

國立臺灣師範大學機電科技學系
碩士論文

指導教授：陳美勇 博士

使用轉角偵測與虛擬網格法重建三維模型

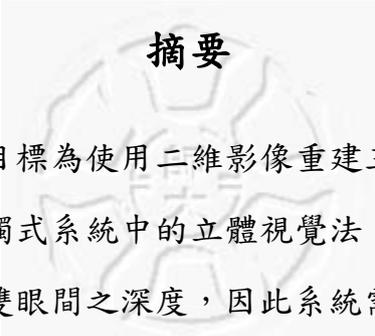
3D Object Model Recovery from 2D Images Utilizing Corner

Detection and Virtual Mesh Grid



研究生：黃盈源 撰

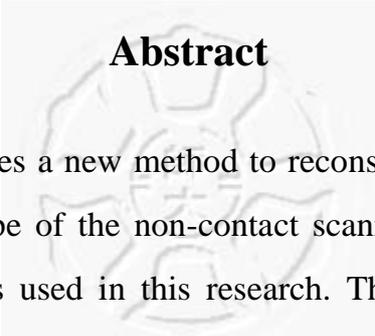
中華民國 一百年 七月



摘要

本研究主要之研究目標為使用二維影像重建三維物件模型。一般研究上所使用的方法為非接觸式系統中的立體視覺法，此方法模擬人類雙眼感知影像進而推算物體與雙眼間之深度，因此系統需要使用兩隻攝影機進行影像的擷取，擷取後的兩張影像進行匹配找出現實空間中一點分別投影至二維影像的投影點，之後利用現實空間與相機座標系統彼此間的幾何轉換關係，藉由兩張影像上的投影點計算出此一點於現實空間中之深度資訊，如此一來即可重建出物體的三維模型。然而立體視覺法中困難之處在於如何從左右兩影像準確地找出相對應的投影點進行深度計算，因此針對此問題在過去的研究提出從外部投影一結構光至物件表面，藉由結構光協助系統定位左右兩影像中相對應的投影點，然而此法受限於物體表面之顏色。故本研究提出之三維重建方法無須藉由投影結構光即可重建出物件之三維模型，針對簡單幾何物件以及曲面物件分別使用轉角偵測以及虛擬網格協助系統定義左右兩影像中相對應的投影點重建出物件之三維模型，簡單幾何物件之特徵點通常出現於輪廓之轉角處，因此系統透過轉角偵測找出左右兩影像中物件輪廓之特徵點，藉由立體視覺法重建起特徵點之深度資訊將特徵點重建至三維座標空間中，再根據重建之特徵點還原出物件之三維模型。在另一方面，由於曲面物件不同於簡單幾何物件在轉角處有明顯之特徵點，因此本研究先於左物件影像建立起虛擬網格，藉由極線幾何原理估測曲面物件左右兩影像中相對應的投影點，於右物件影像建立起相對應的虛擬網格，根據左右兩物件影像之虛擬網格以立體視覺法成功地重建出曲面物件之三維模型。

關鍵字：三維重建、立體視覺法、轉角偵測、虛擬網格。



Abstract

This research proposes a new method to reconstruct the 3D object model from 2D images. One type of the non-contact scanning measurement for the stereo vision algorithm is used in this research. The stereo vision simulates human's eyes to capture the depth information of the object. Therefore, this research uses two CCD Cameras to capture two images of the object. Then, find out the match points from the two images. Using the match points and combine 1)the parameters of the two CCD Cameras and 2)transform matrix between the world coordinate and camera coordinate to get the depth information of each point in the space. Finally, the object's 3D model can be reconstructed. The important issue of the stereo vision theorem is how to find out the match points from the two images accurate. For solving this issue in the past researches, many articles used a projected structure light on the object's surfaces to measure the match points. In this research, the proposed system is able to find out the match points from the two images by the structure light. But this method will be restricted by the color of the object surface. This research proposes a method to reconstruct the 3D model without projecting the structure light. The system uses corner detection and virtual mesh grid to reconstruct the simple geometry and curved the surface of object. The feature points of the simple geometry object are usually on the corner of the contour. So we can find out the feature points by doing the corner detection, and then the system would calculate the depth of the feature points to project the feature points in the 3D coordinated space. And then, the simple geometry object's 3D model would be reconstructed from these feature points. But the curved surface object doesn't have the visible feature points, therefore, this paper build up the virtual mesh grid from the left image. Then, the system would estimate the match points by the epipolar geometry theorem and builds up the virtual mesh grid on the right

image. Finally system reconstructs the 3D model by the stereo vision theorem and virtual mesh grid of the two images successfully.

Keywords : 3D reconstruction, Stereo vision, Corner detection, Virtual mesh grid.

誌謝

隨著炎炎夏日即將進入尾聲，我的學生生涯也將告一段落邁入人生的另一階段，在此特別感謝指導教授 陳美勇 老師在我大學四年與碩士班兩年期間的提攜與教導，老師教導我的不僅只有課業、技術、還有在學術研究上應有的正確觀念以及克服難題之方法，讓我在碩士班的兩年內培養了研究、發現問題與自行解決問題的能力，除此之外最重要的老師還教導我待人處事的基本態度，不僅是學業上的良師也是我人生旅途上重要的指標。

感謝口試委員 莊鎮嘉 博士、黃正民 博士以及 郭景明 博士的協助，給予了我最佳的建議使我的研究以及論文能更為完整嚴謹。

此外也要感謝實驗室的學長、同學與學弟們，在我研究期間能彼此相互扶持勉勵，在研究上能互相討論解決問題，讓我的研究得以更為順利進行，每次報告時給予的建議也奠定了我日後對眾人報告的能力，有幸在求學的生涯裡遇到這群好友豐富了我的研究生活，因此也期勉實驗室的各位能順利完成自己的目標。

最後感謝家裡在我求學的路上一路的扶持，因為有父母親無怨無悔的付出與照顧才能有今日的我，最後期勉我自己在未來的日子裡能朝著目標筆直的前進，回饋於我最愛的家庭，謹以此論文獻給我最愛的家人以及過去所有關心支持我的師長與好友們，感謝你們。

黃盈源

辛卯年·仲夏夜

精密運動控制實驗室



目 錄

摘要.....	I
Abstrcat.....	II
誌謝.....	IV
目錄.....	V
圖目錄.....	VII
表目錄.....	X

第一章 緒論	1
1.1 前言.....	1
1.2 文獻回顧.....	7
1.3 研究動機與目的.....	10
1.4 本論文之貢獻.....	12
1.5 論文架構.....	12
第二章 理論基礎	13
2.1 數位影像基本定義.....	13
2.2 鄰域處理.....	14
2.3 影像邊緣偵測(Edge detection).....	16
2.3.1 Canny 邊緣偵測.....	18
2.4 影像二值化.....	18
2.5 形態學影像處理(morphology).....	19

2.5.1 膨脹(Dilation)與侵蝕(Erosion).....	20
2.6 轉角偵測(corner detection).....	22
2.6.1 Harris 轉角偵測.....	23
2.7 透視投影法.....	23
2.8 立體視覺法(Stereo vision).....	25
2.9 Fundamental matrix.....	26
第三章 系統設計概念與配置.....	29
3.1 三維重建系統設計實現目標.....	29
3.2 三維重建系統架構.....	29
3.3 三維重建系統架設概念.....	30
3.4 三維重建系統配置.....	30
3.5 本重建系統之特色.....	33
3.5 重建系統流程描述.....	33
第四章 三維模型重建系統設計原理.....	35
4.1 系統前處理.....	35
4.2 影像前處理.....	36
4.3 搜尋 match points.....	38
4.3.1 epipolar line.....	38
4.3.2 簡單幾何物件搜尋 match point.....	39

4.3.3 曲面物件搜尋 match point.....	42
4.4 建立虛擬三維空間座標.....	50
4.5 重建三維模型.....	53
4.6 結合運動模型.....	54
第五章 實驗結果與討論.....	56
5.1 系統設備描述.....	56
5.2 三維模型重建系統架設操作流程.....	58
5.3 三維模型重建流程.....	59
5.3.1 相機參數與求解 Fundamental matrix.....	59
5.3.2 重建簡單之幾何物件.....	62
5.3.3 重建曲面物件.....	67
5.4 三維模型重建結果分析.....	72
第六章 結論及未來展望.....	76
參考文獻.....	77





圖 目 錄

圖1-1	指紋辨識.....	1
圖1-2	車牌辨識.....	1
圖1-3	人臉辨識.....	1
圖1-4	(a) X-ray 所照射的骨骼二維影像[4]；(b)骨骼之三維模型.....	2
圖1-5	(a)(b)人臉辨識系統藉由攝影機所照射的二維人臉影像重建三維的人臉模型.....	2
圖1-6	(a)環境探勘機器人；(b)環境三維模型；(c)(d)利用三維模型所建置出來的環境地圖以及路徑規畫.....	3
圖1-7	(a)六個自由度機械手臂；(b)重建物件之三維模型資訊使機械手臂能以最佳姿勢進行夾取.....	3
圖1-8	(a) (b)CMM 使用探針量測物件表面輪廓.....	4
圖1-9	(a) (b)使用雷射掃描系統對於環境進行掃描，並重建其三維環境模型.....	4
圖1-10	(a)立體視覺法使用兩攝影機進行影像擷取；(b) 投影一結構光至物件表面協助系統進行重建；(C)立體視覺法其原理是根據兩相機間的幾何關係；(d)立體視覺法重建之三維模型.....	4
圖1-11	(a) 色度成型法利用物件表面之反射光來估測物件的表面輪廓；(b) 重建結果.....	6
圖1-12	(a)光學法利用不同角度的光源進行照射；(b)(c)(d)結合立體視覺法；(e)(f)重建結果.....	6

