。
- (b) 將所有特徵點對應組合以資料矩陣的型態表示，利用 2.1 節提到的特徵分析，我們可找到最小的特徵值 λ_1 ，如(2.3)式，若 λ_1 其對應的特徵向量為 u_1 ，我們可以求得基本矩陣 $F = u_1$ 。

- (c) 基本矩陣 F 透過 SVD 分解，加上秩為二之限制可得 \hat{F} 。
- (d) 計算 $r_i = m_i^T \hat{F} m_i$ 。
- (e) 計算每一組特徵點對應的影響值 T_i 。
- (f) 刪除 T_i 最大的一組對應點。
- (g) 若左影像的第 i 組極線其 x 、 y 項的係數為 a_i 和 b_i ，右影像的第 i 組極線其 x 、 y 項的係數為 c_i 和 d_i ，則 $\omega_i = \left(\frac{1}{a_i^2 + b_i^2 + c_i^2 + d_i^2}\right)^2$ 。
- (h) $d_i = \omega_i r_i$ 。

其他根據 RANSAC 演變而來的方法還有 Virtual Parallax Method 和 Linear Subspace Method[15] 等等，以上相關文獻的探討可以讓我們了解許多人關於改善特徵點對應的研究，在第六章我們將以這些演算法為基礎，提出一個改進特徵點對應的方法，希望能找出兩張影像間正確的對應組合。

第三章 基本矩陣與極線幾何

3.1 極線幾何

兩張不同視點的立體影像間存在極線幾何(epipolar geometry)的關係，這個幾何關係可用左右影像平面和以兩相機中心連線為軸旋轉的所有極線平面(epipolar plane)的交集來表示。要了解極線幾何，通常會從兩張立體影像之間的特徵點對應開始。

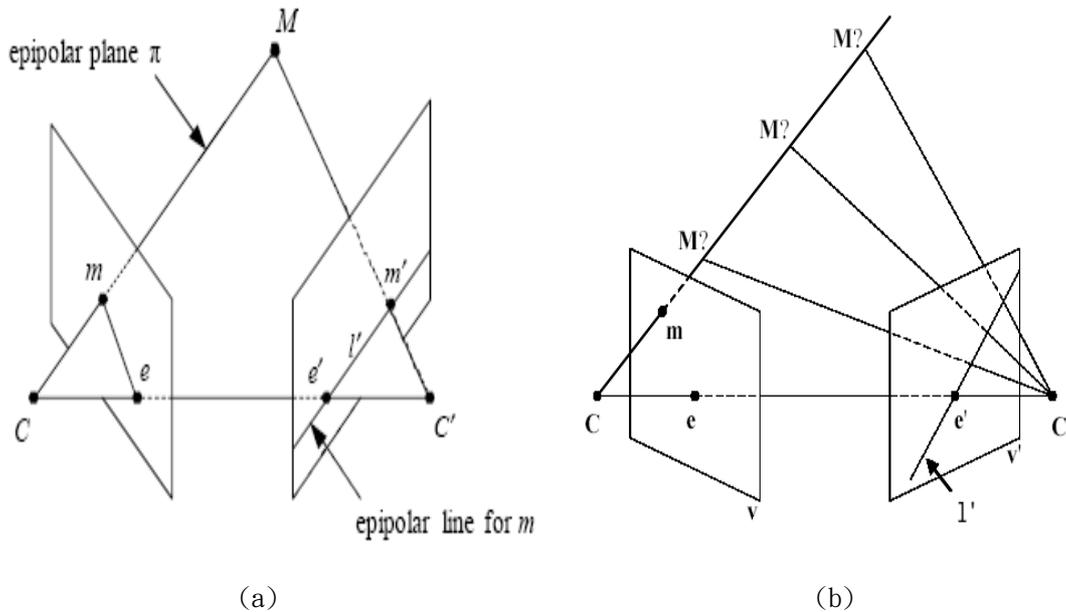


圖 3.1 (a)特徵點對應 (b)極線限制

如圖 3.1(a)所示，假設空間中存在一點 M 及兩台相機， C 和 C' 分別為兩台相機的相機中心， P 和 P' 是兩台相機的投影矩陣，場景點 M 透過投影矩陣分別投影到左影像平面 V 上的 m 和右影像平面 V' 上的 m' 。 m 和 m' 稱為左右影像平面的一組對應點，投影點 m 、 m' 和 M 的關係可表示為 $s_1 m = PM$ 和 $s_2 m' = P'M$ (s_1 、 s_2 為比例常數)，其中 m 、 m' 和 M 均以齊次座標表示，且 m 、 m' 、場景點 M 和兩相機中心共平面 Π ，此平面稱為極線平面。隨著

場景點 M 位置的改變，所有的極線平面將以兩相機中心連線為軸做旋轉，這些極線平面的集合又稱為 **epipolar pencil**。

從另一個角度思考，如果知道左影像的點座標 m ，要如何在右影像找到正確的對應點？已知場景點 M 在 C 和 m 的射線上（同理亦在 C' 和 m' 射線上，且兩射線共平面， M 為兩射線的交點），由射線 \overline{Cm} 和兩相機中心連線可求出極線平面 Π ，極線平面分別交左影像平面 V 和右影像平面 V' 於直線 l 和直線 l' ， l' 同時為射線 \overline{Cm} 在右影像上的投影，也就是所謂的極線 (epipolar line)。

我們可得到一個結論，給定一個左影像平面的點 m ，在右影像平面必對應一條極線 l' ，且 m 的對應點 m' 必落在 l' 上，這種對應關係稱為極線限制 (epipolar constraint)，如圖 3.1(b) 所示，在立體影像特徵點對應的過程中，利用極線限制找 m 的對應點，可以不用搜尋整張右影像，只要找對應的極線上有無對應點即可。

Hartly 等人於 [6] 中提到，兩相機中心 C 和 C' 的連線分別交兩影像平面 V 和 V' 於點 e 和 e' ，這兩個點又稱為極點 (epipole)。極點也可以看成是一個相機中心對另一個相機中心的投影，也就是

$$e' = P'C \quad (3.1)$$

左影像平面 V 上不同的點會對應右影像平面 V' 上不同的極線，在對應正確的情況下，這些極線都會通過極點 e' ，所以極點也可以視為任兩條極線的交點，如圖 3.2 所示。



圖 3.2 極點可視為任兩條極線的交點

由投射投影幾何可知，左影像平面的點 m 和它在右影像平面對應的極線 l' 有以下的線性關係。

$$l' = Fm \quad (3.2)$$

其中 F 稱為基本矩陣，關於基本矩陣的特性將在下一節介紹。因左影像平面的點 m 在右影像的對應點 m' 會落在極線 l' 上，所以 m' 的齊次座標和極線 l' 方向向量的內積為零，亦即，

$$m'^T l' = m'^T Fm = 0 \quad (3.3)$$

考慮對稱性，將(3.3)式左右同步轉置，可得(3.4)式，所以 m' 在左影像對應的極線 $l = F^T m'$ 。

$$l'^T m' = m'^T F^T m' = 0 \quad (3.4)$$

若所取左影像平面的點恰為極點 e ，則相機中心 C 可取代場景點 M 的角色，射線 \overline{Ce} 交右影像平面 V' 於 e' ， e 所對應的極線 l' 為 Fe ，且極線 l' 在右影像平面 V' 上會退化成極點 e' ，所以極線 $l' = Fe = 0$ 。同理，若右影像平面上的點恰為極點 e' ，由(3.4)式可推得 $l = F^T e' = 0$ 。

3.2 基本矩陣之特性

在電腦視覺的領域裡，基本矩陣(**fundamental matrix**)已被廣泛的討論，尤其是未校正的相機和影像，因為它包含兩張影像對同一物體的幾何資訊，它的應用包括仿射和投影重建(affine and projective reconstruction)、投射不變性之計算(computation of projective invariants)、未校正立體影像特徵點對應(uncalibrated stereo matching)、影像分析、影像校正(image rectification)等等[7]。

基本矩陣是極線幾何的代數表示法，描述每個左影像點和它在右影像

$$F = [e']_x [p_1 \quad p_2] \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} p_2^T \\ -p_1^T \end{bmatrix} [e]_x \quad (3.9)$$

由上式可知兩個極點座標和極線轉換矩陣的四個參數 a 、 b 、 c 、 d 可影響 F 的自由度，所以 F 的自由度為 7，也就是說只要有 7 組以上的對應點就可以解基本矩陣。

3.3 必要矩陣

必要矩陣(**Essential matrix**)[8]是基本矩陣的特別化，一般而言，必要矩陣的發展早在基本矩陣之前，一旦必要矩陣已知，相機矩陣便可由必要矩陣求得，基本矩陣可視為移除已校正像機的一般化必要矩陣。



3.3.1 正規化影像座標

若 A 、 R 、 t 分別代表校正矩陣、旋轉矩陣和平移矩陣，則左右影像的投影矩陣 P 、 P' 可分解為 $P = A[I|0]$ 、 $P' = A[R|t]$ 。以右影像為例，已知 $m' = P'M$ (忽略比例常數)，左右同乘 A^{-1} 可得 $\tilde{m}' = [R|t]M$ ， \tilde{m}' 相當於是右影像點的正規化座標，同理可得左影像點正規化座標 \tilde{m} ，則必要矩陣可寫成

$$E = [t]_x R = R[R^T t]_x = SR \quad (3.10)$$

考慮 $W = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ 、 $Z = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ ，則 $Z = \text{diag}(1,1,0)W$ ， $S = \rho UZU^T$ 為斜

對稱矩陣， U 為正交矩陣。Longuet-Higgins[8]提出 Longuet-Higgins 方程式表示必要矩陣，如(3.11)式，將 $\tilde{m}' = A^{-1}m'$ 、 $\tilde{m} = A^{-1}m$ 代入(3.11)式可導出 $m'^T A'^{-T} EA^{-1}m = 0$ 。因此，必要矩陣和基本矩陣之間的對應關係如(3.12)式。

$$\tilde{m}'^T E \tilde{m} = 0 \quad (3.11)$$

$$E = A'^T F A \quad (3.12)$$

3.3.2 必要矩陣之特性

旋轉矩陣和平移矩陣的自由度各為三，但矩陣長度不固定，矩陣元素可以等比例縮放，所以必要矩陣的自由度為五，必要矩陣較基本矩陣少了兩個自由度卻多了一些限制。

Faugeras[8]提出下列三個對等的條件再加上必要矩陣秩為二的限制，可確保矩陣 E 為必要矩陣。

1. 必要矩陣經 SVD 分解後，兩個非零奇異值相等且第三個為零。
2. $(R_1^2 + R_2^2 + R_3^2) = 4(\|R_1 \times R_2\|^2 + \|R_2 \times R_3\|^2 + \|R_1 \times R_3\|^2)$ ， R_i 代表 E 的列向量。
3. $\text{trace}^2(EE^T) = 2\text{trace}((EE^T)^2)$ 。

相較於必要矩陣，基本矩陣的限制只有秩為二，但由於座標系的不同，基本矩陣並不全然滿足必要矩陣所有的代數限制。

3.3.3 透過必要矩陣計算相機矩陣

必要矩陣可直接由(3.11)的正規化影像座標計算，也可由(3.12)的基本矩陣關係式獲得。必要矩陣經 SVD 分解則可獲得旋轉和平移矩陣。

若第一部相機投影矩陣 $P = A[I|0]$ ，必要矩陣經 SVD 分解後得到 $E = U \text{diag}(1,1,0)V^T$ ，則旋轉矩陣有兩種可能的解 UWV^T 或 $UW^T V^T$ ，平移矩陣為 U 的最後一行，也有正負號兩種選擇，所以第二部相機的投影矩陣共有四種可能的幾何表示，如圖 3.4。

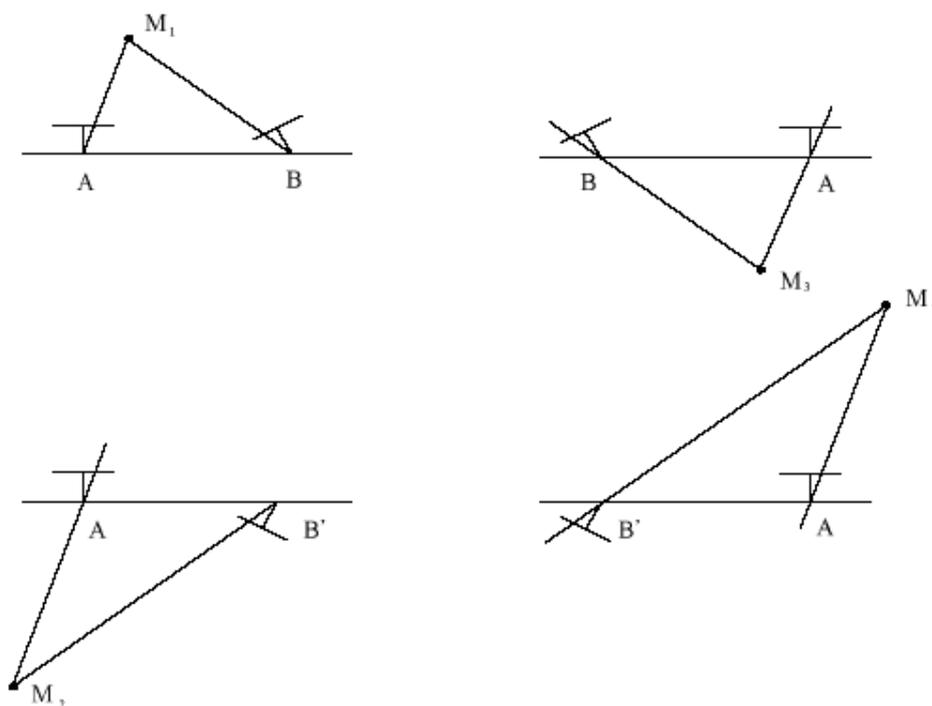


圖 3.4 透過必要矩陣做校正重建後的四種可能解

根據成像原理，重建後的點必須落在兩部相機的前方，所以只要測試一個點是否落在相機前，就可以知道計算出的投影矩陣何者為真，圖 3.4 的例子中，左上角的示意圖才是正解。

第四章 特徵點擷取與對應

特徵點的擷取與對應是電腦視覺相關研究中的一個關鍵部份。舉例來說，大多數影像比對(image matching)、物體辨識(object recognition)、三維重建(3D reconstruction)、相機校正(camera calibration)及運動分析(motion analysis)等應用，都需要在兩張或多張影像上找到特徵點，進而獲得影像間的對應關係。

目前為止仍沒有一個完美的演算法能解決特徵點對應的問題，早期的研究主要針對連續兩張視角變化小或差異不明顯的影像(small baseline algorithm)，最近也有人開始討論適合角度差異較大影像的演算法(wide baseline algorithm)，本篇論文的重點在前者，下面的小節裡我們將分別探討特徵點擷取及對應的演算法。



4.1 特徵點的擷取

角點(corner)是影像中重要的特徵，在本文裡通稱為特徵點，特徵點的擷取在影像分析和電腦視覺領域已有三十年以上的歷史，其應用包括影像檢索(image retrieval)、移動追縱(motion tracking)、場景分析(scene analysis)等等。

特徵點擷取的方法大致可分為兩類，一類為直接分析影像的灰階或色彩，計算其梯度(gradient)或曲率(curvature)來擷取特徵點，大部份特徵點擷取的研究屬於這一種方法，例如Moravec [10]、Harris [11]、Susan [12]，另一類則是先將邊擷取出來，找出擁有最大曲率的邊，或利用邊與邊的交點來擷取特徵點，如[13][14][15]。後者因為計算複雜需要較多時間，且邊的

偵測只在特殊狀況下不受雜訊影響，所以較少被應用。

Susan演算法是利用一圓形的遮罩(mask)在影像中進行灰階比對，圓形遮罩中灰階變化與圓形遮罩面積的比值可分為幾類，若遮罩中相似灰階值的面積約為全部的二分之一，可能為邊，相似灰階值的面積少於二分之一時，可能為角點，當遮罩中大部分的面積都是相似灰階時，表示目前所在區域灰階變化緩和。

在同一類的方法中，Harris演算法主要概念是從Moravec[10]而來，方法是在利用一方形遮罩，根據遮罩移動方向與灰階變化的程度，決定所在位置為邊或角點，若遮罩所在區域灰階值變化平坦，則不管遮罩往哪個方向移動，灰階值變化皆不明顯，如圖4.1(a)；若遮罩沿著邊的方向移動，則灰階變化較不明顯，當遮罩位於影像中邊的位置時，遮罩移動方向若與邊垂直，灰階將會產生明顯的變化，如圖4.1(b)；若遮罩恰好位於角點上，不管移動方向為何，皆會有劇烈的灰階改變，如圖4.1(c)。