圖1-13	(a)投影網狀結構光至物件表面上；(b)將網格點進行編碼；(c)利用網格點使用立體視覺法進行重建；(d)實驗結果.....	8
圖1-14	(a)實驗設備；(b)投影光闌於物件表面；(c)將影像上之光闌進行編碼；(d)找出 match point 可能出現之區域為極線(epipolar line)；(e)使用光闌與極線之交點定位出 match point；(f)重建結果.....	9
圖1-15	圖 1-15 (a)實驗流程；(b) 標誌圖形；(c)將標誌圖形張貼於物件輪廓與表面上；(d)還原出曲線之三維模型；(e) 重建結果將三維的弧形線條進行曲面的掃出.....	10
圖2-1	二維影像座標.....	14
圖2-2	灰階分佈.....	14
圖2-3	遮罩示意圖.....	15
圖2-4	鄰域處理示意圖.....	15
圖2-5	(a)水平濾波器；(b)垂直濾波器；(c) 45°濾波器.....	16
圖2-6	(a)影像邊緣；(b)一階導數；(c)二階導數示意圖.....	17
圖2-7	影像直方圖.....	19
圖2-8	集合理論(a)A 集合與 B 集合；(b) $A \cup B$ ；(c) $A \cap B$ ；(d) $A - B$ ；(e) $(A)^C$	20
圖2-9	遮罩 A 集合與二值影像 B 集合.....	20
圖2-10	使用遮罩對影像進行膨脹示意圖.....	21
圖2-11	遮罩 A 集合與二值影像 B 集合.....	21

圖2-12	使用遮罩對影像進行侵蝕示意圖.....	22
圖2-13	針孔成型示意圖.....	24
圖2-14	透視投影.....	24
圖2-15	Epipolar geometry.....	25
圖2-16	Epipolar geometry.....	26
圖3-1	三維重建系統配置圖.....	30
圖3-2	相機配置圖(a)相機等高設置;(b)相機相互平行.....	31
圖3-3	偏光鏡(a)反光造成的干擾;(b)加裝偏光鏡濾除反光.....	32
圖3-4	兩 CCD 攝影機與待測物之設置.....	32
圖3-5	左右兩 CCD 攝影機擷取之影像(a)左影像;(b)右影像.....	33
圖3-6	三維重建系統整體流程.....	34
圖4-1	透視投影法.....	35
圖4-2	影像中各灰階值數量的分布情形.....	37
圖4-3	epipolar geometry.....	39
圖4-4	(a)(b)(c)簡單幾何物件(c)(d)(f) 簡單幾何物件之特徵點.....	40
圖4-5	(a)(b)(c)藉由特徵點重建出物件之三維模型.....	40
圖4-6	match point 最佳化.....	42
圖4-7	(a)物件之封閉輪廓;(b)網格產生之起始點與終點.....	43
圖4-8	網格延著 x 軸的方向進行切割.....	44

圖4-9	網格延著 y 軸的方向進行切割.....	45
圖4-10	網格切割所遺失的資訊.....	45
圖4-11	使用 3x3 矩陣做為遮罩對每一個網格點進行掃描.....	46
圖4-12	(a)未偵測到輪廓線；(b)網格點位於輪廓線上；(c)輪廓線位於遮罩左右兩側；(d)輪廓線位於遮罩上下兩側；(e)輪廓線位於遮罩四端..	46
圖4-13	網格點優化步驟流程圖.....	47
圖4-14	網格點優化示意圖.....	48
圖4-15	將網格與物件剪影進行比對修剪.....	48
圖4-16	投影等高示意圖.....	49
圖4-17	投影側視圖.....	49
圖4-18	match point 定位示意圖.....	50
圖4-19	二維影像座標.....	51
圖4-20	相機座標.....	52
圖4-21	相機座標與虛擬三維空間座標.....	52
圖4-22	虛擬三維空間重建示意圖.....	54
圖4-23	沿 Z 軸旋轉示意圖.....	55
圖5-1	三維模型重建系統配置圖.....	56
圖5-2	影像擷取卡.....	57
圖5-3	攝影機.....	57

圖5-4	鏡頭.....	58
圖5-5	三維重建系統實體配置圖.....	59
圖5-6	校正板.....	60
圖5-7	(a)左校正板影像；(b)右校正板影像.....	60
圖5-8	傾斜個角度之左校正板影像.....	61
圖5-9	傾斜個角度之右校正板影像.....	61
圖5-10	已知點於空間之分佈圖.....	62
圖5-11	(a)背景影像；(b)輸入之左影像；(b)物件之影像.....	62
圖5-12	物件影像之輪廓.....	63
圖5-13	轉角偵測.....	63
圖5-14	最佳 match point 之找尋.....	64
圖5-15	(a)實體物件；(b)重建結果；(c)模型前視圖；(d)模型上視圖.....	64
圖5-16	模型之位置資訊.....	65
圖5-17	(a)實體物件；(b)重建結果；(c)模型前視圖；(d)模型上視圖.....	65
圖5-18	模型之位置資訊.....	66
圖5-19	(a)實體物件；(b)重建結果；(c)模型側視圖；(d)模型上視圖.....	66
圖5-20	模型之位置資訊.....	67
圖5-21	(a)物件原始影像；(b)物件之剪影.....	67
圖5-22	網格分佈.....	68

圖5-23	右影像 match points 之分佈.....	68
圖5-24	(a)原始物件前視圖；(b)原始物件上視圖；(c)重建結果；(d)模型前視圖；(e)模型上視圖.....	69
圖5-25	模型之位置資訊.....	69
圖5-26	(a)原始物件前視圖；(b)原始物件上視圖；(c)重建結果；(d)模型前視圖.....	70
圖5-27	模型上視圖.....	71
圖5-28	模型之位置資訊(1).....	71
圖5-29	模型之位置資訊(2).....	71
圖5-30	模型之位置資訊(3).....	72

表目錄

表 5-1: 尺寸、距離比較表.....	73
表 5-2: 尺寸、距離比較表.....	73
表 5-3: 距離誤差表.....	74
表 5-4: 重建時間表.....	75

第一章 緒論

1.1 前言

隨著科技的進步影像處理已從二維影像延伸至三維影像，二維影像的應用已普遍出現於日常生活中了，例如：指紋辨識如圖 1-1[1]，車牌辨識如圖 1-2[2]，人臉辨識如圖 1-3[3]，以及產品檢測系統，但是隨著電腦硬體的迅速發展，影像技術的日新月異，二維平面影像技術已延伸至三維立體影像，三維影像與二維影像最大的差異在於深度資訊，因此如何取的深度資訊成為三維重建技術的重要課題。

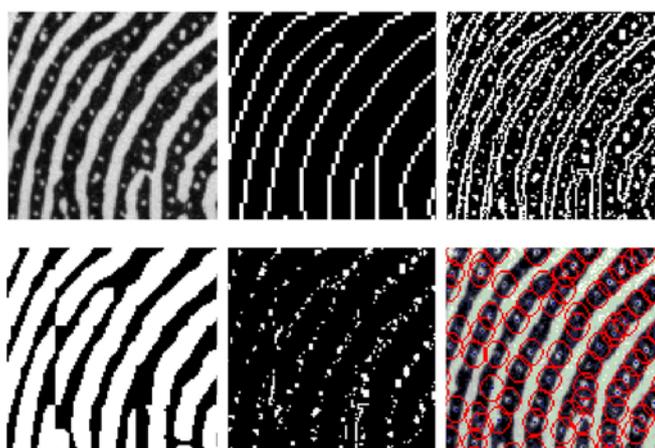


圖 1-1：指紋辨識[1]



圖 1-2：車牌辨識[2]



圖 1-3：人臉辨識[3]

三維重建技術已逐漸成熟目前已有許多領域有相關的應用，例如醫療領域，利用照射 X-ray 所產生的二維影像來重建骨骼模型如圖 1-4[4][5]，

人臉辨識也從二維延伸至三維模型，使用三維人臉模型進行辨識可增加辨識的成功率降低誤判的機率如圖 1-5[6][7]，三維重建技術也可使用在環境探勘機器人，藉由攝影機擷取的影像即時重建起周遭環境的三維模型，可協助機器人路徑的規劃或是障礙物的閃避如圖 1-6[8]，三維重建技術也可用在協助機械手臂夾持物件，重建三維物件的模型可使機械手臂以最佳的姿勢夾持物件如圖 1-7[9]。

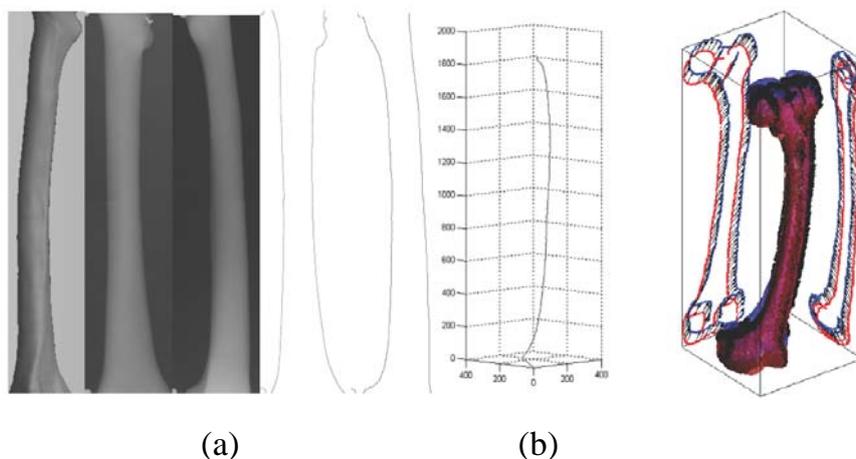


圖 1-4：(a) X-ray 所照射的骨骼二維影像[4]；(b) 骨骼之三維模型[5]