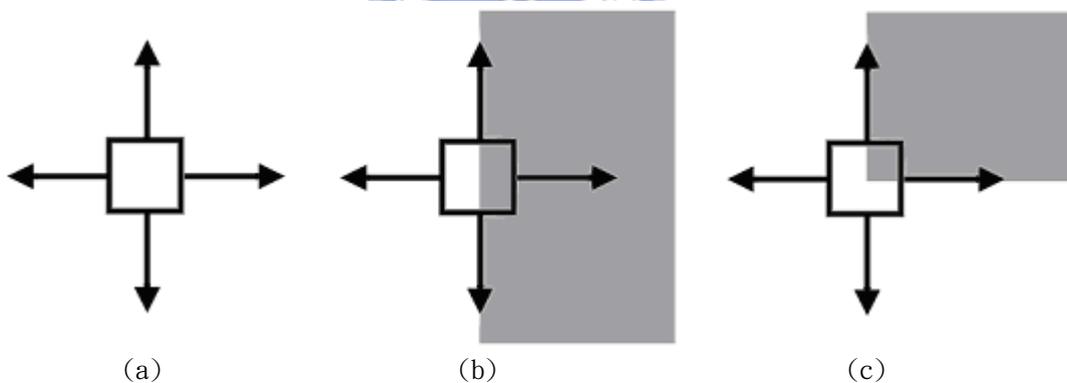


圖 4.1 方形遮罩在影像中移動 (a)遮罩在灰階變化平坦的影像區域中移動 (b)遮罩在邊或線的影像區域中移動 (c)遮罩在具有特徵點的影像區域中移動

其他的方法像是Zheng計算影像梯度(gradient)、Harr利用小波轉換(wavelet transform)、Jolion利用解析度的調整和明暗對比擷取特徵點，各有利弊，圖4.2為各種特徵點擷取演算法結果的比較。



圖 4.2 各種不同特徵點擷取演算法之結果

Moravec 觀察圖 4.1 後，提出(4.1)式來描述方形遮罩在不同方向移動後其灰階變化總和，其中 $W_{u,v}$ 表示方形遮罩，若影像點落在此遮罩之外，則 $W_{u,v}$ 的值為 0，落在遮罩內則 $W_{u,v}$ 之值為 1； I 代表影像， $I_{u,v}$ 為影像點在座標 (u, v) 的灰階值； x 、 y 分別為在 x 方向以及 y 方向的運動向量。

$$E_{x,y} = \sum_{u,v} w_{u,v} |I_{x+u,y+v} - I_{u,v}|^2 \quad (4.1)$$

以泰勒展開式展開(4.1)式後可得(4.2)式，其中 X 和 Y 為影像 I 在 x 方向與 y 方向的梯度，也就是影像 I 分別與 $[-1 \ 0 \ 1]$ 和 $[-1 \ 0 \ 1]^T$ 做旋積 (convolution) 所得到的結果，如(4.3)式、(4.4)式。

$$E_{x,y} = \sum_{u,v} w_{u,v} [xX + yY + O(x^2, y^2)]^2 \quad (4.2)$$

$$X = I \otimes (-1, 0, 1) = \partial I / \partial x \quad (4.3)$$

$$Y = I \otimes (-1, 0, 1)^T = \partial I / \partial y \quad (4.4)$$

分析影像梯度之結果可了解每個特徵點所具有的梯度方向，也可作為特徵點擷取的依據，在這一章接下來的小節裡，我們將介紹更多關於影像梯度的細節。忽略(4.2)式的高次項後，我們可得到(4.5)式。

$$E_{x,y} = Ax^2 + 2Cxy + By^2$$

$$\text{and} \quad A = \left(\frac{\partial I}{\partial x} \right)^2 w_{u,v} \quad B = \left(\frac{\partial I}{\partial y} \right)^2 w_{u,v} \quad C = \left(\frac{\partial I}{\partial x} \right) \left(\frac{\partial I}{\partial y} \right) w_{u,v} \quad (4.5)$$

觀察圖(4.2)後，整體而言，由於 Harris 演算法效果穩定且速度較快，所以被許多相關研究廣為採用，在原來的演算法中，Moravec 利用點和方形遮罩的相對位置以決定 $w_{u,v}$ 之值為 0 或 1，Harris 繼承了 Moravec 演算法的主要概念，但為了減少影像雜訊的影響，Harris 重新定義 $w_{u,v}$ 為高斯函數，而非原來的二值化函數，如(4.6)式，將(4.6)式代入(4.5)式並以矩陣表示，可推得(4.7)式。

$$w_{u,v} = e^{-(u^2+v^2)/2\sigma^2} \quad (4.6)$$

$$E_{x,y} = (x, y)M(x, y)^T \quad \text{and} \quad M = \begin{bmatrix} A & C \\ C & B \end{bmatrix} \quad (4.7)$$

觀察(4.5)式，Moravec 認為 $E_{x,y}$ 之值越大代表灰階變化越明顯，且 $E_{x,y}$ 和矩陣 M 的兩個特徵值(令為 α 、 β)有很大的關係，可分為以下三種情形討論之。

- 若 α 、 β 之值都很小，則表示此遮罩範圍內的灰階變化不明顯。
- 若 α 、 β 之值一大一小，則表示此遮罩沿著某一個方向移動時灰階變

化較大，也就是在此遮罩範圍內有邊或線的存在。

- 若 α 、 β 之值都很大，則表示此遮罩不管往哪一個方向移動其灰階變化都很明顯，也就是在此遮罩範圍內存在特徵點。

根據以上 Moravec 的概念，Harris 提出下式以表示特徵點和邊的關係，他發現若 R 之值大於 0，則為特徵點；若 R 之值小於 0，則為邊或線上的一點。如此我們可從 R 的大小來判斷，而不用計算特徵值 α 、 β 。

$$R = \text{Det}(M) - k \cdot [\text{Tr}(M)]^2 \quad (4.8)$$

其中 $\text{Det}(M) = AB - C^2 = \alpha\beta$ $\text{Tr}(M) = A + B = \alpha + \beta$

Harris 經過實驗以後，建議常數 k 設為 0.04，為了便於和 Moravec 提出的 (4.5) 式比較，Harris 在 (4.8) 式代入 A 、 B 、 C 、 k ，如此可得

$$R = 0.92AB - C^2 - 0.04(A^2 + B^2) \quad (4.9)$$

在實際的應用裡，我們會對不同的輸入影像設不同的門檻值 R ，當 R 越大，符合條件的特徵點越少；反之則符合的特徵點越多，適當的門檻值可避免特徵點過多或過少的情形，圖(4.3) 是一個特徵點擷取的例子。

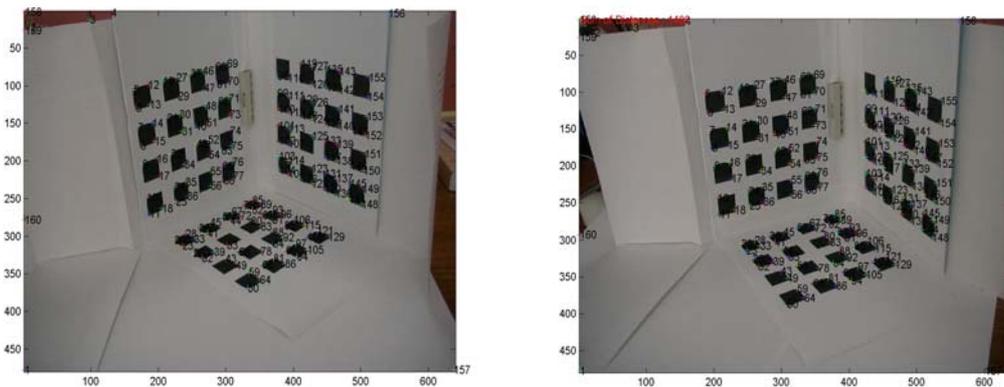


圖 4.3 使用 Harris 演算法擷取一對立體影像之特徵點

4.2 彩色碼

Harris 演算法計算過程裡所得的 X 、 Y 分別為影像在 x 方向和 y 方向的梯度，如(4.3)、(4.4)式，對每一個影像點來說，梯度向量 ∇I 可以表示為(4.10)式。有了梯度向量，我們可以很容易計算出梯度向量的方向，也就是梯度向量 ∇I 的餘切函數，如(4.11)式。

$$\nabla I = (X, Y) = \left(\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y} \right) \quad (4.10)$$

$$\cot \theta = \frac{\nabla I_x}{\nabla I_y} = \frac{\partial I / \partial x}{\partial I / \partial y} \quad (4.11)$$

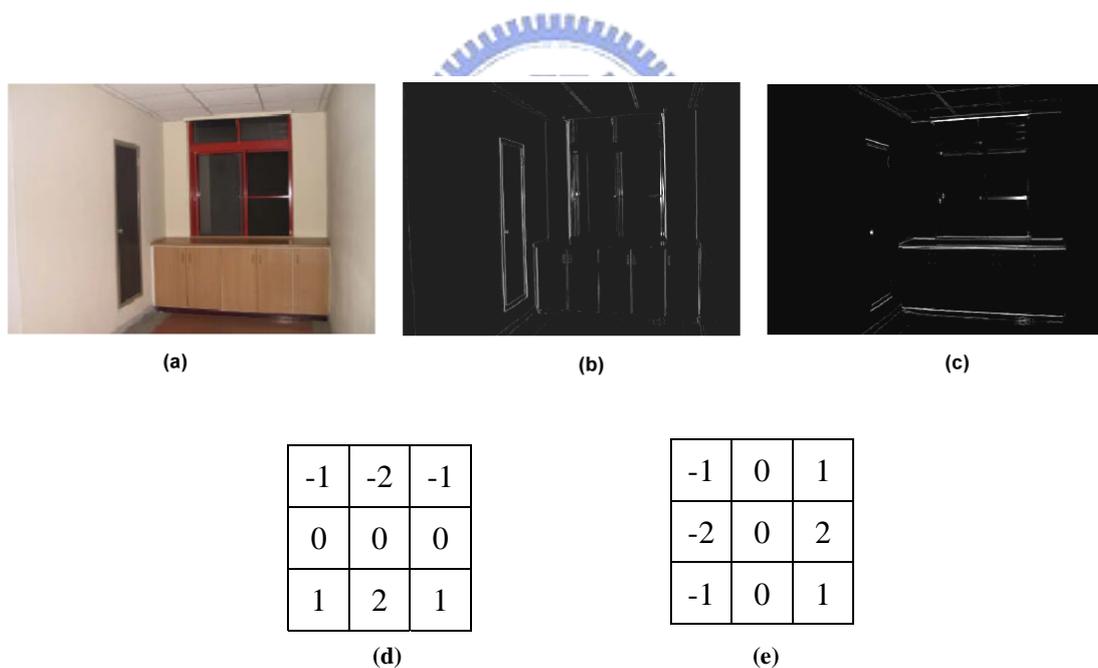


圖 4.4 (a)原始影像 (b)做 $\partial I / \partial x$ 後的影像 (c)做 $\partial I / \partial y$ 後的影像
(d)Sobel operator G_x (e)Sobel operator G_y

從圖 4.4(b)和 4.4(c)可以看到原始影像中 x 方向和 y 方向的梯度變化，這種效果其實和 Sobel operator 的結果非常類似，Sobel operator 常用於邊的偵測。

如此一來，利用 Harris 演算法，我們不但能夠擷取影像的特徵點，也

可得到每一像素周圍灰階變化的情形和方向，分析周圍灰階的改變，可作為特徵點對應的線索，若場景點 M 分別投影在左右影像的點 m 、 m' ，則 m 、 m' 可視為一組對應的特徵點，理論上 m 、 m' 周圍的灰階變化強度和方向應該相去不遠，這個線索將是特徵點對應的重要依據。

若影像點位於邊與邊的交點，或者是邊或線上，這些點通常會有較大的灰階變化且梯度方向也較明顯，和 Harris 演算法相同，為了避免雜訊的影響，也必須設適當的門檻值，在這裡只保留 $\sqrt{\frac{\partial I^2}{\partial x} + \frac{\partial I^2}{\partial y}}$ 大於 10 的影像點。

影響，也必須設適當的門檻值，在這裡只保留 $\sqrt{\frac{\partial I^2}{\partial x} + \frac{\partial I^2}{\partial y}}$ 大於 10 的影像點。

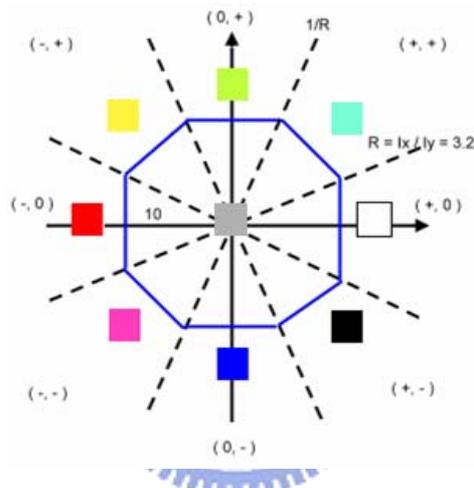


圖 4.5 以不同顏色代表不同的梯度方向

刪除影像灰階變化較不明顯和可能為雜訊的點後，便可針對梯度向量的方向進行分析，利用(4.11)式計算梯度向量的餘切函數後，我們將所有的梯度方向歸納為八類，如圖(4.5)，按反時針的順序依次為 0° 、 45° 、 90° 、 135° 、 180° 、 225° 、 270° 、 315° ，其中中間灰色的部分為 x 方向梯度與 y 方向梯度值皆小於 10 的區域，由於此區域中灰階變化並不明顯，或是雜訊本身影響了梯度計算，使得我們無法判別梯度方向；以 90° 的區域為例，它代表的意義是 y 方向的梯度大於 10 且 y 方向梯度和 x 方向梯度的比值大於 3.2，也就是說 y 方向的灰階變化比 x 方向明顯，此時像素周圍灰階分佈由上至下呈現深到淺；以 135° 的區域為例，梯度向量的長度大於 10 且 x 方向梯度和 y 方向梯度的比值需小於 -0.3，大於 -3.2，此時像素周圍灰階分佈由上至下呈現深到淺、由左至右呈現淺到深的變化，如此我們可用不同

的顏色代表梯度方向的改變。

利用 Harris 演算法擷取圖 4.6(a)的特徵點，再用圖(4.5)的結果對這些特徵點的梯度方向分類，可得圖 4.6(b)，觀察圖中的梯度變化或顏色改變，我們可以發現同一個邊或線上的點，其梯度向量之方向及顏色都很相似，由於每一個特徵點其灰階變化和梯度方向不盡相同，在特徵點的對應上，利用本節的方法加以分析將可順利找出對應的特徵點，這樣的方法我們稱之為彩色碼(color code)。

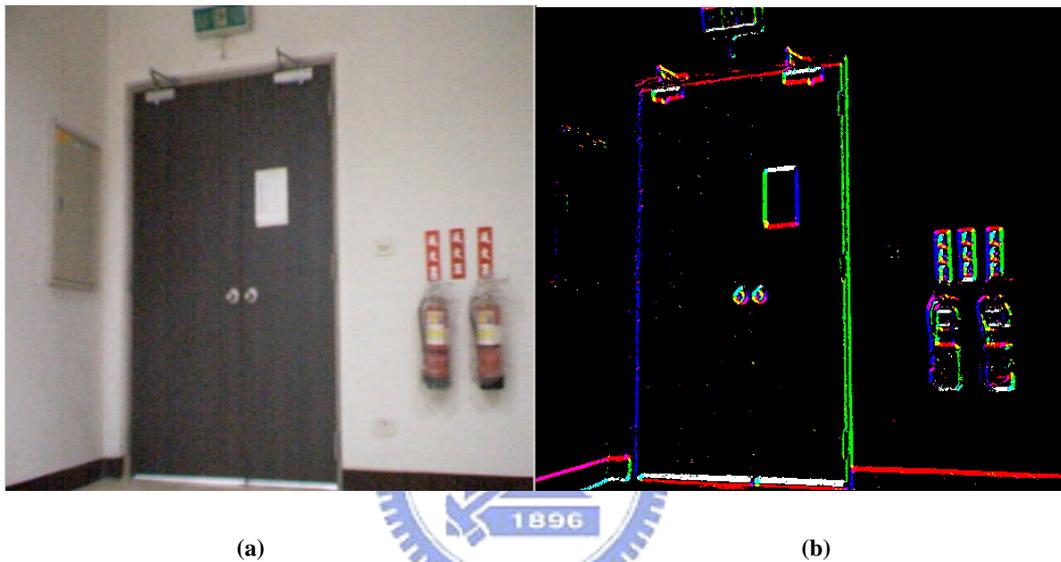


圖 4.6 (a)原圖 (b)將影像梯度方向分類

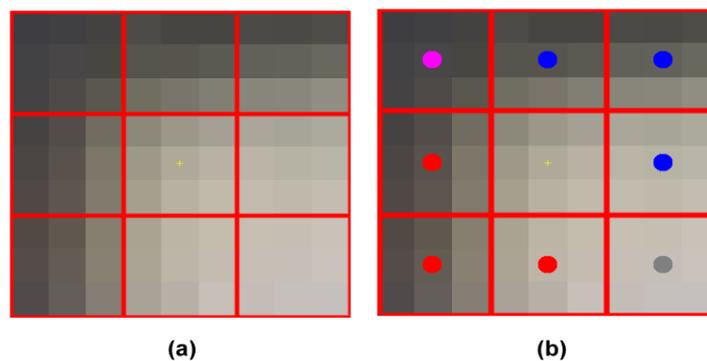


圖 4.7 (a)特徵點四周的八個 3×3 小視窗 (b)以彩色碼代表每個視窗中的梯度方向

在 Harris 演算法擷取特徵點的同時，每一個像素的梯度方向也被計算出來，我們在特徵點的四周建立八個 3×3 的視窗，如圖 4.7(a)，每一個小視窗包含九個像素，由於每一個像素對應一個梯度方向，我們可以找出這

小視窗裡相對多數的梯度方向代表整個視窗的梯度方向，這樣一來，特徵點其四周八個小視窗的梯度方向可以用彩色碼表示，如圖 4.7(b)。

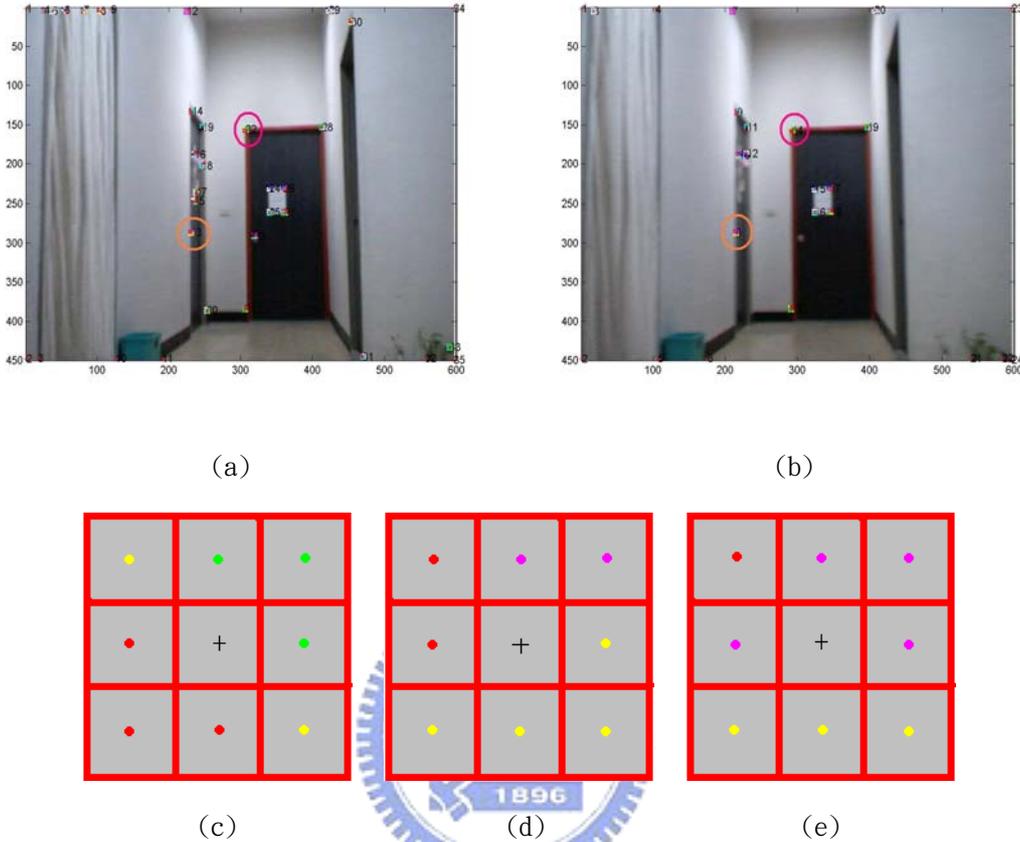


圖 4.8 (a)左影像 (b)右影像 (c)左右影像紅色圓圈特徵點之彩色碼 (d)左影像橘色圓圈特徵點之彩色碼 (e)右影像橘色圓圈特徵點之彩色碼

以實際的室內場景為例，圖 4.8(a)是左影像、圖 4.8(b)為右影像，橘色圓圈和紅色圓圈分別是兩組對應的特徵點，紅色圓圈之特徵點皆擁有相同的彩色碼，如圖 4.8(c)，左影像橘色圓圈和右影像橘色圓圈之特徵點經過量化梯度方向之後，彩色碼的分佈分別為圖 4.8(d)和圖 4.8(e)，由圖可知彩色碼有兩個方向不同，借由彩色碼的比對，可用於特徵點的比對。

分析目前各種特徵點擷取演算法，樣板比對在找出特徵點後，必須對每一組候選點都計算相關係數(correlation)，Zheng[1]在找出特徵點後，對每一組候選點計算對應強度(strength of the match)，兩者都需要額外的計算，而彩色碼可由 Harris 計算過程中的梯度方向產生，若利用彩色碼的概念在特徵點的對應上，將能獲得快速且正確的對應。

4.3 立體影像特徵點的對應

經過前小一節之後，我們獲得彩色碼的特徵，接下來可經由彩色碼獲得特徵點的對應。首先利用距離的限制挑選出在合理距離內的候選點，由於每一候選點皆對應一個彩色碼，根據彩色碼的比對結果便可以決定在這些候選點之中哪一個是最佳的對應點，最後加入雙向對應一致性加強對應正確的機率。

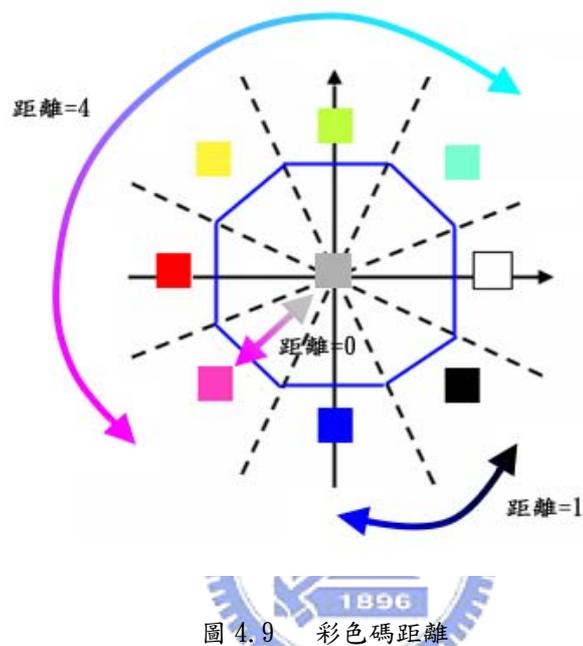
4.3.1 彩色碼之特性

觀察並分析圖 4.8 左右影像對應點的彩色碼，我們可以發現就算是互相對應的特徵點，彩色碼也不一定完全相同，主要原因在於(1)Harris 演算法所獲得的特徵點位置並不一定相同，可能會有幾個像素距離的誤差，導致在周圍環境的灰階變化也會有些許的不同。(2)因取像失真造成灰階變化的差異。考量以上的關係，我們定義了兩個彩色碼的特性：

- **彩色碼相異數**：若彩色碼相同的個數越多，表示左右影像特徵點周圍八個方向的梯度方向越相近，也越有可能為對應點。
- **彩色碼距離**：雖然左右影像特徵點之彩色碼並不完全相同，但若彩色碼的顏色方向差距越小，表示梯度方向的差異較小，也就愈有可能為對應點。我們把這樣的差距稱為彩色碼距離。

考慮以上兩個因素，在決定彩色碼相異數的時候，我們可以容忍梯度方向在兩個相鄰顏色區域的邊緣抖動的情形，如此可避免一點點灰階變化造成彩色碼相異數增加，也就是說兩個彩色碼的顏色為相鄰，我們仍然認定這兩個彩色碼是相同的；當兩個彩色碼不相同且不為相鄰兩色時，則這

兩個彩色碼是相異的。另外由於灰色所代表的區域是灰階變化不明顯的區域，當兩個彩色碼均為灰色時，我們定義這兩個彩色碼相同；當一個彩色碼為灰色，另一個為其餘八種顏色之一時，則不考慮這兩個彩色碼的比較結果，也就是既不相異也不相同。以圖 4.5 為例，紅色和黃色是相鄰的區域，我們把彩色碼當成相同，而灰色和其餘八種顏色則不比較。



彩色碼距離是我們定義出來的另一個標準，當候選點的彩色碼相異數都相同時，彩色碼距離可用來做為判斷的依據。彩色碼距離最小為 0，表示兩個顏色相同；最大為 4，表示兩個顏色所代表的梯度方向經過量化後是完全反方向的顏色，以圖 4.9 的彩色碼分佈為例，當一個彩色碼為灰色，一個為其餘八種顏色之一，則不考慮彩色碼距離，可視為 0；兩個彩色碼皆為紅色時，則距離為 0；若一為白色，另一為綠色或是深藍色，則距離為 2；若一為白色一為紅色，則距離為 4，此時代表其梯度方向完全相反。利用彩色碼相異數和彩色碼距離，只要設合理的門檻值，我們可以在多組候選點裡找出對應的候選點。

4.3.2 雙向對應一致性

在特徵點對應的過程中，有時候在一張影像上出現的特徵點，並不會出現在另外一張影像中，造成這種情況的原因很多，一種是由於相機取像角度的不同，使得景深大的物體被景深較小的物體遮蔽；一種是因為光線變化改變了影像的灰階值…等，由於這些因素的影響，不管是利用極線限制、灰階比對、特徵比對或是彩色碼的方法，都無法找出正確的對應，觀察這個錯誤發生的原因，主要是因為特徵點對應的過程並不是一個對稱 (symmetric) 的步驟，在對應的時候只有可見點能夠對應。

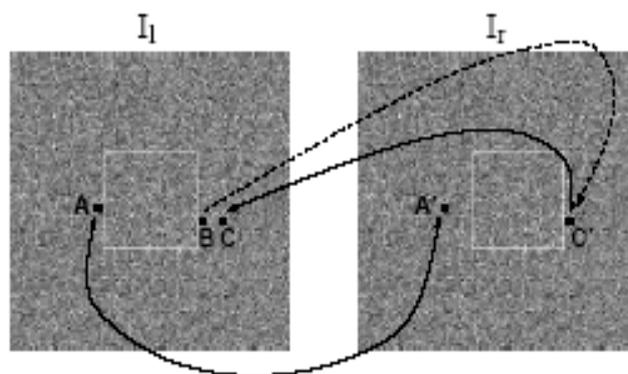


圖 4.10 雙向對應一致性

Fusiello[16]等人提出 left-right consistency 的概念，如果把左影像和右影像的角色互換，我們可以得到新的對應組合，而真正的特徵點對應組合必須滿足雙向的對應關係，也就是不論以左影像或右影像為基準，都必須能在另外一張影像中找到對應，如圖 4.10 所示，A 和 A'、C 和 C' 的對應都具有對稱關係，但是左影像的 B 點在右影像並沒有出現，經過對應的步驟後最有可能的對應點為 C'，但 C' 反推的對應點為 C，所以 B 和 C' 的對應關係不具有對稱性，我們稱這項特性為雙向對應一致性，加入雙向對應一致性，將能減少錯誤的對應。

4.3.3 利用彩色碼進行特徵點的對應

利用彩色碼相異數、彩色碼距離和雙向對應一致性，我們可以得到初步的對應關係，但是在對應的過程中，僅利用以上三種特性並無法解決模稜兩可(ambiguity)的問題。尤其是當影像中具有很多類似校正板影像的重複圖形時，每一個特徵點的梯度及梯度方向幾乎相同，更增加我們判斷的困難性。

從 Harris 演算法中對搜尋範圍的限制，我們可以聯想到特徵點和鄰近點之間也許存在一些可供比對的訊息，觀察一對彼此對應的特徵點和鄰近對應特徵點之間的關係，我們發現它們的距離、角度很相似，在特徵點對應的過程中，有些特徵點只須透過彩色碼的比對就可得到正確的對應結果，有些則需要額外的比對，如果先利用彩色碼找出一些特徵點對應，那麼其餘難以判定的對應點就可利用已知的對應結果作為相鄰點，再比對距離和角度，以挑選出正確的對應點，這樣的方法可以解決模稜兩可的問題。

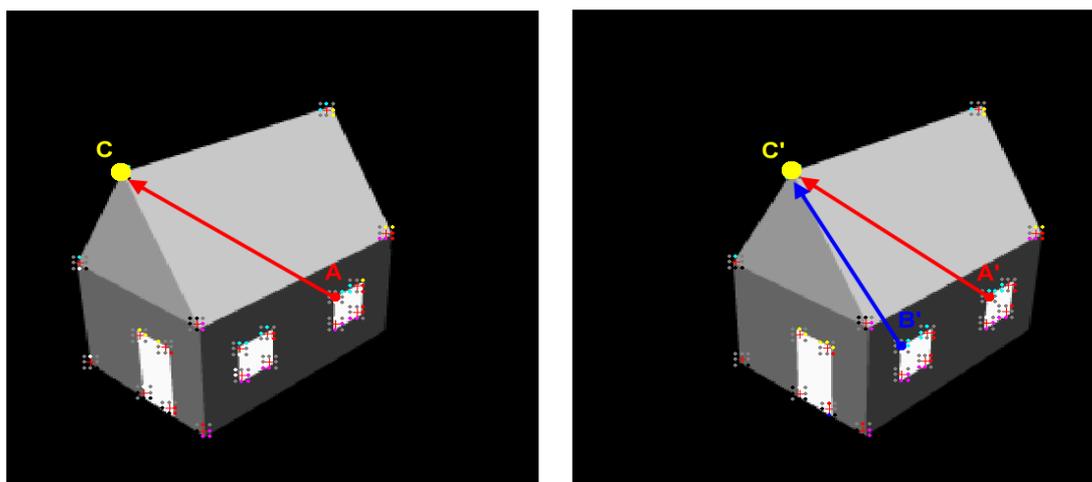


圖 4.11 利用已知對應點解決模稜兩可的對應問題

以圖 4.11 為例， C 和 C' 是利用彩色碼比對所找出一對特徵點，而左影像的 A 點在右影像中有兩個無法用彩色碼來判斷對應關係的候選點 A' 和

B' ，從距離的觀點來考量，我們可以分別計算向量 \overline{AC} 、 $\overline{A'C'}$ 、 $\overline{B'C'}$ 的長度 $|\overline{AC}|$ 、 $|\overline{A'C'}|$ 、 $|\overline{B'C'}|$ ，我們發現在長度上 $|\overline{AC}|$ 與 $|\overline{A'C'}|$ 比起 $|\overline{AC}|$ 與 $|\overline{B'C'}|$ 較為類似，從角度的觀點來考量，向量 \overline{AC} 與 $\overline{A'C'}$ 的夾角較 \overline{AC} 與 $\overline{B'C'}$ 小，如此可判斷出 A' 為 A 正確的對應點。

綜合彩色碼、距離、角度的概念，我們可以分別給一個評估分數，再利用評估分數來判斷候選點的好壞，針對一個想要找對應點的特徵點 A ，計算 A 至已知的對應點 P_i 的距離 $|\overline{P_iA}|$ ，並計算 A 的所有候選點 A_j' 至 P_i 的對應點 P_i' 的距離 $|\overline{P_i'A_j'}|$ 。