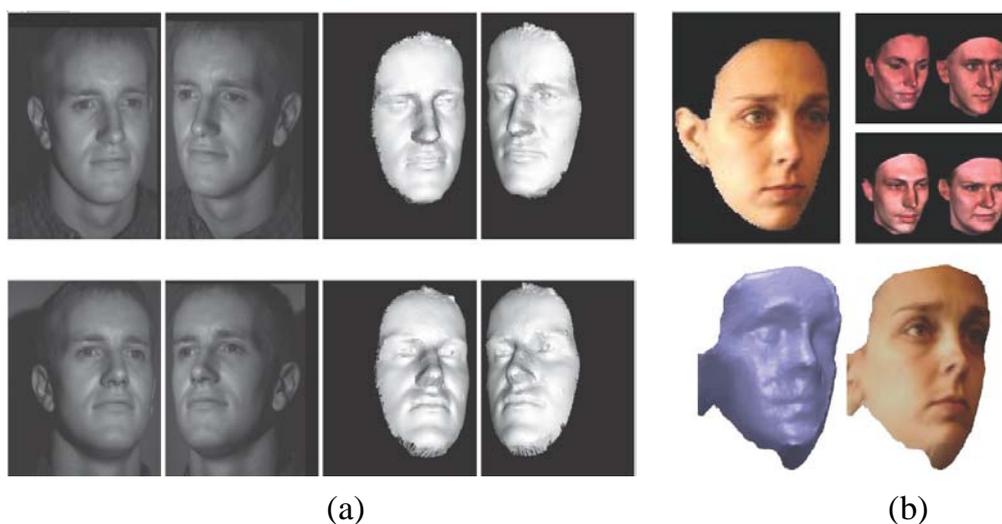


圖 1-5：(a) (b)人臉辨識系統藉由攝影機所照射的二維人臉影像重建三維的人臉模型[6][7]

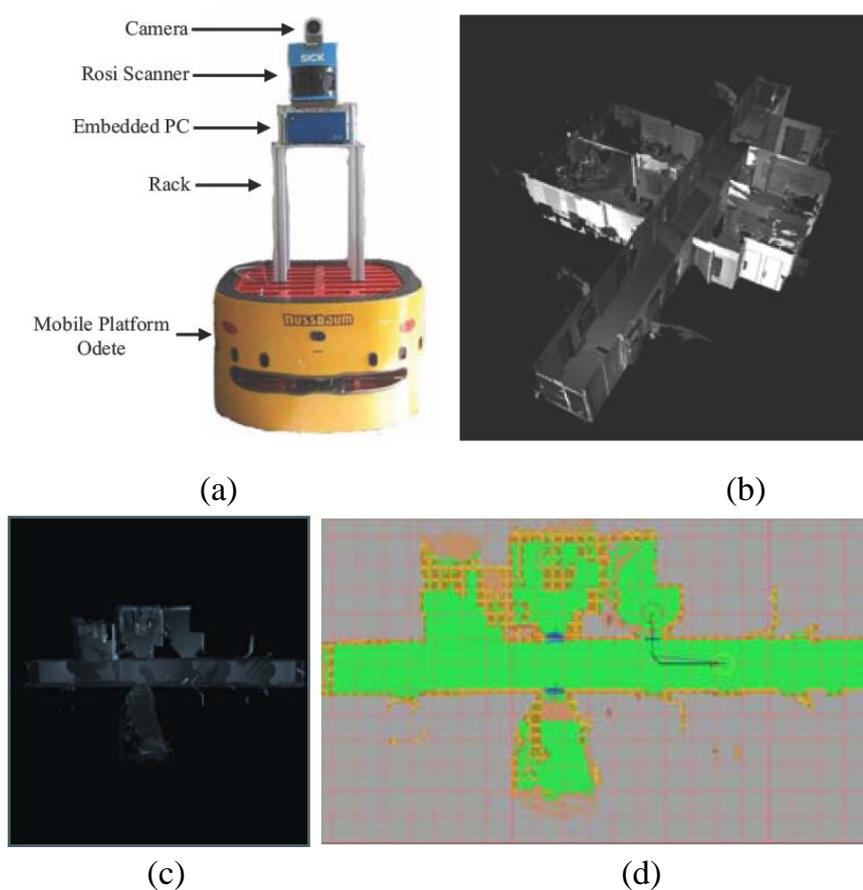


圖 1-6：(a)環境探勘機器人；(b)環境三維模型；(c)(d)利用三維模型所建置出來的環境地圖以及路徑規劃[8]

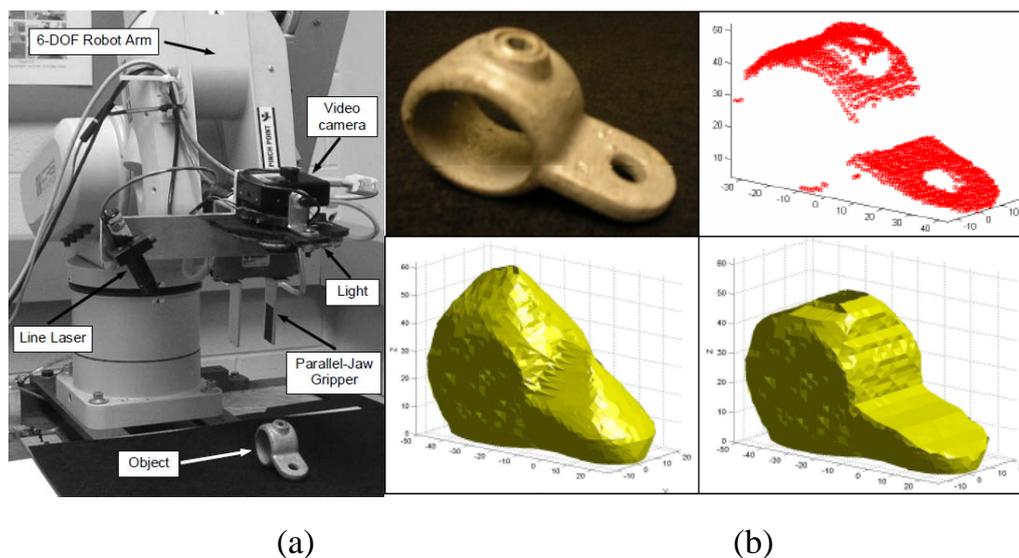


圖 1-7：(a)六個自由度機械手臂；(b)重建物件之三維模型資訊使機械手臂能以最佳姿勢進行夾取 [9]

目前三維重建技術主要分為接觸式以及非接觸式，接觸式技術需利用探針接觸物體表面以擷取物體的表面輪廓資訊，此方法能精確的量測出物件的表面輪廓，因此常被用於製造業，常見的座標測量機(CMM, Coordinate Measuring Machine)即是接觸式的三維掃描儀器，但是因為探針需與量測物件有所接觸(如圖 1-8)因此容易造成物件表面損傷，且需要花較長的時間完成物件的掃描[9]。

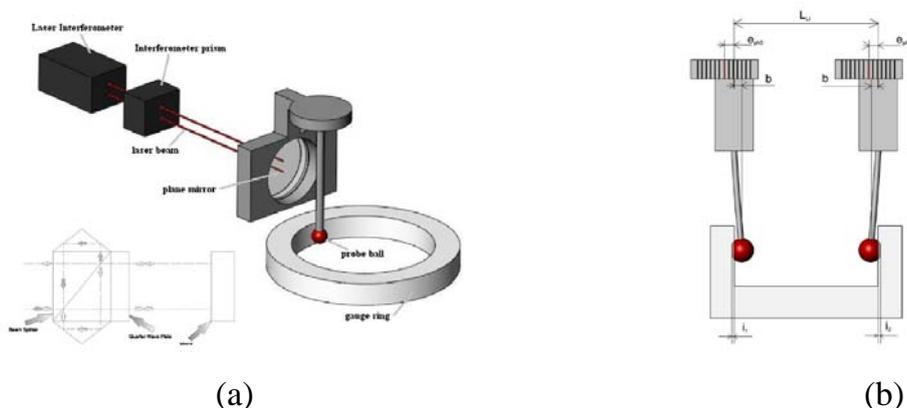


圖 1-8：(a) (b)CMM 使用探針量測物件表面輪廓 [10]

非接觸式的三維重建技術可分為主動式(active)與被動式(passive)的重建系統，主動式常見的有利用雷射來擷取深度，藉由雷射掃描物件表面或是環境進而重建出物件的三維輪廓或是環境三維模型如圖 1-9[11]，使用雷射掃描的優點是準確率高但相對的其花費時間較長設備也較為昂貴。

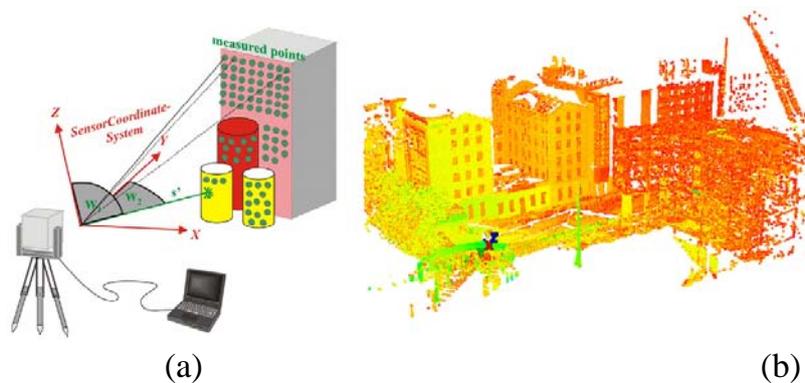


圖 1-9：(a) (b)使用雷射掃描系統對於環境進行掃描，並重建其三維環境模型 [11]

被動式則利用攝影機擷取二維影像，利用二維影像來換算深度資訊以及表面輪廓，進而重建出三維的模型，例如：立體視覺法(Stereo vision)其原理如同人類的雙眼，使用兩台攝影機擷取物體的二維影像，藉由兩張影像間的幾何關係求出其深度資訊，此方法常用於物件的追蹤，利用兩攝影機擷取影像藉由立體視覺法迅速的計算出目標物的位置以及距離等資訊，立體視覺法也可搭配由外部投影一結構光至物件表面上，輔助系統擷取物件表面的細節資訊以及對應點的定位進而準確的重建出物件的三維模型如圖 1-10[12]，此方法可以延伸至三台攝影機或是多台攝影機[13]，而本篇論文也是基於此方法進行延伸。