首先考慮角度的評估分數，若定義向量 $\overline{P_iA}$ 和 $\overline{P_i'A_j'}$ 的夾角為 θ_i ，當 θ_i 越接近 0° 時， $\cos \theta_i$ 越接近 1 ，我們令餘弦值總和為 C_j ，則評估分數 C_j 可由(4.12)式計算，其中 i 代表已正確對應的組數，若 C_j 越小代表所有 $\overline{P_iA}$ 和 $\overline{P_i'A_j'}$ 的夾角總和越小， A 和 A_j' 越有可能為正確的對應點，我們從這 N 組候選點計算出的 C_j 挑出最大值 C_{\max} 和最小值 C_{\min} 以做為接下來的加權使用。

$$\forall j \quad C_j = \sum_{i=1}^N (1 - \cos \theta_i) = \sum_{i=1}^N \left[1 - \left(\frac{\overline{P_iA} \cdot \overline{P_i'A_j'}}{|\overline{P_iA}| |\overline{P_i'A_j'}|} \right) \right] \quad (4.12)$$

$$C_{\max} = \text{Max} \{ C_j \}$$

$$C_{\min} = \text{Min} \{ C_j \}$$

其次考慮距離的評估分數，我們將所有 $\overline{P_iA}$ 和 $\overline{P_i'A_j'}$ 的差相加後可得到 A 和 A_j' 對於已知對應點的總距離差 D_j ，如(4.13)式，若 D_j 越小代表 A 和 A_j' 越有可能為正確的對應點，同樣的我們從這 N 組候選點計算出的 D_j 挑出最大

值 D_{\max} 和最小值 D_{\min} 以做為接下來的加權使用。

$$\forall j \ D_j = \sum_{i=1}^N \left(\left| \overline{P_i A} \right| - \left| \overline{P_i A_j} \right| \right) \quad (4.13)$$

$$D_{\max} = \text{Max} \{ D_j \}$$

$$D_{\min} = \text{Min} \{ D_j \}$$

除了角度和距離之外，彩色碼的特性也值得考慮，在我們的程式中要成為一組對應點最多只能有 3 個彩色碼不同、且這幾個彩色碼距離和必須小於等於 10，這兩個門檻值的大小影響了特徵點的判斷，為解決模稜兩可的問題，可用(4.14)式計算A與所有候選點 A_j 的彩色碼差異值。我們從所有 CG_j 挑出最大值 CG_{\max} 和最小值 CG_{\min} 以做為接下來的加權使用， CG_j 最小值為 0，此時兩點彩色碼完全一致，由於彩色碼相異數為 3，彩色碼距離最大為 4，經過計算後 CG_j 最大值為 15。

$$CG_j = (\text{彩色碼相異數}) + (\text{彩色碼距離}) \quad (4.14)$$

$$CG_{\max} = \text{Max} \{ CG_j \}$$

$$CG_{\min} = \text{Min} \{ CG_j \}$$

對角度、距離和彩色碼的分數取加權後，每一個候選點 A_j 我們都可得到最後的評估分數 G_j ，如(4.15)式，整體而言，當 G_j 值越小，表示 A_j 和其他候選點在各方面比起來較相似，因此可當作正確的對應點。

$$\forall j \ G_j = \frac{D_j - D_{\min}}{D_{\max} + D_{\min}} + \frac{C_j - C_{\min}}{C_{\max} + C_{\min}} + \frac{CG_j - CG_{\min}}{CG_{\max} + CG_{\min}} \quad (4.15)$$

第五章 計算基本矩陣

在得到多組特徵點對應之後，本章將介紹目前各種計算基本矩陣的相關方法，不同的演算法考慮了對於基本矩陣不同的限制，目的皆為找出最正確的基本矩陣。

在第三章曾提到，要計算基本矩陣，至少需要七組以上的對應點，令

一組對應點 $m_i = (u, v, 1)^T$ 、 $m_i' = (u', v', 1)^T$ ， F 為
$$\begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix}$$
，因 $m_i'^T F m_i$

$= 0$ ，每組對應點的關係式可寫成 $U_i f = 0$ ，其中

$$u_i = [uu' \quad uv' \quad u \quad vu' \quad vv' \quad v \quad u' \quad v' \quad 1]^T \quad (5.1)$$

$$f = [f_{11} \quad f_{12} \quad f_{13} \quad f_{21} \quad f_{22} \quad f_{23} \quad f_{31} \quad f_{32} \quad f_{33}] \quad (5.2)$$

如有 n 組對應點，待解的線性方程式可寫成 $U_n f = 0$ ， $U_n = [u_1, \dots, u_n]^T$ ，

若利用七組對應點計算基本矩陣，因為 U_n 的秩為七、維度是二，透過 SVD 分解，可找到兩個基底 f_1 、 f_2 ，分別對應兩個矩陣 F_1 、 F_2 ，基本矩陣 F 可定義為 F_1 、 F_2 的線性組合，寫成 $\alpha F_1 + (1-\alpha)F_2$ ，又 F 的行列式值必須為零，可得到一個 α 的三次方程式，所以 α 可解出一或三個實根，代入 α 可求得一或三個 F 。下一小節將討論適用於八點或多於八點的演算法。

5.1 八點演算法

八點演算法是由 Longuet-Higgins[8]提出，在實際的應用裡，經過特徵

點擷取後，常常會超過七組對應點，如果忽略基本矩陣秩為二的限制，我們可使用最小平方方法解目標函數(5.3)式，將(5.1)式和(5.2)式代入(5.3)式可得(5.4)式，分列如下：

$$\text{MIN}_F \sum_i (m_i^T F m_i)^2 \quad (5.3)$$

$$\text{MIN}_F \|U_n f\|^2 \quad (5.4)$$

(5.4)式在 $f=0$ 時有最小值，但基本矩陣為零並不是我們要的解，解決的方法是對基本矩陣的係數加上一些限制，下面將介紹幾種不同的限制方法，這些方法理論上僅需八組對應點就可解出基本矩陣，但為了求精確的基本矩陣，通常取超過八組對應點，因此超過八組對應點的情形也適用。

5.1.1 線性最小平方方法



第一個方法是令基本矩陣的某個係數為 1，然後用線性最小平方方法解(5.3)式，在不失一般性的情況下，可以假設向量 f 的最後一個元素 f_9 不為零，若設 $f_9 = -1$ ，則 $\|U_n f\|^2$ 可寫成

$$\|U_n f\|^2 = \|U_n' f' - c_9\|^2 = f'^T U_n'^T U_n' f' - 2c_9^T U_n' f' + c_9^T c_9 \quad (5.5)$$

其中 U_n' 是由 U_n 前八行組成的 $n \times 8$ 矩陣， c_9 為 U_n 之第九行。令 $\|U_n f\|^2$ 的一階導函數為零，也就是(5.5)式左右同時對 f' 做偏微分，可推導出最後的解 $2U_n'^T U_n' f' - 2c_9^T U_n' = 0$ 或 $f' = (U_n'^T U_n')^{-1} U_n'^T c_9$ ，但這方法的問題是我們無法預知哪一項係數為零，若 f_9 實際上遠小於其他元素或是零，結果將會錯的離譜，解決的辦法是將基本矩陣的九個係數分別設為 1，再用這一小節的方法分別計算，並保留最佳解。

5.1.2 特徵分析(Eigen analysis)

第二個方法是限制 f 的長度 $\|f\|=1$ ，和 5.1.1 小節相比， f 的九個元素中沒有任何一個元素具有預設長度，如此(5.4)會變成(5.6)式，透過一個 Lagrange multiplier λ ，可將(5.6)式轉換成無限制最小化的問題，如(5.7)式。

$$\underset{F}{\text{MIN}} \|U_n f\|^2 \quad (\text{且 } \|f\|=1) \quad (5.6)$$

$$\underset{F}{\text{MIN}} F(f, \lambda)$$

$$\text{其中 } F(f, \lambda) = \|U_n f\|^2 + \lambda(1 - \|f\|^2) \quad (5.7)$$

令 $F(f, \lambda)$ 對 f 之一階導函數為零，可得 $U_n^T U_n f = \lambda f$ ，所推出的解 f 必為 9×9 矩陣 $U_n^T U_n$ 的特徵向量，且 λ 為對應的特徵值，又因矩陣 $U_n^T U_n$ 具有對稱的特性，所以九個特徵值皆為實數或零，假設九個特徵值由小到大分別為 $\lambda_1, \dots, \lambda_9$ 且 $\lambda_1 \sim \lambda_9 \geq 0$ ，將這個結果代回(5.7)式可得 $F(f, \lambda) = \lambda_i$ for $i=1 \sim 9$ ，因為我們的目的是將 $F(f, \lambda)$ 最小化，所以要取最小特徵值 λ_1 。

5.1.3 限制基本矩陣的秩為二

5.1.1 節的線性最小平方法的好處是可以產生分析的解，但即使在資料量很大的情況下，對雜訊仍然很敏感，其中一個原因是因為沒有滿足基本矩陣秩為二的限制。

基本矩陣透過 SVD 分解(singular value decomposition)可得 $F = USV^T$ ，其中 $S = \text{diag}(\sigma_1, \sigma_2, \sigma_3)$ 為一對角矩陣，其中 $\sigma_1 \geq \sigma_2 \geq \sigma_3$ 、 σ_i 為第 i 個奇異值， U 、 V 為正交矩陣，為了滿足秩為二的限制，可以對基本矩陣做下修改

$$\hat{F} = U\hat{S}V^T \quad \hat{S} = \text{diag}(\sigma_1, \sigma_2, 0) \quad (5.8)$$

在基本矩陣計算過程中加入秩為二的限制是由 Faugeras 提出，我們可以不失一般性的假設 $f = [g^T, f_8, f_9]^T$ ，其中 g 為包含基本矩陣前七個元素的向量，若 c_8 、 c_9 是矩陣 U_n 的最後兩個行向量， C 是包含 U_n 前七行的 $n \times 7$ 矩陣，又 $U_n f = 0$ ，我們可推得 $Cg = -f_8 c_8 - f_9 c_9$ ，假設 C 的秩為七，我們可用最小平方法解 g ，如下

$$g = -f_8 (C^T C)^{-1} C^T c_8 - f_9 (C^T C)^{-1} C^T c_9 \quad (5.9)$$

(5.9)式的解根據兩個自由變數 f_8 、 f_9 的不同而有所改變，如果令基本矩陣行列式值為零，我們可得到 f_8 和 f_9 的三次齊次方程式，可求出 f_8 和 f_9 的比例關係，因為三次方程式存在一個或三個實根解，所以至少可得到一組基本矩陣，分為下列兩種情形：

- 存在一個實根解：因為這個解求出的基本矩陣長度大小並不固定，僅差異一個比例常數，我們可以將基本矩陣正規化使得向量長度為一。
- 存在三個實根解：分別將三個實根代入方程式，可求得三個基本矩陣，接著將三個基本矩陣正規化，選擇使 $U_n f$ 向量長度最小的基本矩陣和其對應的實根。

這個方法和 5.1.1 節提到的線性最小平方法稍有不同，相異之處在於 Faugeras 提出的方法是在線性最小平方法之後再加上秩為二的限制，計算出的基本矩陣會比僅用線性最小平方法稍好。

5.1.4 線性關係的幾何表示

線性關係的另一個問題在於我們想最小化的關係式 $\|U_n f\|^2$ 本身不具任何實際意義，因為不管是左右影像的特徵點，其座標都以二維的方式表示，所以應該在影像平面上測量才具有物理意義，其中一種具實際物理意義的度量方式是特徵點到其對應極線的距離，我們稱為極線距離(epipolar distance)，以右影像為例，極線距離可寫成

$$d(m_i', l_i') = \frac{m_i'^T l_i'}{\sqrt{l_1' + l_2'}} = \frac{1}{b_i} \sqrt{m_i'^T F m_i'} \quad (5.10)$$

其中 $b_i = \sqrt{l_1' + l_2'}$ ， l_1' 、 l_2' 分別代表右影像極線方程式 l_i' 中 x 、 y 的係數，如此(5.3)式可推得 $\underset{F}{\text{MIN}} \sum_i b_i^2 d^2(m_i', l_i')$ ，由此可知在求解的過程中，我們同時對無具體意義的量 b_i 和有具體意義的量 $d(m_i', l_i')$ 求最小化。

5.2 極線距離之非線性最小化

由 5.1.4 節的討論可知，該方法針對具有實際物理意義的極線距離做最小化，但(5.10)式僅針對右影像的部分討論，考慮對稱性，由於左右影像的角色不盡相同，也就是說(5.10)式僅考慮右影像上的極線，並不適用於左影像，所以有必要修正(5.10)式，若將左右影像的角色互換，左影像極線距離之最小化可寫成 $\underset{F}{\text{MIN}} \sum_i d^2(m_i, l_i)$ ，如此(5.10)式可用(5.11)式代替：

$$\text{MIN}_F \sum_i d^2(m_i', l_i') + d^2(m_i, l_i) \quad (5.11)$$

令左影像上的極線 $l_i = F^T m_i' = [l_1, l_2, l_3]^T$ ，右影像上的極線 $l_i' = F m_i = [l_1', l_2', l_3']^T$ ，又我們已知 $m_i'^T F m_i = m_i^T F m_i'$ ，根據(5.10)式，我們可把(5.11)式改寫成

$$\text{MIN}_F \sum_i w_i^2 (m_i'^T F m_i)^2$$

$$\text{其中 } w_i = \left[\left(\frac{1}{l_1^2 + l_2^2} + \frac{1}{l_1'^2 + l_2'^2} \right) \right]^{1/2} = \left(\frac{l_1^2 + l_2^2 + l_1'^2 + l_2'^2}{(l_1^2 + l_2^2)(l_1'^2 + l_2'^2)} \right)^{1/2} \quad (5.12)$$

接下來我們將介紹如何利用(5.12)式計算基本矩陣。

5.2.1 反覆式線性法



前一節裡所討論的極線距離最小化，即(5.12)式，亦可用權重式線性最小平方法解之，在計算基本矩陣的過程中，如果分配給每組對應點一個權重 w_i ，則對應的線性方程式必須乘上權重，而八點演算法的算式可修改成 $U_n f = 0$ ，其中 $U_n = [w_1 u_1, \dots, w_8 u_8]^T$ ，若每一組對應點使用相同的權重，相當於使用(5.12)式來估計基本矩陣。

上述方法碰到的問題是每組對應點的權重 w_i 在不同的基本矩陣並不完全一樣，為了克服這個問題，我們使用反覆式線性法。首先假設每一組對應點的權重均為 1，利用改良後的八點演算法計算基本矩陣的初始值，再從基本矩陣初始值估算權重 w_i ，重複很多次以後可以得到每一組對應點的新權重 w_i 。

雖然這個方法採用極線距離，也很容易撰寫程式，但和原來 5.1.1 節的線性最小平方法相比，結果並無顯著的改進，主要的原因是沒有考慮基本矩陣秩為二的限制。

5.2.2 參數空間之非線性最小化

Loung 在[17]中提到，基本矩陣有幾種參數化的方法，從本章的討論中，我們知道一個事實，如果利用基本矩陣秩為二的特性，也就是基本矩陣中的某一行(或某一列)是其他兩行(或其他兩列)的線性組合，將能夠有效獲得正確的基本矩陣。以下我們將討論參數化的過程。

若基本矩陣 F 的行向量分別以 c_1 、 c_2 、 c_3 表示，秩為二之限制相當於 (5.13) 式，且秩為二之 3×3 基本矩陣可為 (5.14) 式的形式之一。

$$c_{j_0} + \lambda_1 c_{j_1} + \lambda_2 c_{j_2} = 0 \quad (\exists j_0, j_1, j_2 \in [1, 2, 3]) \quad (\exists \lambda_1, \lambda_2 \in R) \quad (5.13)$$

$$\begin{bmatrix} c_1 & c_2 & \lambda c_2 \end{bmatrix} \quad \begin{bmatrix} c_1 & 0_3 & c_3 \end{bmatrix} \quad \begin{bmatrix} c_1 & c_2 & 0_3 \end{bmatrix} \quad (5.14)$$

假設我們建立一個三維向量，使得 λ_1 、 λ_2 分別為第 j_1 和 j_2 個座標，且 1 為第 j_0 個座標，此向量即為基本矩陣 F 的特徵向量，在這裡也就是基本矩陣的極點(epipole)，用參數化的方法可以直接計算極點座標，而不是經由基本矩陣反推極點位置。

由 3×3 基本矩陣的參數空間之非線性最小化可知，除了行向量外，也可以從列向量的角度考慮，列向量的分解方式和行向量一樣，其中一列(或一行)可表示為其餘兩列(或兩行)的線性組合，所以我們可將參數集合分成九個映對。

矩陣的參數化可以利用左影像的極點座標 x 和 y 及右影像的極點座

x' 和 y' ，與對應於極線轉換矩陣的四個元素 a 、 b 、 c 、 d ，極線轉換矩陣可將左影像的極線映射至右影像的極線，以指標在第三列第三行為例，基本矩陣的參數化可寫成(5.15)式

$$F = \begin{bmatrix} a & b & -ax-by \\ c & d & -cx-dy \\ -ax'-cy' & -bx'-dy' & (ax+by)x'+(cx+dy)y' \end{bmatrix} \quad (5.15)$$

考慮基本矩陣的絕對長度並不固定，且允許一個比例常數的差異，所以把 a 、 b 、 c 、 d 除以四個元素當中最大元素的絕對值將矩陣正規化，這樣一來我們可以得到 36 種基本矩陣參數表示法。

■ 選擇最佳之參數表示法

給定基本矩陣和左右影像的極點座標，我們必須從 36 種參數表示法找出基本矩陣最佳的表示方式，一旦選擇某一種參數表示法，同時也決定了行列指標的位置，以(5.15)式為例，指標在第 i_0 列($i_0=3$)和第 j_0 行($j_0=3$)，為了和本章一開始的符號區隔，我們以 f_{i_0, j_0} 代表這種參數化表示法的基本矩陣，我們的目的是透過 i_0 及 j_0 的選擇使 9×8 的 Jacobian matrix 的秩最大，其中 Jacobian matrix 可用(5.16)式表示。

$$J = \frac{d(f_{i_0, j_0})}{dp} = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ -a & -b & 0 & 0 & -x & -y & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ -c & -d & 0 & 0 & 0 & 0 & -x & -y \\ 0 & 0 & -a & -c & -x' & 0 & -y' & 0 \\ 0 & 0 & -b & -d & 0 & -x' & 0 & -y' \\ ax'+cy' & bx'+dy' & ax+by & cx+dy & xx' & yx' & xy' & yy' \end{bmatrix} \quad (5.16)$$

其中 $p = [x, y, x', y', a, b, c, d]^T$ ，要使 J 的秩最大可用另一個角度思考，首先建立一個 9 維向量，9 個元素分別是 J 的九個 8×8 子矩陣的行列式值，求此

向量長度使之最大，在此同時也可使 J 的秩最大，經過簡單計算可得(5.17)式，換句話說，三十六種基本矩陣參數表示法當中，最好的基本矩陣能夠使(5.17)式的值最大。

$$(ad - bc)^2 \sqrt{x^2 + y^2 + 1} \sqrt{x'^2 + y'^2 + 1} \quad (5.17)$$

僅用一種參數表示法存在一個缺點，當極點在無窮遠處時，基本矩陣的計算可能產生錯誤，但不管極點在哪裡，利用三十六種基本矩陣參數表示法都可以求解，代價是需要付出較多的計算時間。

■ 非線性最佳化

對(5.16)式求最佳化還有很多其他的方法，在此不一一介紹，另外值得一提的是：在最佳化的過程中，基本矩陣參數表示法可能會改變，最佳化程序開始的參數表示法可能不適合最後的基本矩陣，通常會以八點演算法計算出的基本矩陣當成初始值。



5.3 梯度分析

對(5.3)式 $\sum_i (m_i^T F m_i)^2$ 求最小化並不代表一定能夠得到好的基本矩陣，因為每一組對應點算出的 $(m_i^T F m_i)^2$ 其變異量不盡相同，令 $f_i = m_i^T F m_i$ ，且每一組 f_i 的變異量為 $\sigma_{f_i}^2$ ，當每一組 f_i 的變異量相同時，最小平方法會得到一個最佳解，考慮變異量的不同，目標函數(5.3)式可修改成

$$\text{MIN}_F \sum_i f_i / \sigma_{f_i}^2 \quad (5.18)$$

因為我們是以相同方法擷取左右影像上的特徵點，假設特徵點的雜訊呈現獨立且相似的高斯分佈，其共變異量矩陣(covariance matrix)可用

$\Lambda_{m_i} = \Lambda_{m'_i} = \sigma^2 \text{diag}(1,1)$ 表示，其中 σ 為未知的雜訊程度，透過一階導函數逼

近， f_i 的變異量 $\sigma_{f_i}^2$ 可用下式表示

$$\sigma_{f_i}^2 = \frac{\partial f_i}{\partial m_i} \Lambda_{m_i} \frac{\partial f_i}{\partial m_i} + \frac{\partial f_i}{\partial m'_i} \Lambda_{m'_i} \frac{\partial f_i}{\partial m'_i} = \sigma^2 [l_1^2 + l_2^2 + l'_1{}^2 + l'_2{}^2] \quad (5.19)$$

其中 $l_i = F^T m_i = [l_1, l_2, l_3]^T$ ， $l'_i = F m'_i = [l'_1, l'_2, l'_3]^T$ ，若將(5.18)式的分子分母同

乘 σ^2 ，並不會影響最小化的結果，(5.18)式可寫成(5.20)式，其中 $g_i^2 =$

$\sqrt{l_1^2 + l_2^2 + l'_1{}^2 + l'_2{}^2}$ 代表 f_i 的梯度(gradient)。

$$\min_F \sum_i (m_i^T F m_i)^2 / g_i^2 \quad (5.20)$$

5.4 正規化



Hartley[18]從計算的觀點分析八點演算法的問題，在所有解基本矩陣的線性法裡，以八點演算法最常被利用，如果直接使用影像中的特徵點座標當輸入，再用線性方法計算基本矩陣，計算的結果將不穩定。基於上述原因，Hartley 提出等向性縮放法(istropic scaling method)，若得到影像中的特徵點座標後，先對點座標做平移和縮放，再代入八點演算法求解，將會得到較好的結果。影像的正規化可分為下列三個步驟：

- 若 m_i 、 m'_i 分別為一組互相對應的左右影像特徵點座標，透過 $\hat{m}_i = T m_i$ 、 $\hat{m}'_i = T' m'_i$ 可以轉換影像座標，其中 T 、 T' 為包含平移和縮放的正規化轉換矩陣。
- 若 \hat{m}_i 、 \hat{m}'_i 為一組對應的特徵點，利用線性最小平方法解基本矩陣 \hat{F} ，

並加上秩為二之限制。

- F 為 m_i 、 m_i' 之間對應的基本矩陣，則影像座標正規化前後的基本矩陣

F 、 \hat{F} 存在 $F = T^T \hat{F} T$ 的關係。

將每張影像的特徵點平移和縮放，使這些特徵點的中心落在座標系原點，這樣的過程稱為影像的正規化，假設影像特徵點的齊次座標表示法為 $[x, y, z]^T$ ，在 x 和 y 方向選擇適當的等向縮放比例，經過正規化轉換後，每個特徵點齊次座標的三個分量將具有相同的平均值，且各點到原點的平均距離為 $\sqrt{2}$ 。以上我們介紹了目前各種基本矩陣的計算方法，若已有超過七組以上的特徵點對應組合，即可求出基本矩陣。



第六章 利用極線幾何改善特徵點之對應

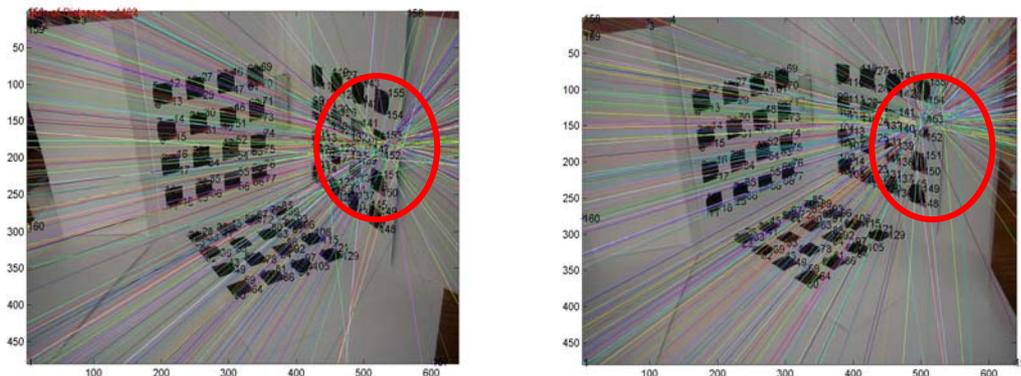


圖 6.1 右影像和左影像的極線

利用第四章的特徵點對應演算法找出特徵點對應之後，便可由八點演算法求出基本矩陣，有了基本矩陣，每一個特徵點所對應的極線方程式便可由(3.2)式、(3.4)式得到。理論上，若是特徵點對應的結果越精確，所得到的基本矩陣將會越正確，特徵點會越接近所對應的極線，而所有的極線也會交於極點。

以圖 6.1 為例，由於影像中存在錯誤的對應組合，所以左右影像中的極線並不全然交於極點。在接下來的小節裡，我們將利用兩個極線幾何的限制來改善特徵點的對應，以求出正確的基本矩陣，並以兩種計算方式反推極點位置，證明對應的正確性。

6.1 極線距離之限制

從 3.1 節中，我們了解極線限制的關係：若給定一個左影像平面的點 m ，在右影像平面必對應一條極線 l' ，且 m 的對應點 m' 必落在 l' 上，也就

是說若此組對應點是正確的對應，則左右影像的極線距離應該趨近於零，值得注意的是，極線距離為零並不代表正確的對應點，不過我們仍可以設合理的門檻值來去除錯誤的對應點。

考量左影像和右影像的差異，在左右影像其中一方或兩方的極線距離都很大或都很小時，若對極線距離設相同的門檻值，可能會失之偏頗，若對極線距離設不同的門檻值，又不禁使我們懷疑此門檻值的實用性，且門檻值可能只適合用於某一場景，故不適合分開考慮。在這裡我們套用 2.3 節 LMeDs 的方法，但稍微做一點修改。首先計算左右影像上每一組特徵點的極線距離，分別對左右影像每一組的極線距離取平方後相加，並找出左右影像 n 組極線距離平方和的中位數，令為 M_J ，將 M_J 代入(6.1)式可到 σ ，其中 n 為對應的特徵點組數。

$$\sigma = 1.4826 \times [1 + 5/(n-8)] \times \sqrt{M_J} \quad (6.1)$$

根據 σ ，我們可以看出衡量對應點正確與否的標準，如(6.2)式，若 $\omega_i=1$ ，代表此組對應可當作正確的對應；若 $\omega_i=0$ ，則此組對應可視為錯誤的對應，經由極線距離的限制後，我們可以刪除錯誤的對應組合。

$$\omega_i = \begin{cases} 1 & \text{if } r_i^2 \leq (2.5\sigma)^2 \\ 0 & \text{otherwise} \end{cases} \quad (6.2)$$

$$r_i^2 = d^2(m_i', Fm_i) + d^2(m_i, F^T m_i') \quad (6.3)$$

6.2 極點到極線的距離之限制

一對正確對應的特徵點除了滿足極線距離的限制外，其對應的極線也必須交於極點之上，因此我們定義了極點到極線的距離，為了計算這個距

離，我們必須先計算極點的位置，在接下來的小節裡，我們將介紹兩種極點的計算方式。

6.2.1 利用 pseudo-inverse 計算極點

在得到左右影像特徵點座標和基本矩陣之後，我們可經由 $l = F^T m'$ 、 $l' = Fm$ 計算左右影像的極線方程式，其中 $l = [l_1, l_2, l_3]^T$ 、 $l' = [l'_1, l'_2, l'_3]^T$ 、左影像極點 e 的座標 (x, y) 、右影像極點 e' 的座標為 (x', y') ，也就是說左影像和右影像的極線方程式可分別表示為 $l_1x + l_2y + l_3 = 0$ 、 $l'_1x' + l'_2y' + l'_3 = 0$ ，若影像中有 n 組對應點，則左右影像將各有 n 個極線方程式。下面將以左影像為例，說明利用 pseudo-inverse 計算極點 e 的過程，同理可推得右影像的極點 e' 。



$$\text{若左影像的 } n \text{ 個極線方程式為 } \begin{cases} a_1x + b_1y = c_1 \\ a_2x + b_2y = c_2 \\ \bullet \quad \bullet \quad \bullet \\ \bullet \quad \bullet \quad \bullet \\ a_nx + b_ny = c_n \end{cases}, \text{ 則極點可視為所有極線}$$

的交點，以矩陣型態表示極線方程式，可得(6.4)式，值得我們注意的一點是， $a_1x + b_1y = c_1$ 和程式所計算出的極線 $l_1x + l_2y + l_3 = 0$ 之常數項相差一個負號，如果稍有不慎，將會造成極點的計算錯誤。

$$\begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \\ \bullet & \bullet \\ \bullet & \bullet \\ a_n & b_n \end{bmatrix}_{n \times 2} \begin{bmatrix} x \\ y \end{bmatrix}_{2 \times 1} = \begin{bmatrix} c_1 \\ c_2 \\ \bullet \\ \bullet \\ c_n \end{bmatrix}_{n \times 1} \quad (6.4)$$

由於 $n \times 2$ 矩陣不一定為方陣，所以反矩陣不一定存在，為了計算的便

利性，我們令 $C = \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \\ \bullet & \bullet \\ \bullet & \bullet \\ a_n & b_n \end{bmatrix}_{n \times 2}$ 、 $H = \begin{bmatrix} c_1 \\ c_2 \\ \bullet \\ \bullet \\ c_n \end{bmatrix}_{n \times 1}$ ，根據 pseudo-inverse 的定義，可

用下式求 x 和 y 的解

$$\begin{bmatrix} x \\ y \end{bmatrix}_{2 \times 1} = (C^T C)^{-1} C^T H \quad (6.5)$$

求得極點座標後，我們便可以計算極點到每一條極線的距離，並利用這個距離當作篩選的標準。

6.2.2 利用基本矩陣計算極點

以右影像為例，任意選取右影像上除極點 e' 之外的一點 m ，極線方程式可表示為 $l = Fm$ ，又極線 l 應會通過極點 e' ，所以極點 e' 必滿足 $e'^T (Fm) = (e'^T F)m = 0$ ，如此可得到 $(e'^T F) = 0$ ，亦即 $(F^T e') = 0$ ，也就是說 e' 是基本矩陣的 left null-space，同理可證， e 是基本矩陣的 right null-space。

若 $F = \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix}$ 、左影像極點 e 的座標 (x, y) 、右影像極點 e' 的座

標為 (x', y') ，觀察程式中所求出之基本矩陣，我們發現第三列(第三行)的元素比其他兩列(其他兩行)為大，利用基本矩陣秩為二之限制，可將第三列(第三行)的元素視為其他兩列(其他兩行)的線性組合， $(Fe) = 0$ 和 $(F^T e') = 0$ 的關係可分別用(6.6)式和(6.7)式表示。

$$\begin{cases} f_{11}x + f_{12}y = f_{13} \\ f_{21}x + f_{22}y = f_{23} \end{cases} \quad (6.6)$$

$$\begin{cases} f_{11}x' + f_{21}y' = f_{31} \\ f_{12}x' + f_{22}y' = f_{32} \end{cases} \quad (6.7)$$

利用行列式求解，可得兩個極點 e 、 e' 的座標，如下，

$$\begin{aligned} x &= \frac{|f_{13} \times f_{22} - f_{12} \times f_{23}|}{|f_{11} \times f_{22} - f_{12} \times f_{21}|} & y &= \frac{|f_{13} \times f_{21} - f_{11} \times f_{23}|}{|f_{11} \times f_{22} - f_{12} \times f_{21}|} \\ x' &= \frac{|f_{31} \times f_{22} - f_{21} \times f_{32}|}{|f_{11} \times f_{22} - f_{12} \times f_{21}|} & y' &= \frac{|f_{31} \times f_{12} - f_{11} \times f_{32}|}{|f_{11} \times f_{22} - f_{12} \times f_{21}|} \end{aligned} \quad (6.8)$$

有極點座標後，我們可以計算極點到每一條極線的距離，並利用這個距離當作篩選的標準。

舉例來說，若 $F = \begin{bmatrix} 2.9121e-006 & 3.5235e-005 & -0.0068865 \\ -3.0898e-005 & 5.6597e-006 & 0.014985 \\ 0.0038465 & -0.019687 & 0.99966 \end{bmatrix}$ ，將基本矩

陣 F 代入(6.8)式，可求得左影像的極點座標(513,153.04)、右影像的極點座標(530.69,174.51)。

6.3 不同門檻值的影響

在 6.2 節中，我們介紹了兩種極線幾何的限制，分別是極線距離的限制和極點到極線的距離的限制，對這兩種限制設不同的門檻值，可以刪除一些錯誤的對應組合，進而改善特徵點的對應。