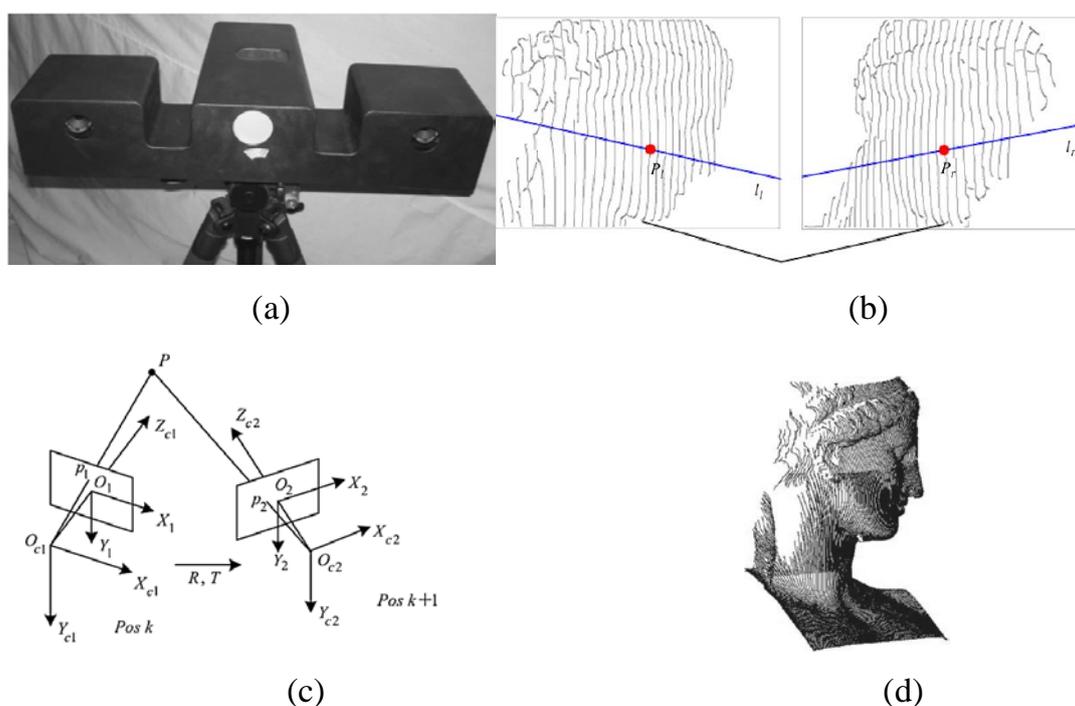


圖 1-10：(a)立體視覺法使用兩攝影機進行影像擷取；(b) 投影一結構光至物件表面協助系統進行重建；(C)立體視覺法其原理是根據兩相機間的幾何關係；(d)立體視覺法重建之三維模型[12]

色度成型法(Shape from shading)由 B.K.P Horn 提出，不同於立體視覺法，色度成型法將單一影像上每個像素的亮度值代入事先建立的色度模型中進而求出物體表面的深度資訊[14]，此方法後來延伸出立體光學法。立

體光學法(Photometric Stereo)改進色度成型法來獲得更多的表面資訊，一台攝影機固定從一方向照射物件，由不同方向的光源照射至物件上，每次僅有一光源進行照射，將所有的影像綜合利用漫射模型求出物件表面的梯度向量，進而得到此物件的三維模型[15]。

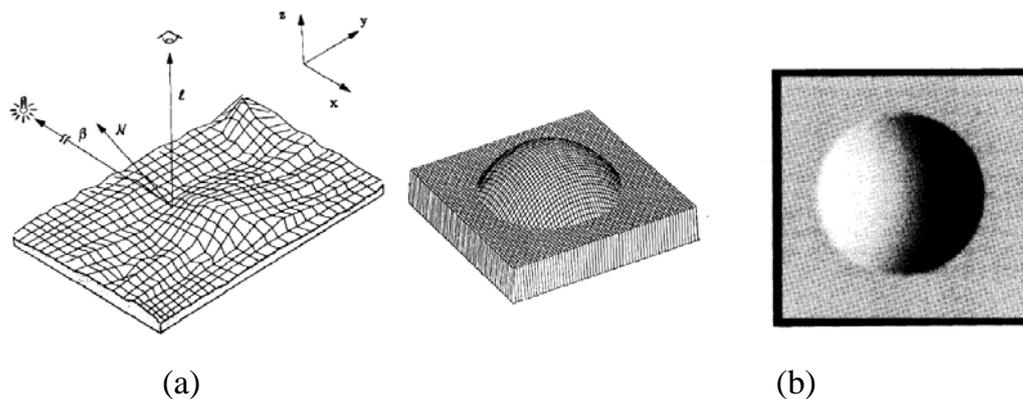


圖 1-11：(a) 色度成型法利用物件表面之反射光來估測物件的表面輪廓；(b)重建結果[14]

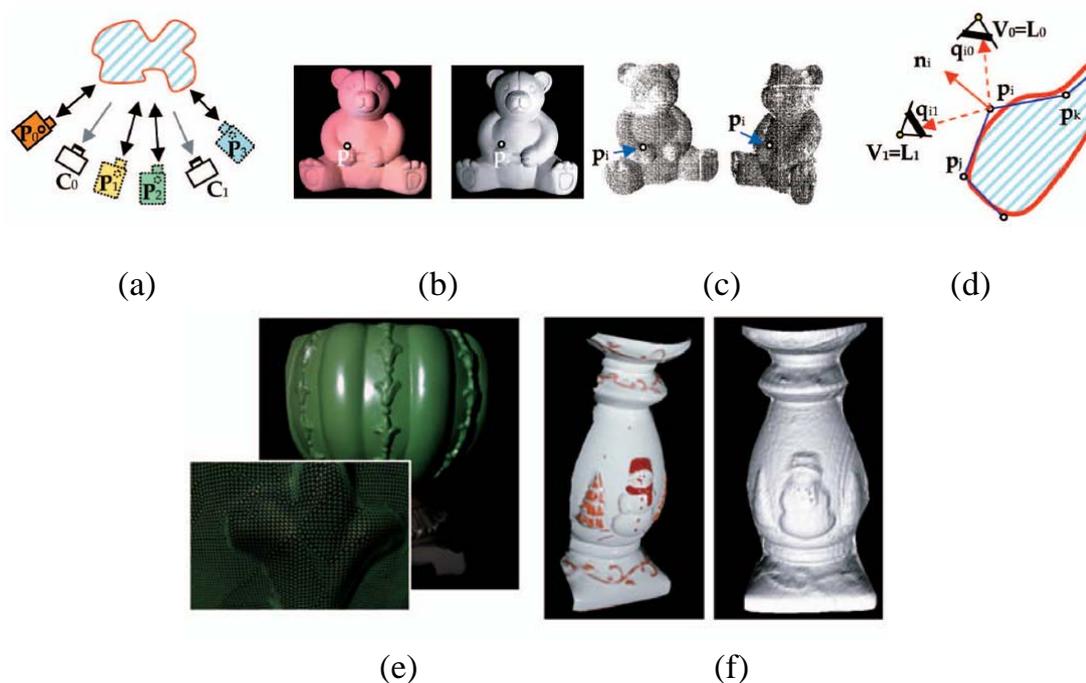


圖 1-12：(a)光學法利用不同角度的光源進行照射；(b)(c)(d)結合立體視覺法；(e)(f)重建結果[15]

比較上述多種的三維模型重建技術，接觸式的重建技術優點是精確度較高，細節部分也重建的出來，因此此種重建方法大多應用於製造業使用此技術重建工件的三維輪廓，或是對工件的表面輪廓進行檢測，但缺點就是需要花費大量的時間進行輪廓的掃描，也因為高精密度所以設備成本高。相對的被動式使用相機作為重建的工具，優點是成本低、重建速度快，但是其重建的準確度較低細節部分也難以重建，因此此種重建技術常用於環境探勘機器人或是輔助機械手臂進行物件的夾取，因為環境探勘機器人以及機械手臂需要的是即時的資訊，因此需要迅速的建立起三維模型，且只需要大致的物件輪廓即可，綜觀起來使用者可根據自己的需求以及目的選擇適當的重建方法來進行三維模型的重建。

1.2 文獻回顧

本研究重建三維物件模型主要目的是希望能提供物件之外觀以及距離等資訊給機械手臂進行夾取或是環境探勘機器人等所使用，因此重建的速度必須迅速如此一來才能即時提供物件的三維資訊，除此之外重建的三維物件模型也須維持一定之準確率，根據以上之需求本研究選擇立體視覺法(stereo vision)作為重建之基礎理論，立體視覺法主要是模擬人類的雙眼感知影像進而推算物體深度，因此使用立體視覺法系統需要使用兩隻攝影機進行影像的擷取，擷取後的兩張影像進行匹配找出現實空間中之一點分別投影至左右兩設影機的投影點，之後利用現實空間與相機座標系統彼此間的幾何轉換關係，藉由兩張影像上的相對應投影點(match points)計算出此一點於現實空間中之深度資訊，如此一來即可重建出物體的三維模型，也可計算出此物件與攝影機間之距離。

使用立體視覺法最為困難也最為主要的部分就是如何從左右兩影像找出相對應的投影點進而還原出此點現實空間中之深度資訊，因此如何準

確且有效率的從左右兩影像找出相對應的投影點成為大家改善的主要議題，以下將介紹幾篇研究其主要目的就是使用各種方法準確的從左右兩影像找出相對應的投影點。

Philippe Lavoie, Dan Ionescu, and Emil M. Petriu [16]等學者，提出一方法投影網狀的結構光於物件表面上，之後使用兩攝影機同時擷取此物件影像，首先對擷取進來的影像進行影像處理得到清楚的物件輪廓，之後對兩影像上的網格點進行標記，因此只要從兩影像上找出標記為同一點之網格即可準確的重建此點之三維座標，由於網格點是由外部用結構光投影至物件表面上的，因此使用此法找出的相對應投影點誤差小，但缺點是結構光未照射之部分則無法重建，所以這篇研究先重建起所有網格點之三維座標，細部部分則根據周圍已重建的網格點之三維座標使用線性內差法進行重建，如此一來即可使重建起的三維物件模型外觀更為平順，其實驗架構如圖 1-13 所示。