不同的門檻值對於改善特徵點對應的演算法有很大的影響。不管是極線距離之限制或是極點到極線的距離之限制，距離越小代表越有可能為正確的特徵點對應，距離越大代表越有可能是錯誤的特徵點對應，找出這樣

的關係後，根據實際需要的不同，我們可以對這兩種距離限制設定不同的門檻值，並保留距離限制小於門檻值的對應組合。當門檻值越大，表示我們的條件限制越寬鬆，符合的特徵點對應越多；門檻值越小，表示我們的條件限制越嚴格，符合的特徵點對應越少。

我們以圖 6.1 為例，說明門檻值的大小對於特徵點對應組數的影響。經過初步的特徵點對應後，左右影像的 160 組極線分佈如圖 6.1，觀察圖 6.1，所有的極線並不完全交於極點上，所以特徵點的對應仍有改進空間，為了改善特徵點的對應，我們分別對兩種距離限制代入不同大小的門檻值，可得到剩餘的特徵點對應組數和門檻值大小的趨勢關係，為了方便說明，先以極點至極線的距離之限制為例，將結果加以整理分析可得表 6.1，至於詳細的實驗結果，將等到第七章再做完整整理。

表 6.1 門檻值的大小對於特徵點對應組數的影響

原來的特徵點對應組數	極點到極線的距離門檻值(左影像/右影像)	限制的程 度	更新後的特徵點對應組數
160	10/10	寬鬆	130
160	5/5	中等	107
160	1/1	嚴格	51

6.4 特徵點對應之改善

在第四章我們介紹了特徵點對應的步驟，主要是以彩色碼比對為基礎並結合與已知對應點的相對關係來解決模糊對應的問題。在距離、角度以及雙向對應一致性的限制之下，建立一對立體影像上的特徵點對應，分析

這些特徵點的對應組合並觀察這兩張立體影像的極線分佈，若還有改善的空間，就可使用本篇論文的演算法。

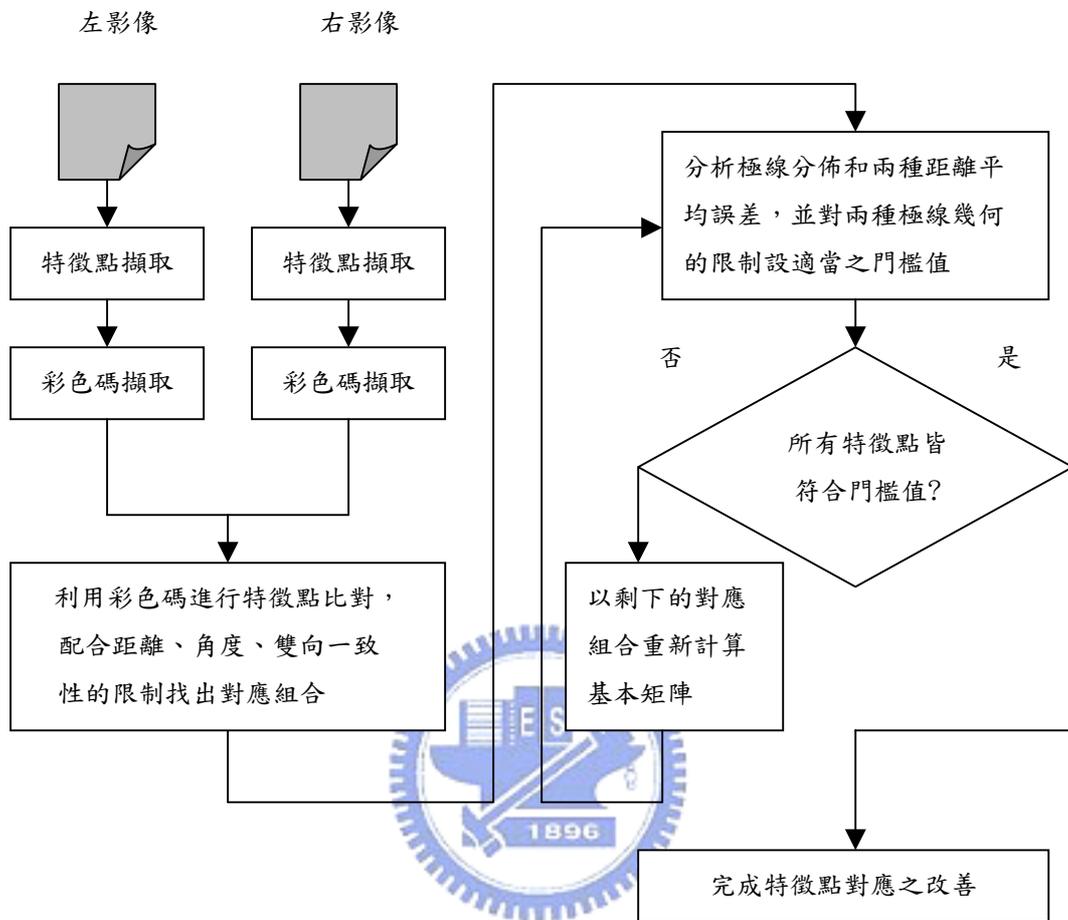


圖 6.2 特徵點對應之改善

圖 6.2 為改善特徵點對應的流程，一開始輸入兩張立體影像，並擷取其中的角點作為特徵點以及特徵點所對應的彩色碼，取得彩色碼之後我們就可以依據彩色碼比對機制來尋找對應點，接著利用極線幾何的兩種限制刪除誤差較大的對應點，蒐集剩下較好的對應點並重新計算基本矩陣，重複刪除和計算基本矩陣的步驟，直到不能再去除對應的特徵點為止。在下一章中，我們將用不同的門檻值進行實驗，實驗結果顯示不管是極線分佈圖或是極線距離平均誤差、極點至極線距離的平均誤差都比原來的結果好。

第七章 實驗結果

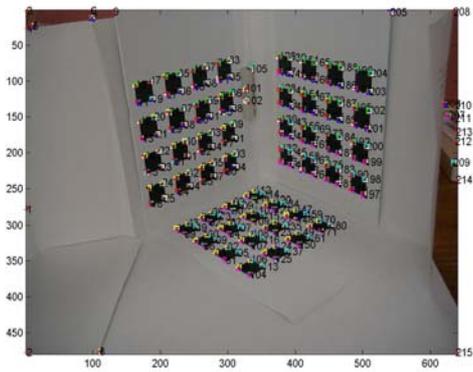
本篇論文提出一個以彩色碼比對為基礎，並利用極線限制改善特徵點對應的方法。在這一章裡，我們將用實際從數位像機取得的立體影像進行實驗，以證明此演算法能夠找出正確的特徵點對應組合。實驗的硬體環境是 Pentium IV 2.0GHz，作業系統為 Windows XP，程式發展環境為 Matlab 6.5。本章的實驗主要可分為以下三個部分，下面的小節裡我們將深入探討這三個主題。

- 利用兩種極線限制(極線距離和極點到極線的距離)刪除錯誤的對應
- 重算基本矩陣和不重算基本矩陣的影響
- 極點位置的不確定性

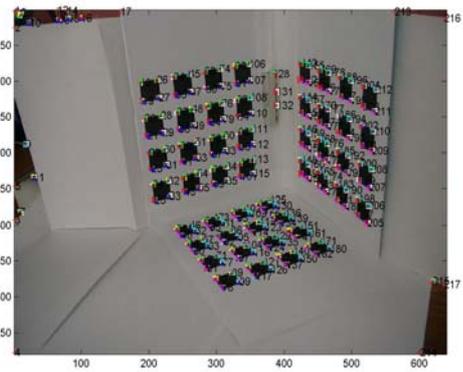


7.1 利用極線限制刪除錯誤的對應

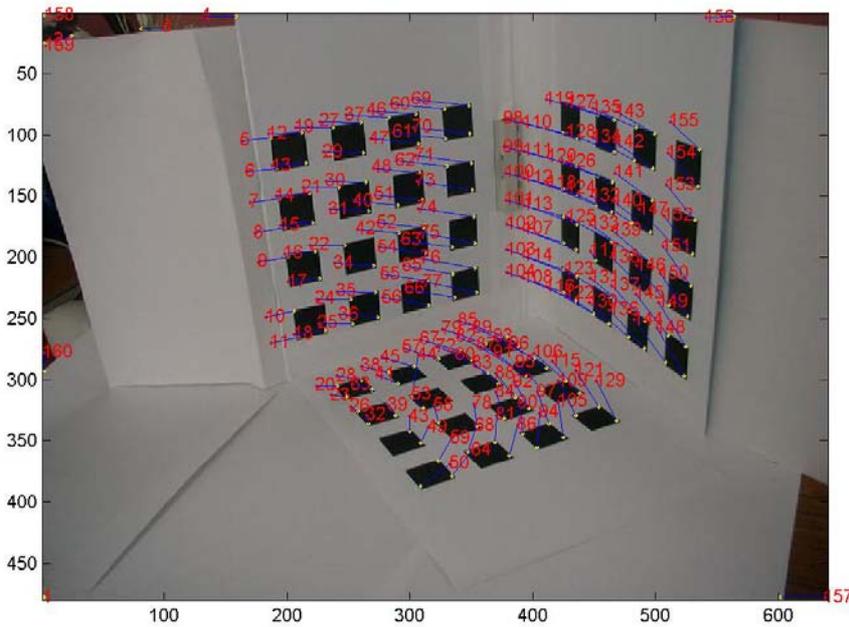
第一個要討論的主題是有關於如何利用極線限制刪除錯誤的對應組合。下面的例子是以 AGFA 1680 數位相機拍攝兩張不同角度的校正板影像，每一張影像大小為 640×480 ，這兩張影像的的差異除了平移之外，並向右旋轉了約 30° ，調整 Harris 演算法的各項參數分別對兩張影像擷取特徵點後，左影像可得 215 個特徵點，而右影像可得 217 個特徵點，利用彩色碼進行特徵點對應後，我們一共得到 160 組的對應點，如圖 7.1(c)，其中紅色點代表左影像特徵點位置，黃色點代表右影像特徵點位置，藍色線段表示左右影像之間的對應關係。左影像擷取角點的時間為 22.7 秒、右影像擷取角點的時間為 23.39 秒、對應時間一共是 42.16 秒。



(a)



(b)



(c)

圖 7.1 (a)左影像特徵點和彩色碼 (b)右影像特徵點和彩色碼 (c)160 組對應特徵點

將對應的特徵點座標帶入計算公式求得基本矩陣後，我們可在兩張影像上繪製極線，左影像和右影像的極線分佈如圖 7.2。左右影像的極線距離平均誤差為 10.473、9.308，極點至極線的距離的平均誤差為 5.23、6.985，由於存在錯誤的對應組合，所有極線並不完全交於一點，接下來我們使用

6.4 節的演算法對極線距離和極點至極線的距離設不同門檻值，以改善特徵點的對應，並討論實驗結果。

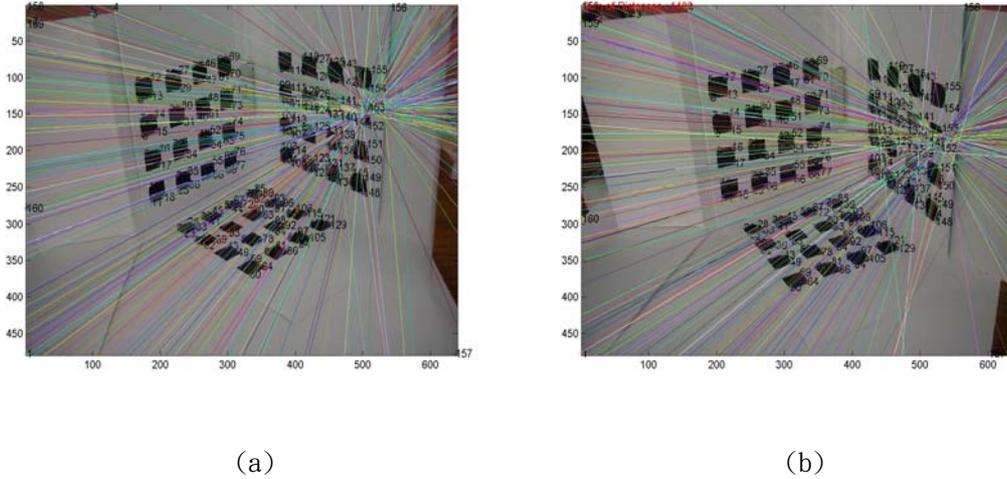


圖 7.2 (a)160 組對應點左影像極線分佈圖 (b)160 組對應點右影像極線分佈圖

7.1.1 僅對極線距離設限制



首先計算左右影像每一組特徵點的極線距離，對左右影像每一組的極線距離取平方後相加，並找出左右影像 n 組極線距離平方和的中位數，令為 M_J ，將 M_J 代入(7.1)式可到 σ ，其中 n 為對應的特徵點組數。

$$\sigma = 1.4826 \times [1 + 5/(n - 8)] \times \sqrt{M_J} \quad (7.1)$$

根據 σ ，我們可以找出衡量對應點正確與否的門檻值 $(2.5\sigma)^2$ ，如(7.2)式，若 $\omega_i=1$ ，代表此組對應可當作正確的對應；若 $\omega_i=0$ ，則此組對應可視為錯誤的對應，經由極線距離的限制後，我們可以刪除錯誤的對應組合。

$$\omega_i = \begin{cases} 1 & \text{if } r_i^2 \leq (2.5\sigma)^2 \\ 0 & \text{otherwise} \end{cases} \quad (7.2)$$

$$r_i^2 = d^2(m_i', Fm_i) + d^2(m_i, F^T m_i') \quad (7.3)$$

表 7.1 僅對極線距離設限制

對應的特徵點組數	160 組	150 組	145 組
左影像和右影像極線距離平方和的門檻值	1235.9	789.65	573.558
極線距離平均誤差 (左影像/右影像)	10.473 / 9.308	7.474 / 7.809	5.218 / 5.075
極點至極線的距離平均誤差 (左影像/右影像)	5.23 / 6.985	6.053 / 6.182	0.685 / 0.575

在 160 組對應點的例子中，我們使用極線距離平方和的門檻值加以限制，可刪除 10 組特徵點，以剩餘的 150 組對應點重新計算基本矩陣，並使用程式計算出來新的極線距離平方和的門檻值加以限制，可再去掉 5 組特徵點，再以 145 組對應點重新計算基本矩陣，此時發現所有的特徵點都符合極線距離的限制。表 7.1 為原來兩種距離的平均誤差(極線距離平均誤差、極點到極線的距離平均誤差)和更新後的平均誤差比較，距離誤差的單位均為像素，從表中我們發現不管是極線距離平均誤差或極點到極線的距離平均誤差都呈現下降的趨勢。

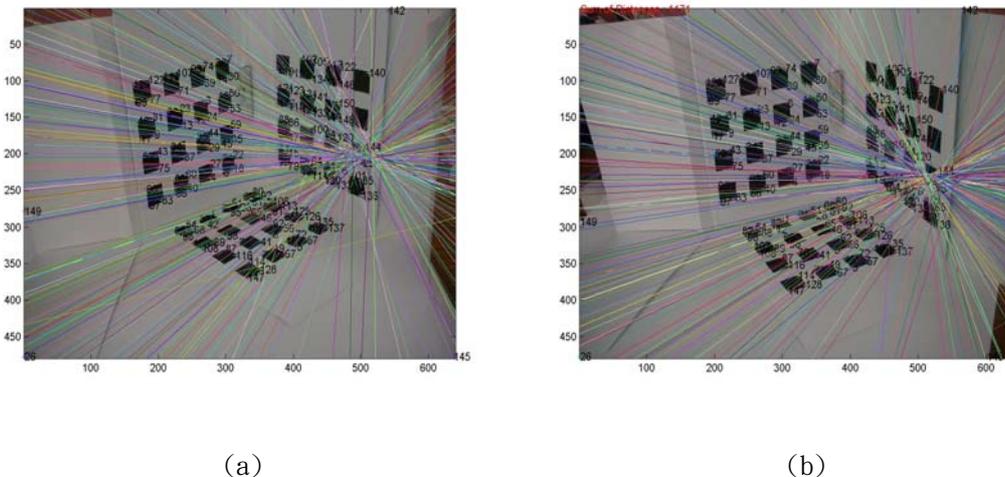
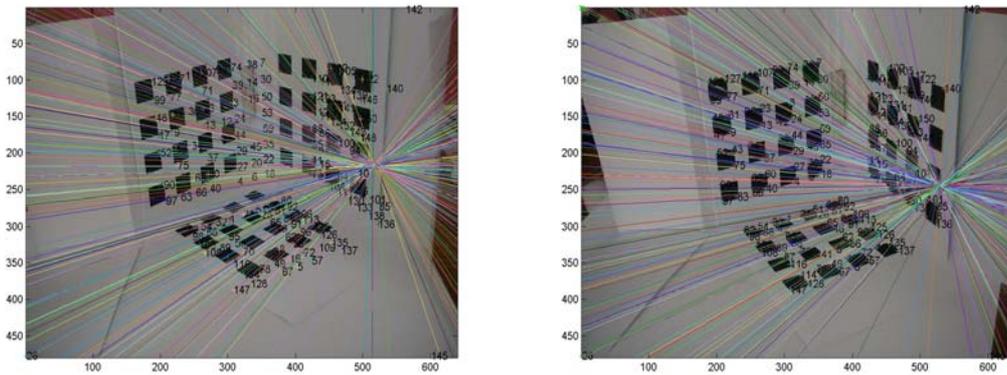


圖 7.3 (a)150 組對應點左影像極線分佈圖 (b)150 組對應點右影像極線分佈圖
(c)145 組對應點左影像極線分佈圖 (d)145 組對應點右影像極線分佈圖



(c)

(d)

圖 7.3 (續)

圖 7.3 為重新計算基本矩陣以後的極線分佈圖，剩下 150 組對應點時的極線分佈圖，不管是圖 7.3(a)或 7.3(b)，都還有極線未交於極點上，而這種情況到了特徵點僅剩 145 組時，已有大幅改進，而改善的過程經過三次計算基本矩陣和驗證，一共花了 76 秒。



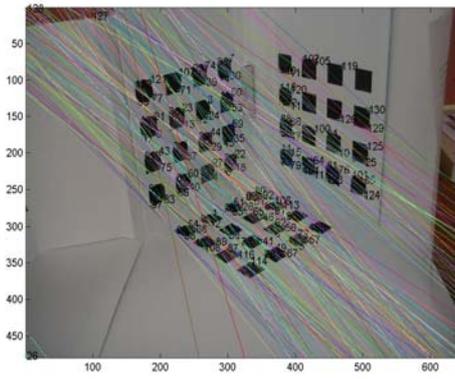
7.1.2 僅對極點到極線的距離設限制

要計算極點到極線的距離，必須先知道極點的座標。在 6.2 節中，我們介紹了兩種計算極點的方式，在程式裡面我們將採用 pseudo-inverse 來計算極點(pseudo-inverse 的計算過程請參考 6.2 節)，然後計算左影像和右影像每一組極線到極點的距離。有了極線到極點的距離後，考慮對稱性，我們分別對左影像和右影像設門檻值，同一組對應點在左影像和右影像都符合限制才視為正確的對應點，以第五列門檻值 5/5 為例，左影像和右影像的距離皆小於 5 才被當作正確的對應點，表 7.2 是 160 組對應點僅針對極點至極線的距離設限制的結果。由上到下門檻值越來越小，條件漸趨嚴格，符合條件的對應點組數也越來越少。

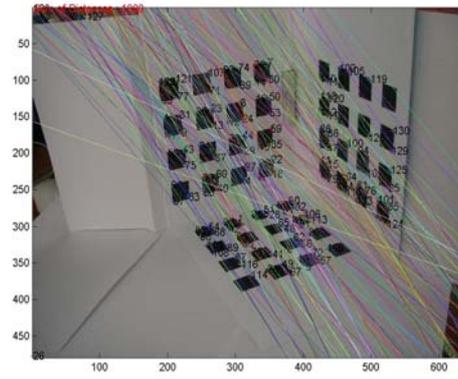
表 7.2 僅對極點到極線的距離設限制

原來的特徵點組數	極點到極線的距離的門檻值 (左影像/右影像)	更新後的特徵點組數	極線距離平均誤差 (左影像/右影像)	極點至極線的距離平均誤差 (左影像/右影像)
160 組	10 / 10	130 組	7.284 / 7.697	66.065 / 71.1
160 組	5/5 or 4/4	107 組	6.023 / 7.066	57.56 / 20.316
160 組	3 / 3	103 組	6.488 / 6.679	29.947 / 13.17
160 組	2 / 2	86 組	3.918 / 2.262	24.94 / 13.69
160 組	1 / 1	51 組	3.959 / 3.812	0.977 / 0.687

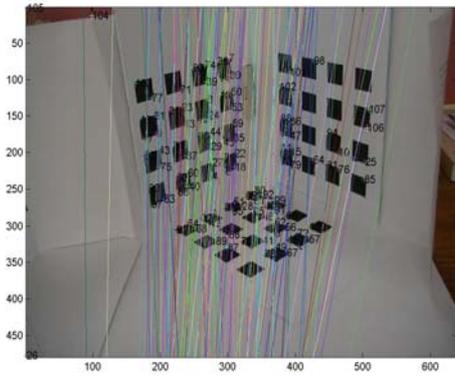
原來 160 組左影像和右影像極線距離平均誤差為 10.473 / 9.308，極點至極線的距離平均誤差為 5.23 / 6.985，觀察表 7.2。我們發現當極點到極線的距離的門檻值越嚴格，符合條件的特徵點組數越來越少，重新計算基本矩陣後，極線距離平均誤差呈現下降的趨勢，除了門檻值從 2 到 1 時，因為極線距離總誤差的下降幅度比不上特徵點組數的減少，所以看起來平均誤差反而變大。極點到極線的距離平均誤差呈現變大的情況，主要是因為如果只限制極點到極線的距離，可能會去掉原本極線距離很小的對應組合，保留錯誤且去掉正確的對應會造成計算出的基本矩陣無法反映左右影像間的幾何關係，表 7.2 裡二至五列的 130 組、107 組、103 組、86 組都屬於這樣的情況，即使極線距離平均誤差下降，但極點至極線距離的平均誤差卻是增加的，因為所有組合的誤差都超過門檻值，所以無法再計算基本矩陣而停止，仔細了解後我們發現：這四個例子中，極線距離前幾名的組合都佔了極線距離總誤差很大的部分，也大大的影響基本矩陣計算的精確性，以 130 組的右影像為例，極線距離總誤差為 1000.06，極線距離最大的五組分別為 147.28、73.57、40.83、26.75、26.18，這五組應該是具有錯誤，因為它們的誤差比平均誤差超過太多。以下我們將挑兩個錯誤的例子(130 組、107 組)和一組正確的例子(51 組)，並繪製它們的極線分佈圖。



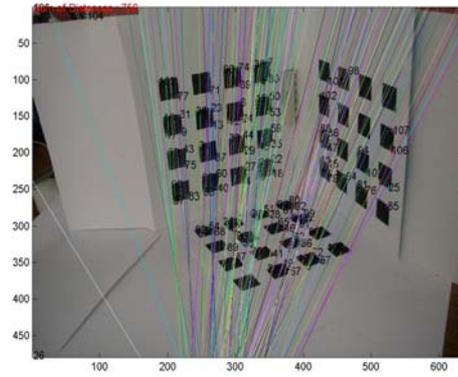
(a)



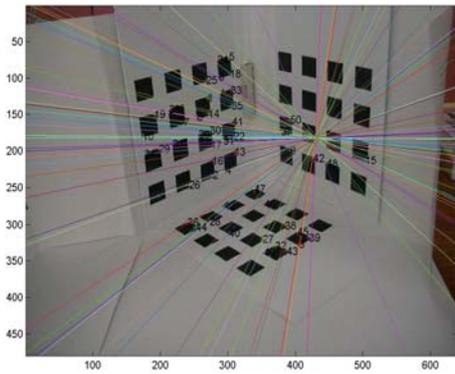
(b)



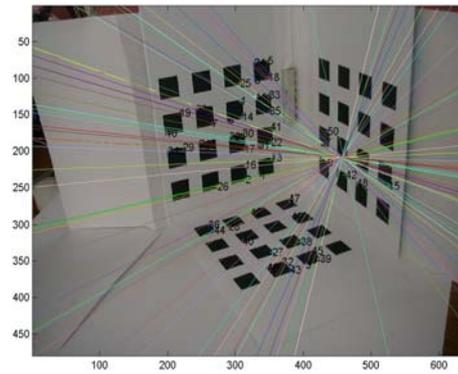
(c)



(d)



(e)



(f)

圖 7.4 (a)130 組對應點左影像極線分佈圖 (b)130 組對應點右影像極線分佈圖
(c)107 組對應點左影像極線分佈圖 (d)107 組對應點右影像極線分佈圖
(e)51 組對應點左影像極線分佈圖 (f)51 組對應點右影像極線分佈圖

表 7.2 的五組實驗結果都經過兩次計算基本矩陣和驗證的過程，實驗時間由上至下依次為 44.469、44.219、43.671、43.25、43.093 秒，由圖 7.4(a)~7.4(d)我們可以發現：若極點到極線的距離的門檻值限制太寬鬆，在下次重新計算基本矩陣的過程可能會去掉一些原本極線距離很小的對應組合，卻保留錯誤的對應組合，導致所有極線無法交於極點上，以 130 組的右影像為例，如果我們刪除極線距離最大的 5 組對應點，再重算一次基本矩陣並畫出極線分佈，如圖 7.5，就可得到我們想要改善特徵點對應的目的。

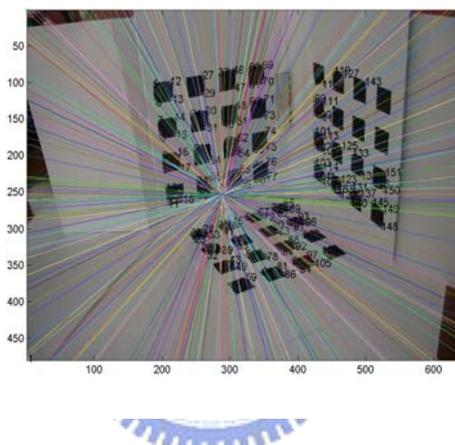


圖 7.5 圖 7.4(b)去掉極線距離誤差最大的 5 組對應點並重算基本矩陣，可得正確的結果

7.1.3 同時對極線距離和極點到極線的距離設限制

在這一節裡，我們將討論同時對兩種距離設限制的實驗結果，並和單獨使用某一種限制的結果比較。在表 7.3 裡，左影像和右影像極線距離平方和門檻值為 1235.9，極點到極線的距離的門檻值皆為 10，符合條件的對應點共有 126 組，利用這 126 組算出新的基本矩陣和新的極線距離門檻值 1060.314，極點到極線的距離的門檻值仍是 10，符合條件的對應點剩下 121 組，重新計算基本矩陣和極線距離門檻值，並設極點到極線的距離的門檻

值為 10 後，發現無法再去除對應點，我們發現重新計算基本矩陣後，極線距離平均誤差和極點到極線的距離平均誤差都呈現下降的趨勢。

表 7.3 同時對兩種距離設限制

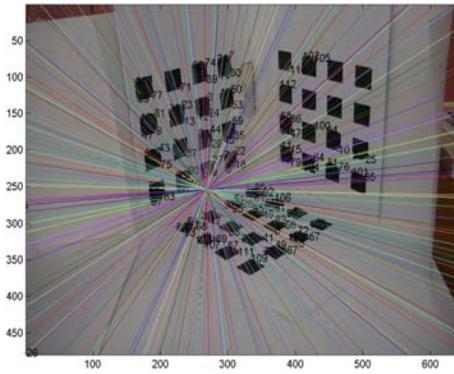
對應的特徵點組數	160 組	→ 126 組	→ 121 組
左影像和右影像極線距離平方和的門檻值	1235.9	1060.314	1032.2
極點到極線的距離的門檻值 (左影像/右影像)	10 / 10	10 / 10	10 / 10
極線距離平均誤差 (左影像/右影像)	10.473 / 9.308	5.964 / 6.416	1.79 / 1.802
極點至極線的距離平均誤差 (左影像/右影像)	5.23 / 6.985	1.521 / 1.664	1.3992 / 1.57

實驗經過三次計算基本矩陣和驗證的過程，一共花了 73.742 秒。

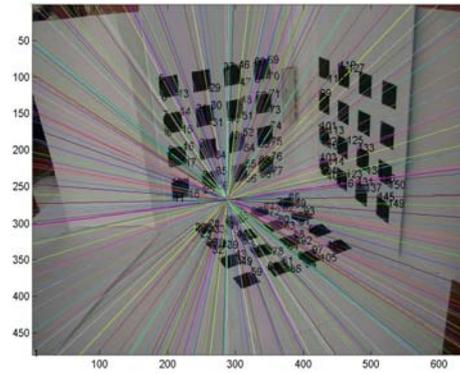
表 7.4 同時對兩種距離設限制

原來的特徵點組數	左影像和右影像極線距離平方和的門檻值	極點到極線的距離的門檻值(左影像/右影像)	更新後的特徵點組數	極線距離平均誤差(左影像/右影像)	極點至極線的距離平均誤差(左影像/右影像)
160 組	1235.6	8 / 8	113 組	6.096 / 6.654	0.2554 / 0.2737
160 組	1235.6	5 / 5	103 組	5.855 / 6.394	0.447 / 0.4987
160 組	1235.6	1 / 1	51 組	3.959 / 3.812	0.977 / 0.687

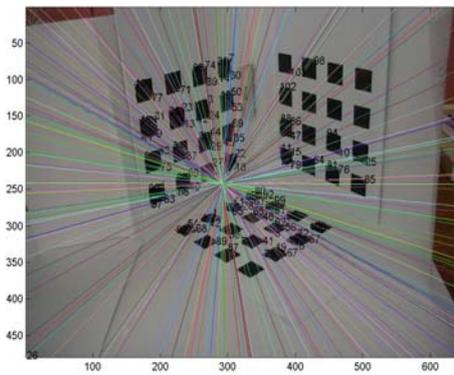
實驗經過兩次計算基本矩陣和驗證的過程，實驗時間由上至下依次花了 44.31、43.7、43.1 秒。表 7.4 和表 7.3 的差別在於極點到極線的距離的門檻值不同，隨著該門檻值的下降，極線距離平均誤差也逐漸減少，極點到極線的距離平均誤差雖然呈現擾動的現象，但那是因為總誤差遞減速度比不上對應點組數的減少，結果仍在可接受範圍內，接著我們將挑出 113 組和 103 組的極線分佈圖做代表，觀察改善特徵點對應的情形，結果如圖 7.6。



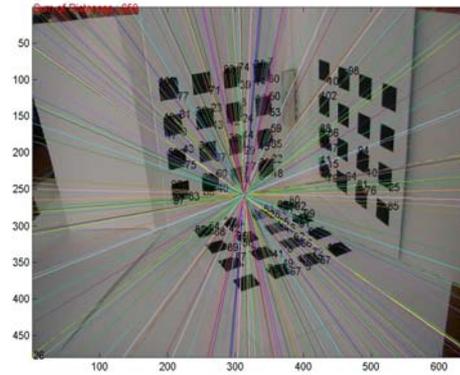
(a)



(b)



(c)



(d)

圖 7.6 (a)113 組對應點左影像極線分佈圖 (b)113 組對應點右影像極線分佈圖
(c)103 組對應點左影像極線分佈圖 (d)103 組對應點右影像極線分佈圖

比較表 7.1~7.4，我們可以得到一個結論：不管是在極線距離平均誤差或是極點到極線的距離平均誤差方面，同時使用兩種距離限制會比只使用某一種限制效果更好，且計算基本矩陣的次數越多，所花時間也越長。

另一個實驗是以 Canon PowerShot G1 數位相機取得的影像，影像大小為 640×450 ，調整 Harris 演算法的各項參數分別對兩張影像擷取特徵點後，左影像可得 35 個特徵點，而右影像可得 24 個特徵點，利用彩色碼進行特徵點對應後，我們一共得到 22 組的對應點，如圖 7.7(c)。左影像擷取角點的時間為 14.453 秒、右影像擷取角點的時間為 14.188 秒、對應時間