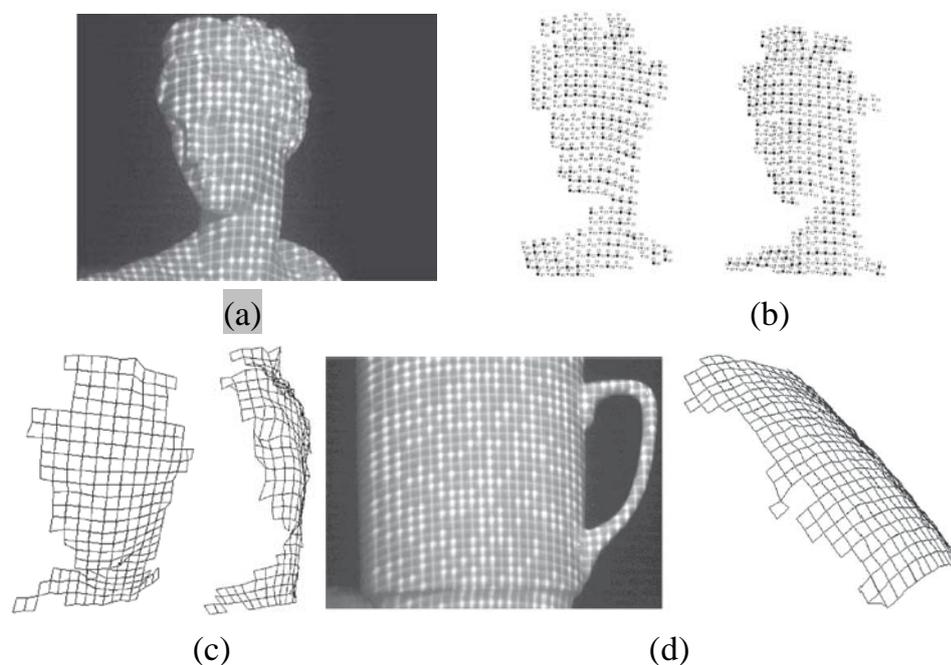


圖 1-13：(a)投影網狀結構光至物件表面上；(b)將網格點進行編碼；(c)利用網格點使用立體視覺法進行重建；(d)實驗結果[16]

LiMei Song , DaNv Wang [17]等學者，提出方法為投影光閘結構於物件表面上，同樣使用兩攝影機擷取物件影像，將影像中之光閘藉由影像處理提取出來，並且同樣的將每條光線進行編碼設定，利用立體視覺法極線幾何原理(epipolar geometry)，左影像上一投影點其右影像相對應之投影點可能出現位置為一直線此一直線稱為極線(epipolar line)，此相對應投影點也同樣會出現在相同編碼上之光線，故只要找出極線與相同編碼之光線交點即可鎖定定位出相對應投影點之位置，接著使用立體視覺法根據找出的相對應投影點重建物件之三維模型，實驗架構、實驗流程與結果如圖 1-14 所示。

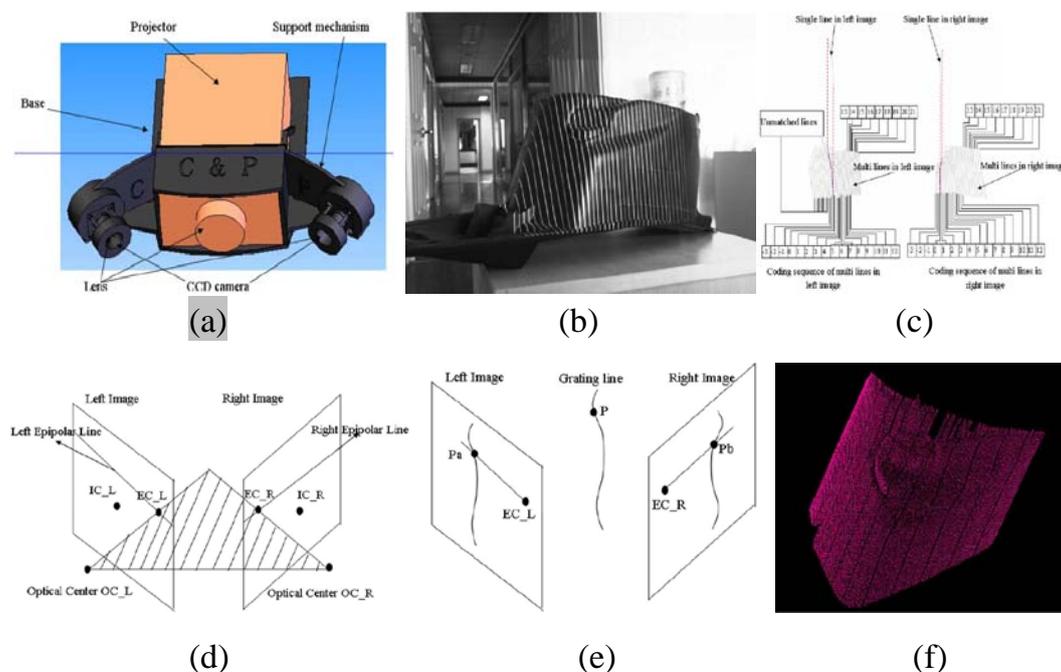


圖 1-14：(a)實驗設備；(b)投影光閘於物件表面；(c)將影像上之光閘進行編碼；(d)找出 match point 可能出現之區域為極線(epipolar line)；(e)使用光閘與極線之交點定位出 match point；(f)重建結果[17]

Zheng Jian-dong, Zhang Li-yan [18]等學者提出半自動式結合標誌圖形重建曲面的三維模型，設計兩種特殊的標誌圖形協助影像的擷取，利用第一種標誌圖形標示出物件的範圍，接著使用另一種標誌圖形貼於物件的

輪廓上，此研究使用手持式攝影機從各角度對物件擷取影像，從影像裡的標誌圖形將物件從背景裡提取出來，也使用這些標誌圖形進行相機校正還原出攝影機的空間座標，再根據物件上的標誌圖形設定物件輪廓的起點與終點，接著進行影像處理使用虛擬的弧形去近似物件的表面輪廓，最後將近似好的虛擬弧形與校正完成的攝影機空間座標以立體視覺法進行三維模型的重建，最後再將三維的弧形線條進行曲面的掃出，重建出曲面的三維模型，實驗流程與結果如圖 1-15 所示。

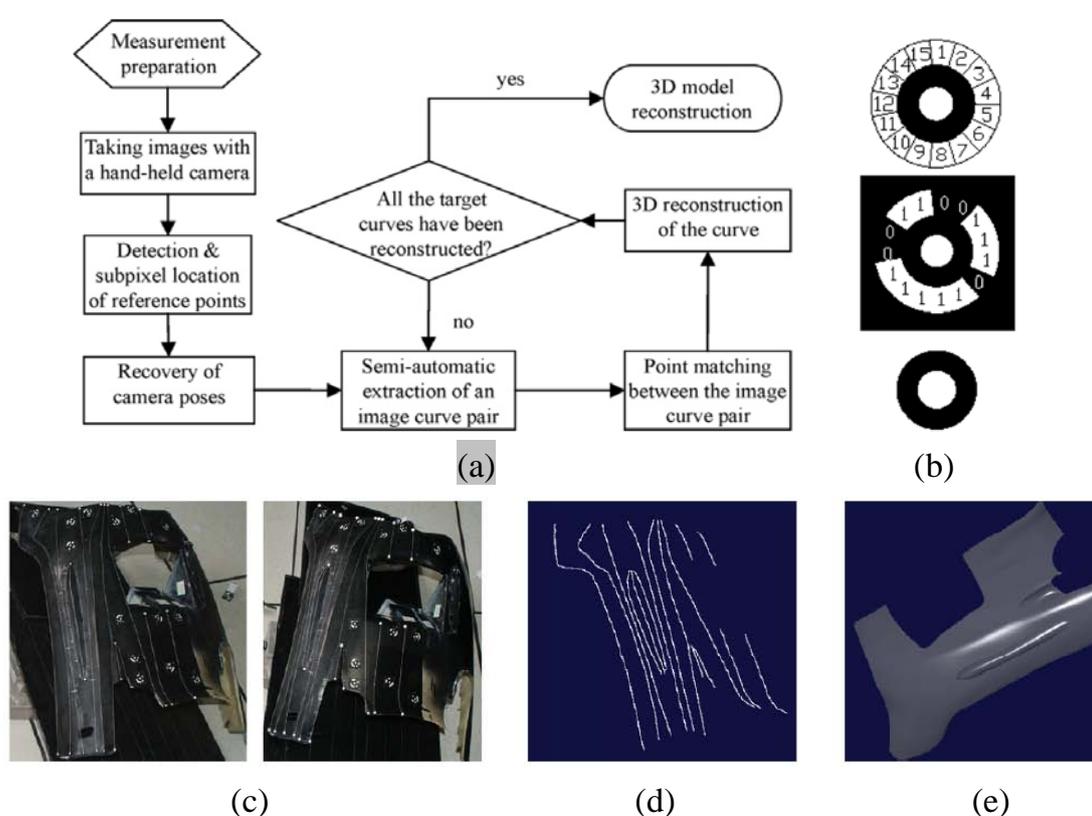


圖 1-15 (a)實驗流程；(b) 標誌圖形；(c)將標誌圖形張貼於物件輪廓與表面上；(d)還原出曲線之三維模型；(e) 重建結果將三維的弧形線條進行曲面的掃出[18]

1.3 研究動機與目的

隨著三維重建技術的逐漸成熟其應用已被廣泛的使用於各領域，其中

視覺技術與機械手臂的結合使機械手臂能有更為強大的功能，以往機械手臂進行工作其夾取的路徑以及目標均為固定且事先設定好的，若是配備視覺技術的機械手臂則能對未知位置的物件進行夾取，由裝置於機械手臂上之鏡頭擷取物件之影像搭配雷射掃描裝置，或是使用二維影像計算物件深度進而提供物件的外觀，距離等資訊協助機械手臂夾取目標物，本研究欲開發三維重建系統，其研究目的希望能提供更為完善以及準確的物件資訊協助機械手臂迅速且有效率的夾取未知位置以及外觀之物件且能以最佳姿勢穩固的夾取物件，為達到此一目的本研究致力於三維模型重建系統的開發，基於以上目的為了使機械手臂能即時且迅速的識別物件外觀以利機械手臂有效率的夾取物件，此三維模型重建系統必須快速、準確的重建出物件外觀模型還需計算出物件之位置，根據以上需求本研究選擇立體視覺法作為重建系統之開發基礎。本研究欲達成之目標列舉如下：

1. **快速重建三維模型**：降低資料之處理量，節省程式運算時間以達到即時提供資訊之目的。
2. **重建系統架構簡單**：本研究使用立體視覺法只需使用兩攝影機擷取物件之二維影像即可重建出物件之三維模型，以及計算出物件之位置資訊。
3. **無需使用額外設備**：不同於過往研究，本研究開發之系統無須額外使用設備投影結構光至物件表面，或是必須先在物件表面做標記以協助重建，提高系統使用之便利性。
4. **整體設備成本低廉**：此重建系統只需兩支攝影機即可進行物件之重建以及定位。

1.4 本論文之貢獻

本研究之貢獻，本研究之硬體設備只需兩支攝影機，無須使用額外的設備投影結構光，也無須於物件表面做記號即可重建出物件之三維模型，以往研究常見以投影設備投影結構光於物件表面，但是結構光容易受物件表面情況，反光或是顏色的影響，本研究的系統架構也無須使用標誌記號於物件表面，因此提供了最佳便利性，重建系統之技術本研究使用立體視覺法進行重建，提出新的方法結合虛擬網格之建立，因此可由虛擬網格重建起物件之三維模型，虛擬網格可針對物件之表面輪廓進行網格解析度之調整，可設以較高解析度之網格重建物件較細微之輪廓，以較大解析度之網格重建物件較為平滑或是大面積之部分，根據建立起來的虛擬網格即可以最少資料量重建起物件三維模型也可同時兼顧重建模型之完整性，如此一來系統可迅速且完整的提供物件三維模型資訊給機械手臂進行最佳姿勢來夾取物件。

1.5 論文架構

本論文一共分為六章，各章內容依序如下。

【第一章】- 緒論

【第二章】- 理論基礎

【第三章】- 系統組成設計與配置

【第四章】- 三維模型重建系統設計原理

【第五章】- 實驗結果與討論

【第六章】- 結論及未來展望

第二章 理論基礎

本章節將講述本研究系統架構所須之基本原理，包含有數位影像基本定義、數位影像處理、影像型態學、數位影像邊緣檢測原理，相機投影原理以及立體視覺法基礎原理，以便作為後續章節重建系統之基礎原理。

2.1 數位影像基本定義

現實空間可被表示成一個三維的函數 $g(x, y, z)$ ，現實空間中的物體透過相機將物體影像投影至投影面上產生二維影像，藉由將此影像數位化我們可使用電腦對數位影像進行處理或是儲存，數位影像如圖 2-1 所示，其中數位影像(digital image)可被表示成一個二維的函數 $f(x, y)$ ，其中 x 和 y 是二維空間座標，於座標軸 (x, y) 其相對應值 f 的大小稱為該點的強度(intensity)，灰階影像以灰階值(gray level)表示其強度如圖 2-2 所示，其強度範圍界於 0-255，因此我們可以將影像以 $n \times m$ 矩陣進行表示如式子 2-1，矩陣內元素的值為相對應之像素強度值。彩色影像每個點帶有 RGB 三個顏色的強度值，R 為紅色(Red)、G 為綠色(Green)、B 為藍色(Blue)，RGB 空間以這三種顏色合成任一顏色。數位影像中 x 、 y 與 f 的值皆為有限的離散量。一幅影像取樣的點越多則其數位影像越接近真實的影像，但相對的其資訊量會越龐大，一幅影像取樣的數量會決定 x 、 y 的值，取樣的多寡於數位影像中稱之為解析度，解析度決定了數位影像的品質以及大小，數位影像處理(digital image processing)是指藉由電腦處理數位影像，數位影像由大量的元素組成，而每個元素有特定的位置與數值，這些構成數位影像的元素稱為照片元素(picture element)、影像元素(image element)或是像素(pixel)，其中像素是最常用來表示數位影像之元素。

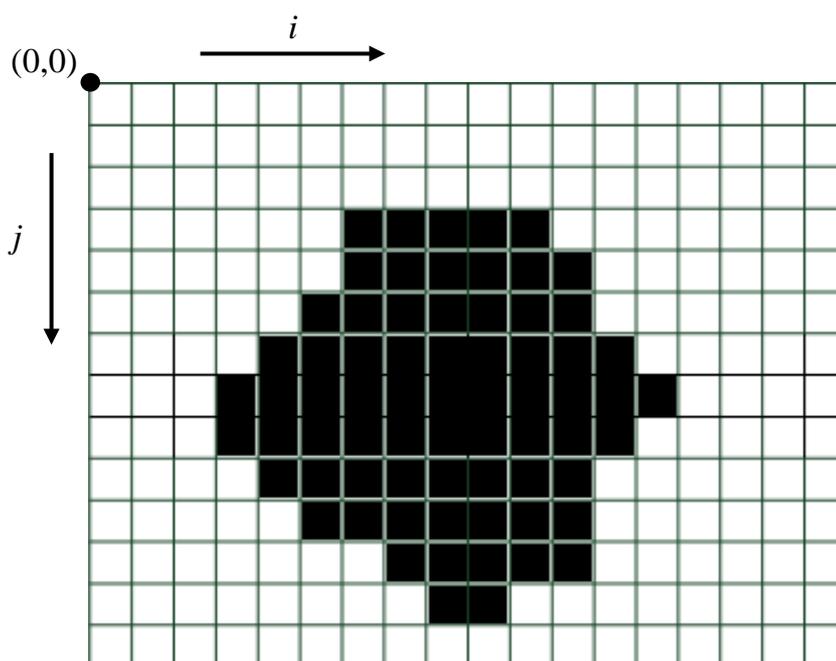


圖 2-1：二維影像座標

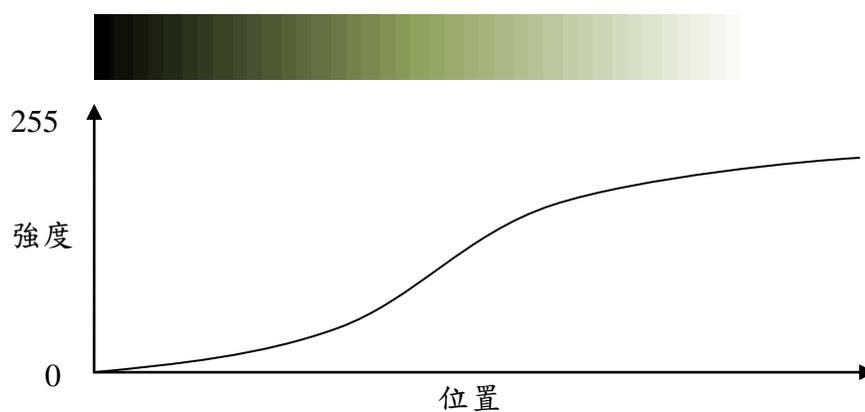


圖 2-2：灰階分佈

$$f(x, y) = \begin{bmatrix} f(0,0) & f(0,1) & \cdots & f(0, N-1) \\ f(1,0) & f(1,1) & \cdots & f(1, N-1) \\ \vdots & \vdots & \cdots & \vdots \\ f(M-1,0) & f(M-1,1) & \cdots & f(M-1, N-1) \end{bmatrix} \quad (2-1)$$

2.2 鄰域處理

鄰域處理的技術普遍使用於影像處理，首先須定義出遮罩(mask)的形狀以及中心點如圖 2-3 所示，遮罩內中心點周遭之區域即為鄰域，將遮罩所包括到區域內的像素值分別帶入使用者的函式中進行計算產生響應 R 如圖 2-4 以及式 2-2 所示，移動遮罩的中心點至下一點產生新的鄰域帶入函式進行計算，重複以上步驟直到整張影像均經過處理，上述步驟中遮罩與函式的結合此運算稱為濾波處理。

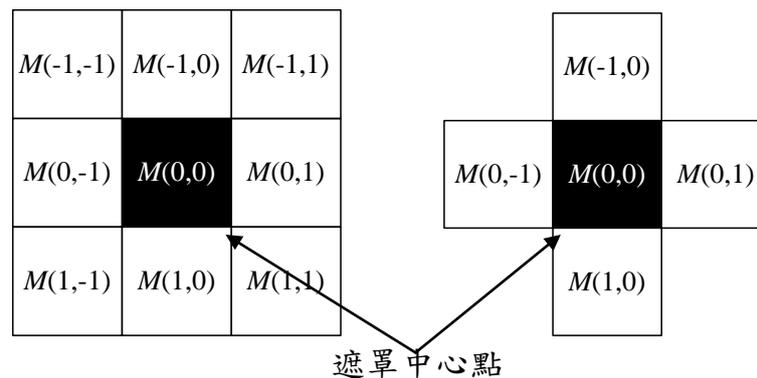


圖 2-3：遮罩示意圖

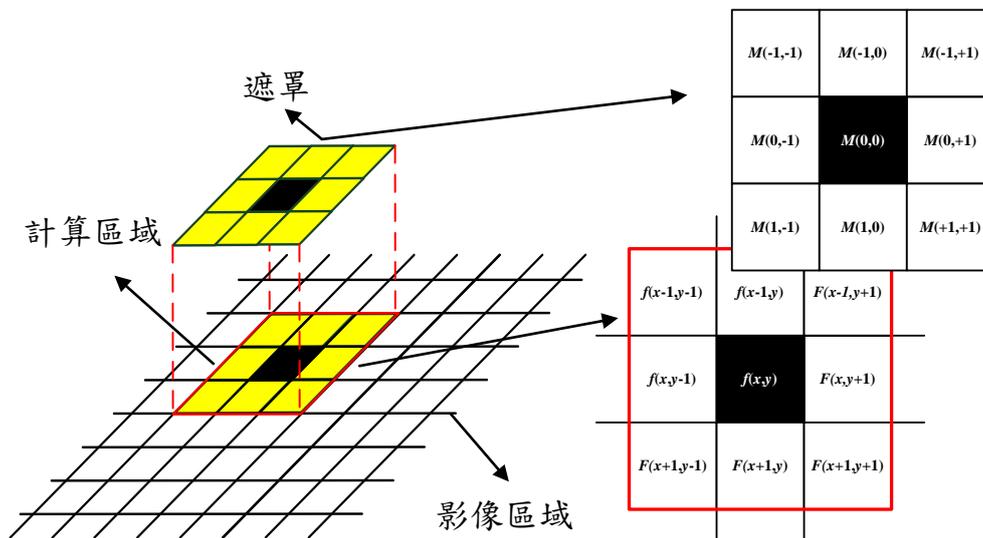


圖 2-4：鄰域處理示意圖

$$R = M(-1,-1)f(x-1, y-1) + M(-1,0)f(x-1, y) + \dots \\ + M(0,0)f(x, y) + \dots + M(1,0)f(x+1, y) + M(1,1)f(x+1, y+1) \quad (2-2)$$