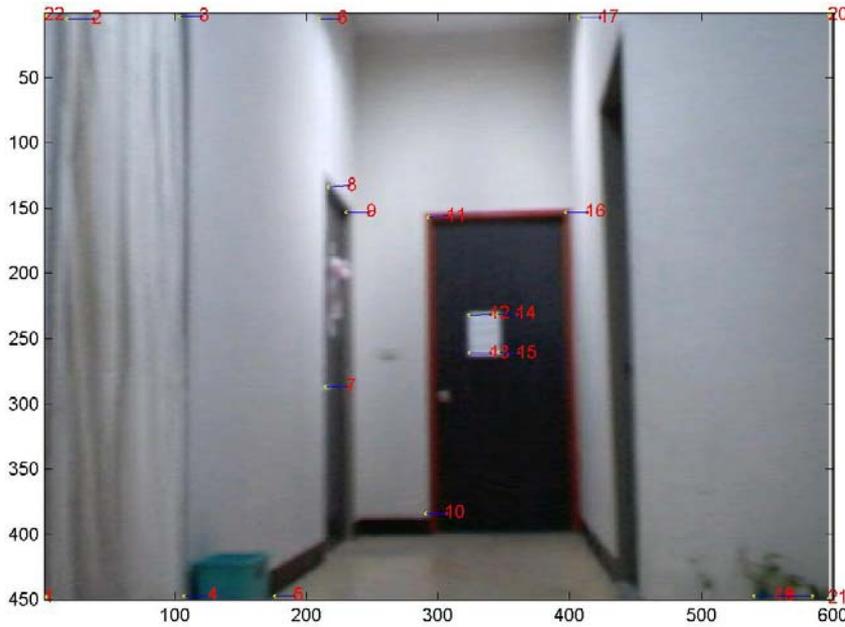
一共是 1.547 秒。



(a)



(b)



(c)

圖 7.7 (a)左影像特徵點和彩色碼 (b)右影像特徵點和彩色碼 (c)22 組對應特徵點

原來 22 組對應點左影像和右影像極線距離平均誤差是 0.828 及 0.827 個像素，極點至極線距離平均誤差為 1.533 及 1.526 個像素，由於特徵點的組數較少，能夠改善的空間也相對有限，以表 7.5 為例，整個改善特徵點對應的流程一共花了 10.34 秒，可見不管是特徵點擷取、對應或是改善

特徵點對應所花的時間都和特徵點的多寡有關。

表 7.5 不同場景的實驗

原來的特徵點組數	左影像和右影像極線距離平方和的門檻值	極點到極線的距離的門檻值(左影像/右影像)	更新後的特徵點組數	極線距離平均誤差(左影像/右影像)	極點至極線的距離平均誤差(左影像/右影像)
22 組	6.4352	10 / 10	20 組	0.515 / 0.516	0.663 / 0.661

7.2 重算基本矩陣和不重算基本矩陣的影響

以原來 160 組對應點的右影像為例，若我們僅刪除影像中錯誤的對應組合，但不重算基本矩陣，則極線距離平均誤差可從表 7.6 看出端倪，在這個例子裡，我們不討論極點至極線的距離平均誤差的原因是因為若採用 pseudo-inverse 計算極點，由於特徵點對應組數的減少，極線數目也會減少，算出的極點位置也不一樣；但基本矩陣並沒有改變，所以基本矩陣反推出來的極點位置仍相同，因此我們不考慮極點至極線的距離平均誤差。

表 7.6 不改變基本矩陣

原來的特徵點的組數	左影像及右影像極線距離門檻值	更新後的特徵點組數	原來極線距離平均誤差	更新後極線距離平均誤差
160 組	25	150 組	9.3081	6.5283
160 組	12	140 組	9.3081	4.8261
160 組	10.4	120 組	9.3081	4.697
160 組	6.35	80 組	9.3081	2.8931
160 組	2.92	40 組	9.3081	1.514

我們可以發現雖然平均誤差呈現遞減的趨勢，但是極線距離平均誤差 1.514 時只剩下 40 組特徵點，雖然節省了一點計算時間，但特徵點組數太

少有時候對於我們了解整個三維場景並沒有幫助，從表 7.3 中，我們所提出的方法在 121 組特徵點的時候就有相同的結果，雖然重新計算基本矩陣，但比起以往 RANSAC 等等的方法計算次數少，且更有助於三維重建。

圖 7.8 是 150、120、80、40 組右影像的極線分佈圖，由於沒有重算基本矩陣，所以即使 40 組特徵點的極線距離平均誤差只有 1.514，但不是所有極線都交於極點之上。

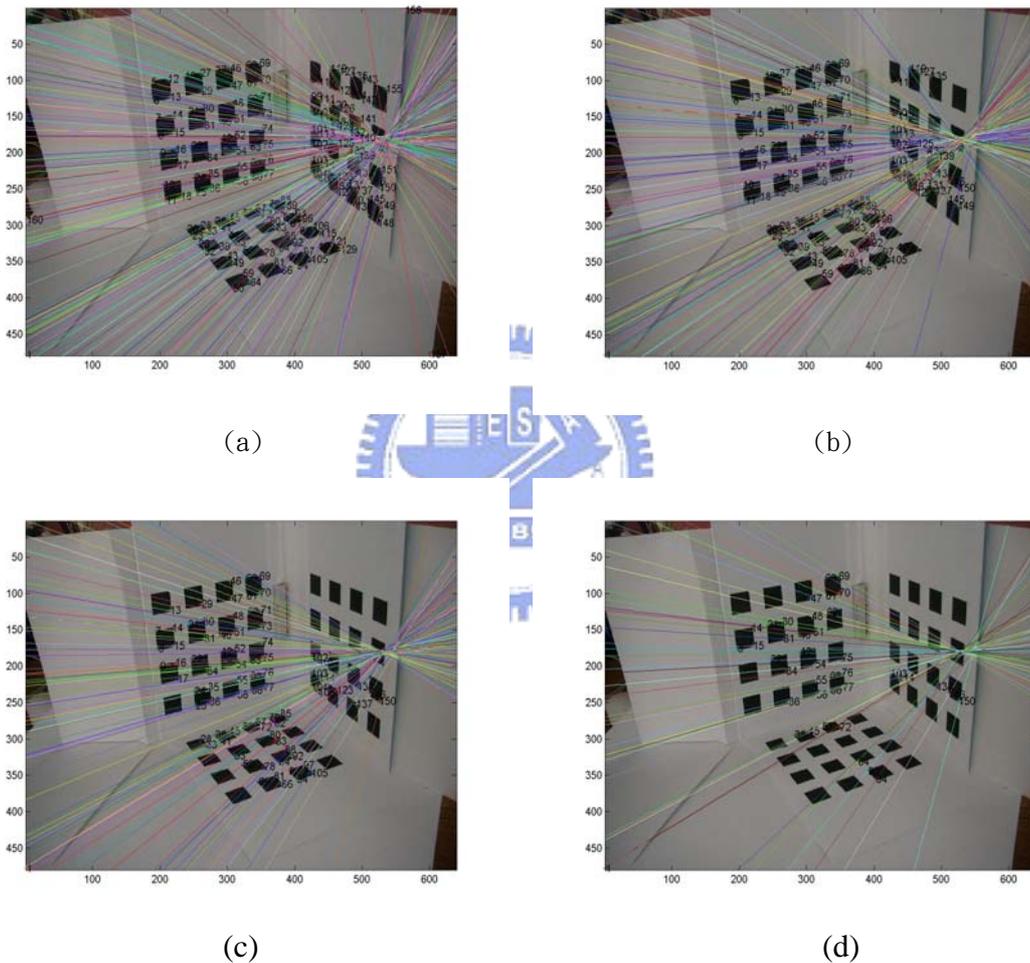


圖 7.8 未改變基本矩陣狀況下 (a)150 組對應點右影像極線分佈圖 (b)120 組對應點右影像極線分佈圖 (c)80 組對應點右影像極線分佈圖 (d)40 組對應點右影像極線分佈圖

7.3 極點位置的不確定性

在 6.2 節中我們提到兩種極點的計算方式，在這一節裡我們將討論實驗過程中產生的極點座標，同樣的可分成三種情況：僅使用極線距離限制、僅使用極點至極線的距離限制、同時使用兩種距離限制。

表 7.7 極點位置(僅對極線距離設限制)

對應的特徵點組數	160 組	→ 150 組	→ 145 組
左影像和右影像極線距離平方和的門檻值	1235.9	789.65	573.558
用 pseudo-inverse 計算極點座標	左 (509.86,149.11) 右 (532.21,179.28)	左 (507.42,205.88) 右 (517.9,229.96)	左 (518.45,217.78) 右 (529.3,244.6)
用基本矩陣反推極點座標	左 (513,153.04) 右 (530.69,174.51)	左 (507.93,210.51) 右 (518.87,224.94)	左 (518.4,218.56) 右 (529.44,243.79)

表 7.8 極點位置(僅對極點到極線的距離設限制)

原來的特徵點組數	極點到極線的距離的門檻值 (左影像/右影像)	更新後的特徵點組數	用 pseudo-inverse 計算極點座標	用基本矩陣反推極點座標
160 組	10 / 10	130 組	左 (185.52,109.09) 右 (348.62,201.09)	左(-3640.6,3077.9) 右(3830.7,-3492.6)
160 組	5/5 or 4/4	107 組	左 (226.4,1126) 右 (262.37,674.92)	左(-105.39,1236.8) 右 (629.02,615.44)
160 組	3 / 3	103 組	左 (169.99,1037.9) 右 (235.26,627.24)	左 (33.232,944.76) 右 (406.38,612.93)
160 組	2 / 2	86 組	左(-512.37,166.22) 右(-221.73,117.65)	左(-570.41,322.93) 右(-202.01,-52.938)
160 組	1 / 1	51 組	左 (413.69,184.63) 右 (440.4,210.01)	左 (414.17,185.47) 右 (440.13,209.07)

表 7.9 極點位置(同時對兩種距離設限制)

對應的特徵點組數	160 組	→ 126 組	→ 121 組
左影像和右影像極線距離平方和的門檻值	1235.9	1060.314	1032.2
極點到極線的距離的門檻值(左影像/右影像)	10 / 10	10 / 10	10 / 10
用 pseudo-inverse 計算極點座標	左 (509.86,149.11) 右 (532.21,179.28)	左 (268.64,243) 右 (288.24,253.46)	左 (273.69,242.56) 右 (293.42,253.42)
用基本矩陣反推極點座標	左 (513,153.04) 右 (530.69,174.51)	左 (268.04,242.3) 右 (288.62,254.32)	左 (273.17,241.97) 右 (293.75,254.14)

表 7.10 極點位置(同時對兩種距離設限制)

原來的特徵點組數	左影像和右影像極線距離平方和的門檻值	極點到極線的距離的門檻值(左影像/右影像)	更新後的特徵點組數	用 pseudo-inverse 計算極點座標	用基本矩陣反推極點座標
160 組	1235.6	8 / 8	113 組	左(268.69,254.28) 右(287.81,267.57)	左(268.57,254.18) 右(287.9,267.7)
160 組	1235.6	5 / 5	103 組	左(292.05,243.06) 右(313.35,257.86)	左(292.23,243.16) 右(313.21,257.74)
160 組	1235.6	1 / 1	51 組	左(413.69,184.63) 右(440.4,210.01)	左(414.17,185.47) 右(440.13,209.07)

觀察表 7.7~7.10，我們發現：不管是利用 pseudo-inverse 或是利用基本矩陣反推極點位置，計算結果都是非常相似的，也呼應了我們去除影像中的雜點後，兩種距離平均誤差呈現下降趨勢的結果，除了表 7.8 的第二列到第五列等四個例子，因為基本矩陣無法完整表達每組特徵點之間相對關

係情形，所以兩種極點計算方式結果相差很大。

Csurka 在[21]中提到極點位置的不確定性，以左影像為例，若計算出的極點座標 $A = [x, y]^T$ 呈現高斯分佈，則 A 落在橢圓內的機率可寫成(7.4)式，其中 Λ_A 是 A 的共變異量矩陣(covariance matrix)， $E[A]$ 是 A 的期望值， k 是任意常數，如此一來，我們可用橢圓方程式表示極點可能出現的位置，

$$(A - E[A])^T \Lambda_A^{-1} (A - E[A]) = k^2 \quad (7.4)$$

我們提出的改善特徵點對應流程，經過這一章的實驗證明具有一定的正確性，可以提供目前特徵點對應問題一個有效的改善方法，而我們的實驗結果同時也符合極點的不確定性。



第八章 結論

我們利用 Harris 演算法擷取特徵點，在 Harris 演算法計算過程中，同時也會產生每一像素的影像梯度，針對影像梯度的方向加以分析，可將梯度方向量化為九個方向，並以九個顏色替代不同的方向，我們稱之為彩色碼。彩色碼足以建立初步的特徵點對應，藉由比對和相鄰對應點的距離、角度的關係，可以解決模稜兩可的對應問題。但在特徵點擷取的途中，Nonmaximum Supression 所選取出 R 值最大的點不一定就是最正確的特徵點，或是找出不屬於特徵點的角點，此時可能會造成對應的錯誤，為提高對應的準確性，我們加上極線限制的特性，去除極線距離誤差和極點至極線的距離誤差過大的對應點後，我們可以找出正確的對應組合，和以往的方法相比，減少了重複計算基本矩陣的次數，可節省時間。

除了極線限制之外，彩色碼的比對仍有改進的空間，因為量化的結果可能產生多組彩色碼相異數和彩色碼距離結果相同的位置，需要靠其他方法來判別，若已知一些成功的對應，也可利用極線之交比(cross ratio)驗證其他對應組合，另外實驗中用兩種極點計算方式求出的極點位置也符合極點的不確定性，這些主題都是未來可以深入探討的方向。

參考文獻

- [1] Z. Zhang, R. Deriche, O. Faugeras, and Q.-T. Luong. A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry. *Artificial Intelligence Journal*, 78:87–119, October 1995.
- [2] Fischler, M. and Bolles, R.: 1981, Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography, *Communications of the ACM* 24, 381 – 385.
- [3] Rousseeuw, P. and Leroy, A.: 1987, *Robust Regression and Outlier Detection*, John Wiley & Sons, New York.
- [4] Torr, P. and Murray, D.: 1993, Outlier detection and motion segmentation, in P. Schenker (ed.), *Sensor Fusion VI, SPIE Vol. 2059*, Boston, pp. 432 – 443.
- [5] Z. Zhang. Determining the epipolar geometry and its uncertainty: A review. *The International Journal of Computer Vision*, 27(2):161–195, 1998.
- [6] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, 2001.
- [7] G. Csurka, C. Zeller, Z. Zhang, and O. Faugeras. Characterizing the uncertainty of the fundamental matrix. *Computer Vision and Image Understanding*, 68(1):18–36, October 1997.
- [8] Longuet-Higgins, H.: 1981, A computer algorithm for reconstructing a scene from two projections, *Nature* 293, 133 – 135.
- [9] Faugeras, O.: 1993, *Three-Dimensional Computer Vision: a Geometric Viewpoint*, The MIT Press.
- [10] H. P. Moravec, “Toward Automatic Visual Obstacle Avoidance,” *Proc. of 5th International Joint Conference on Artificial Intelligence*, pp. 584, August 1977.

- [11] C. Harris and M. Stephens, "A Combined Corner and Edge Detector," *Proc. 4th Alvery Vision Conf. Manchester*, pp. 189-192, August 1988.
- [12] S. M. Smith and J. M. Brady, "SUSAN: A New Approach to Low Level Image Processing," *International Journal of Computer Vision*, pp. 45-78, 1997.
- [13] H. Asada and M. Brady, "The Curvature Primal Sketch," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 8, pp. 2-14, 1986.
- [14] R. Deriche and G. Giraudon, "2-D Curve Matching Using High Curvature Points : Application to Stereo Vision," *Pattern Recognition*, pp. 240-242, June 1990.
- [15] R. Horaud, F. Veillon, and T. Skordas, "Finding Geometric and Relational Structures in an Image," *Proc. First European Conference On Computer Vision, Antibes, France*, pp. 374-384, April 1990.
- [16] A. Fusiello, V. Roberto, E. Trucco. "Efficient Stereo with Multiple Windowing," *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 858-863, June 1997.
- [17] Quang-Tuan Luong, "*Matrice Fondamentale et Calibration Visuelle sur l'Environnement-Vers une plus grande autonomie des syst`emes robotiques.*" PhD thesis, Universit´e de Paris-Sud, Centre d'Orsay, December 1992.
- [18] Hartley, R.: 1995, In defence of the 8-point algorithm, *Proceedings of the 5th International Conference on Computer Vision*, IEEE Computer Society Press, Boston, MA, pp. 1064 – 1070.
- [19] O. Faugeras and Q.-T. Loung, "The fundamental matrix : Theory, Algorithms, and Stability Analysis," *Int'l J. Computer Vision*, 17, pp. 43-75, 1996.
- [20] P.H.S. Torr and D.W. Murray, "The Development and Comparison of Robust Methods for Estimating the Fundamental Matrix," *Intl J. Comput. Vision*, vol. 24, no. 3, pp. 271–300, 1997.
- [21] F. Mokhtarian and R. Suomela, "Robust Image Corner Detection Through Curvature Scale Space," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 20, no. 12, 1998.