遮罩的大小以及其形狀可依使用者自行設計，不同的大小形狀的遮罩會有不同的處理效果，因此使用者可根據需求設計恰當的遮罩，常見的遮罩為 $n \times n$ 的形式，其中又令 n 為奇數最為常見，當 n 為奇數時遮罩會以中心成對稱，許多濾波器均使用 3×3 的遮罩，但也有其他種類遮罩的應用，例如 Dithering 技術使用 n 為偶數的 $n \times n$ 遮罩。

2.3 影像邊緣偵測(Edge detection)

邊緣偵測是影像處理中常用到的一種技術，邊緣偵測的主要目的為找出影像中圖形的邊緣，我們可以使用前述的鄰域處理來進行簡單的邊緣偵測，使用 3×3 的遮罩與函式產生邊緣濾波器如圖 2-5 所示，若是濾波器計算的區域為平滑區域則計算響應 R 值並不明顯，當濾波器處理到邊緣區域時則 R 值會有明顯的響應，由遮罩內不同的函式我們可以設計水平、垂直或是 45° 的邊緣如圖 2-5(a)(b)(c) 所示，圖 2-5(a) 為水平濾波器，當水平邊緣線通過此濾波器的中心時會有最佳的響應，同理圖 2-5(b) 為垂直濾波器，圖 2-5(c) 為 45° 濾波器。

-1	-1	-1	-1	2	-1	-1	-1	2
2	2	2	-1	2	-1	-1	2	-1
-1	-1	-1	-1	2	-1	2	-1	-1
	(a)		(b)		(c)			

圖 2-5：(a)水平濾波器；(b)垂直濾波器；(c) 45° 濾波器

則從影像強度值的分布情況來看，圖形的邊緣為強度值劇烈改變之處，因此我們可以使用一階導數和二階導數來進行邊緣的偵測如圖 2-6 所示。

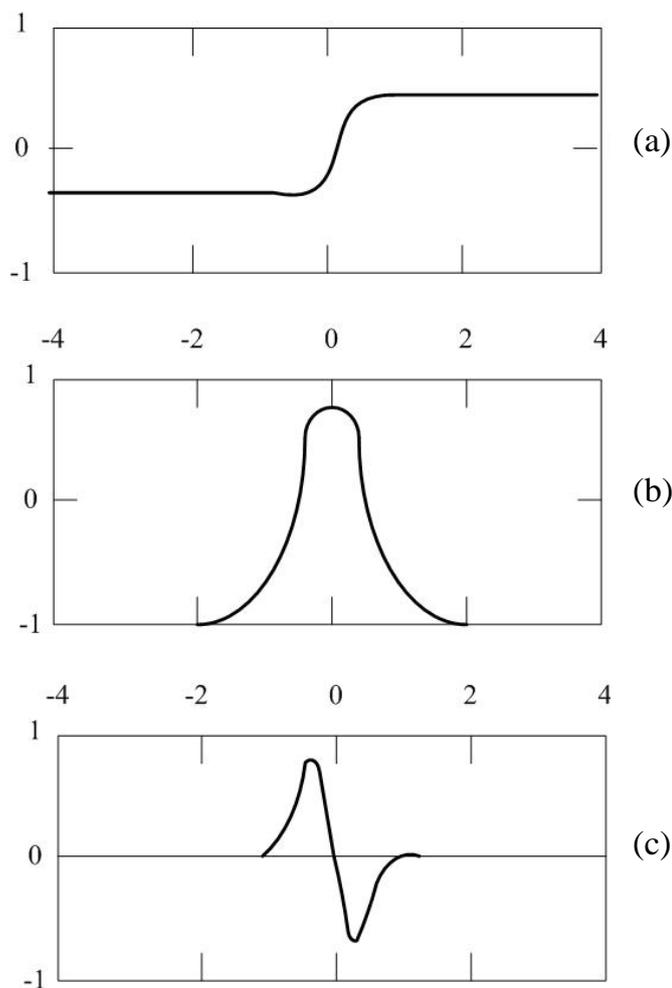


圖 2-6：(a)影像邊緣；(b)一階導數；(c)二階導數示意圖

將二維函數 $f(x,y)$ 的梯度定義為向量如 2-3 所示，梯度大小如 2-4 所示，梯度，梯度向量為 f 在 (x,y) 處之最大改變率發生的角度方向如式 2-5 所示。

$$\nabla f = [G_x \ G_y] = \left[\frac{\partial f}{\partial x} \ \frac{\partial f}{\partial y} \right] \quad (2-3)$$

$$\text{Magnitude } (\nabla f) = \sqrt{G_x^2 + G_y^2} \quad (2-4)$$

$$Direction (\nabla f) = \tan^{-1}\left(\frac{G_y}{G_x}\right) \quad (2-5)$$

2.3.1 Canny 邊緣偵測[19]

Canny 偵測法為 Canny 於 1986 提出的邊緣偵測方法，Canny 邊緣偵測法步驟如下

1. 首先使用高斯濾波器平滑影像降低雜訊的干擾如式 2-6 所示， I 為輸入之原始影像， G 為高斯函數如式 2-7 所示， σ 為特定標準差，將原始影像 I 與高斯函數 $G(x,y)$ 作旋積得到平滑影像 I_g 。

$$I_g = I * G \quad (2-6)$$

$$G = e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (2-7)$$

2. 每點都計算出梯度向量 ∇f ，以及梯度向量之大小與方向如前述所示，若是此點在此梯度方向上為局部最大強度的點，則被定義為邊緣點。
3. 將前一步驟所產生的邊緣點在梯度大小影像上產生山脊，追蹤這些山脊將山脊以外之像素皆令為 0，接著使用者設定兩門檻值 $T1$ 與 $T2$ ，其中 $T1 < T2$ ，山脊上的像素若是大於 $T2$ 則歸類為強邊緣像素，若是介於 $T1$ 與 $T2$ 之間則歸類為弱邊緣像素。
4. 最後將強邊緣像素周遭之弱邊緣像素併入完成邊緣的連接。

2.4 影像二值化

一般灰階畫影像其灰階值的範圍介於 0-255 之間，二值化的影像則為 0 與 1 所組成的邏輯陣列，若是要將灰階影像轉成二值影像則需透過門檻值的建立來進行轉換，圖 2-7 為影像直方圖，影像直方圖為各灰階強度其

像素的數目，如圖 2-7 所示此影像主要有兩個群集，欲區分此兩群集我們可透過兩波峰之間的波谷來決定門檻值 T ，式 2-8 藉由門檻值我們將此影像二值化。

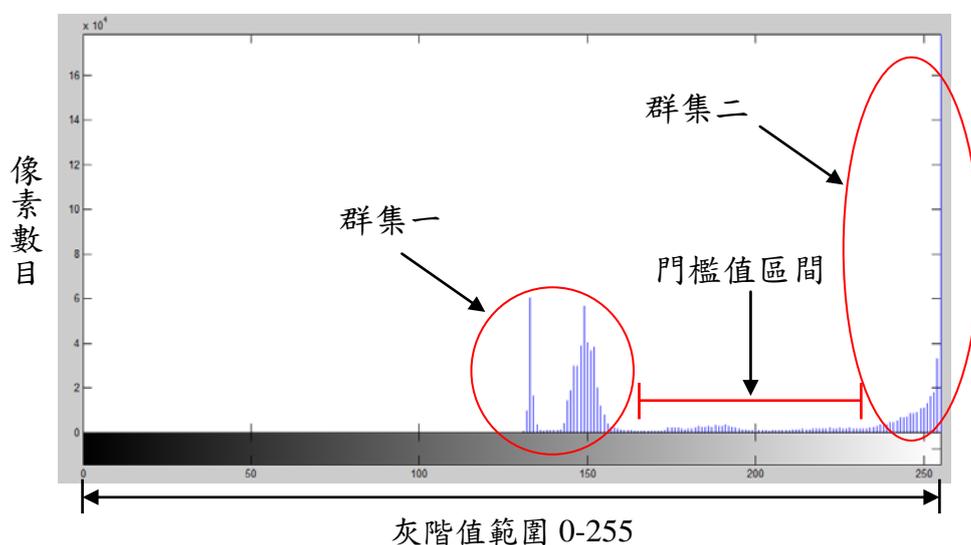


圖 2-7：影像直方圖

$$I = \begin{cases} 1 & \text{if } I \geq T \\ 0 & \text{if } I < T \end{cases} \quad (2-8)$$

2.5 形態學影像處理(Morphology)

形態學影像處理主要透過遮罩以及數學上的集合理論對影像作型態上的處理例如膨脹或是侵蝕，形態學的基礎為集合理論因此首先須了解集合理論， A 為一集合 B 為另一集合如圖 2-8(a)所示，(b) A 與 B 集合之聯集，(c) A 與 B 集合之交集，(d) A 集合與 B 集合之差集，(e) A 集合之補集。

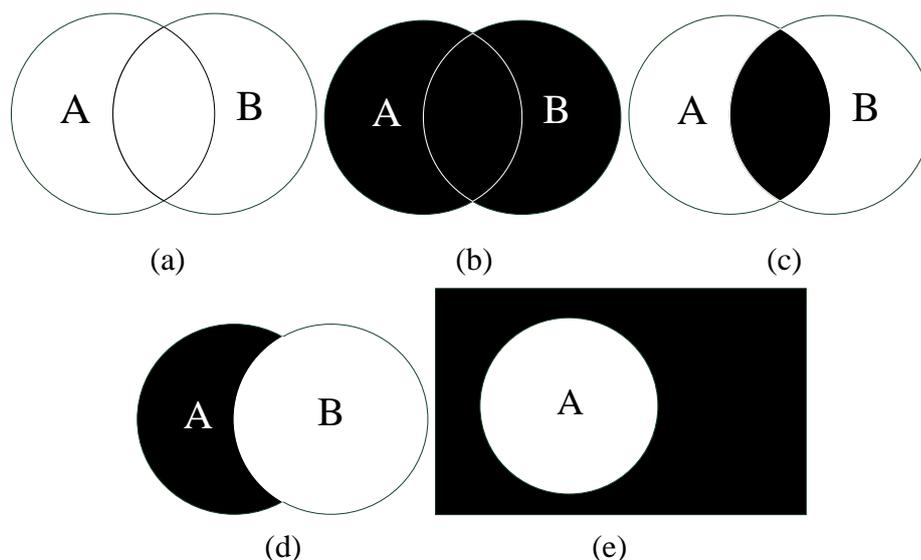


圖 2-8：集合理論(a)A 集合與 B 集合；(b) $A \cup B$ ；(c) $A \cap B$ ；(d) $A - B$ ；(e) $(A)^C$

2.5.1 膨脹(Dilation)與侵蝕(Erosion)

透過遮罩與二值化之影像作集合處理可將影像中之物件產生膨脹或是侵蝕的效果，首先需定義遮罩之型式，使用遮罩對二值影像作鄰域處理，令二值影像內像素值為 1 之集合為 A 集合，遮罩之區域為 B 集合如圖 2-9 所示， $A \oplus B$ 表示 A 藉由 B 產生膨脹其數學式子如式 2-9 所示，當遮罩中心點偵測到影像中值為 1 時將遮罩與影像產生聯集，膨脹過程如圖 2-10 所示，透過膨脹處理我們可以將二值影像中的物件增大、變厚或是增強。

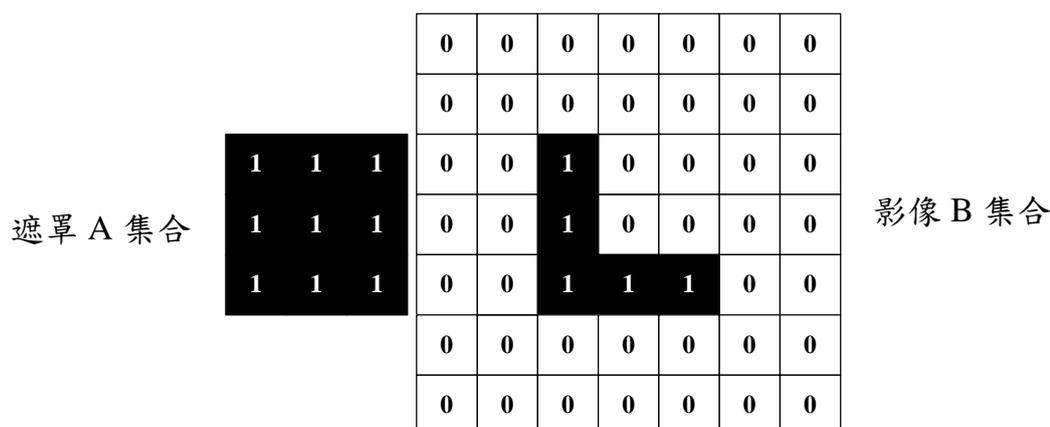


圖 2-9：遮罩 A 集合與二值影像 B 集合

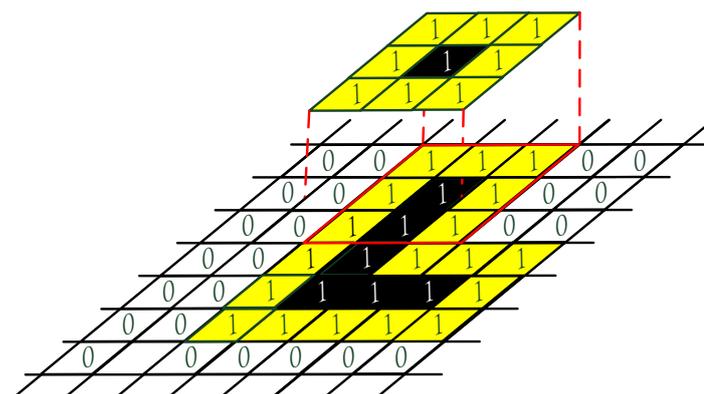


圖 2-10：使用遮罩對影像進行膨脹示意圖

$$A \oplus B = \{x+w/x \in A, w \in B\} \tag{2-9}$$

同樣的透過遮罩與集合理論我們可以使用遮罩對二值影像進行侵蝕處理，使二值影像中的物件收縮或是變薄， $A \ominus B$ 表示 A 藉由進行侵蝕其數學式子如式 2-10 所示，其侵蝕原理為使用遮罩對二值影像逐步進行比對，當偵測到影像局部範圍完全符合遮罩時則令中心點為 1 其他區域為 0，藉此達到侵蝕的效果，其侵蝕過程如圖 2-12 所示。

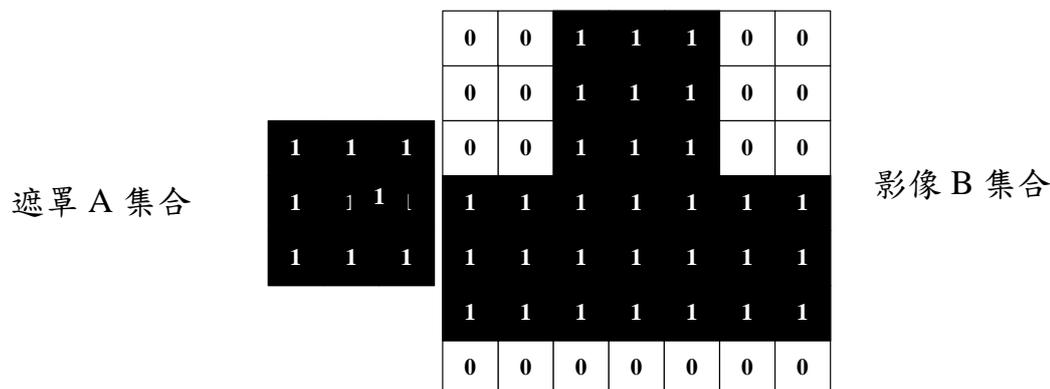


圖 2-11：遮罩 A 集合與二值影像 B 集合

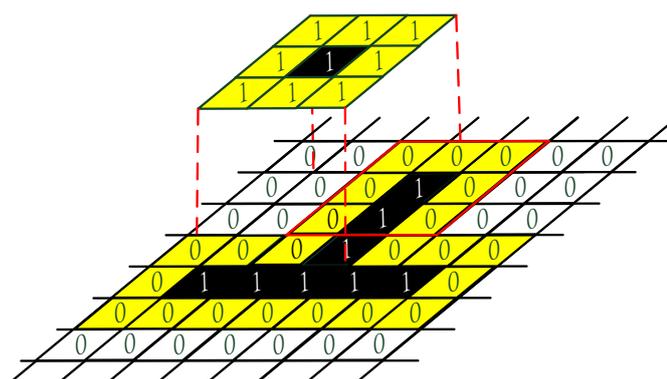


圖 2-12：使用遮罩對影像進行侵蝕示意圖

$$A \ominus B = \{w \in A / B_w \subseteq A\} \quad (2-10)$$

形態學中的膨脹與侵蝕為基礎技術，同一張二值影像可藉由膨脹或是侵蝕的組合達到特殊的效果，例如先將二值影像進行侵蝕將細微雜訊去除或是將細部去除，之後再使用膨脹將影像強化還原達到將原始二值影像斷開的效果，斷開的表示式為 $A \circ B$ 如式 2-11 所示，或是先將二值影像先進行膨脹處理後再進行侵蝕處理，可將不連續處連接起來達到閉合的效果，閉合的表示式為 $A \bullet B$ 如式 2-12 所示，除此之外透過不同的膨脹與侵蝕的組合以及不同的遮罩形狀可以達到特殊的效果，例如雜訊的濾除、邊緣偵測、細線化、骨架化或是影像增強等效果。

$$A \circ B = (A \ominus B) \oplus B \quad (2-11)$$

$$A \bullet B = (A \oplus B) \ominus B \quad (2-12)$$

2.6 轉角偵測(Corner detection)

轉角偵測是常見的擷取物件特徵點的方法，轉角偵測廣泛的被應用於許多領域，物件的特徵點常出現於邊緣或是轉角，因此透過轉角偵測我們可以擷取出物件的特徵訊息，轉角偵測可透過遮罩的鄰域處理或是使用梯

度演算法找出物件輪廓的轉角點。

2.6.1 Harris 轉角偵測[20]

Harris轉角偵測演算法首先於影像中設定一區域遮罩，遮罩會沿各方向作微小移動，接著計算遮罩內平均像素強度的變化，當強度的變化值大於預先設定的門檻值時將遮罩中心點標記為轉角點，令像素 (x,y) 的強度為 $f(x,y)$ ， $w(x,y)$ 為高斯係數對影像作平滑處理，遮罩移動 (u,v) 時其強度變化以式 2-13 所示 $Q(x,y)$ 如式 2-14 所示， I_x, I_y 為像素水平與垂直導數，透過式 2-15 計算轉角點響應偵測出物件之轉角點。

$$\begin{aligned} E(x, y) &= \sum_{u,v} w(x, y) (I(x+u, y+v) - f(x, y))^2 \\ &\approx \sum_{u,v} w(x, y) \left([I_x, I_y] \begin{bmatrix} u \\ v \end{bmatrix} \right)^2 \\ &= [u, v] Q(x, y) \begin{bmatrix} u \\ v \end{bmatrix} \end{aligned} \quad (2-13)$$

$$Q(x, y) = \sum_{u,v} w(u, v) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} \quad (2-14)$$

$$R = \det(Q) - 0.04 \operatorname{tr}(Q)^2 \quad (2-15)$$

2.7 透視投影法

一般最常見的相機是透過針孔成像將實景投影至投影面上，針孔成像模型如圖 2-13 所示，透視投影法較接近一般視覺效果，其投影線會集中於投影中心點如同人類視網膜成像原理，如圖 2-14 所式，其中 f 為相機投影之焦距，透過透視投影法，我們可以求出現實空間投影至二維影像上之投影關係矩陣，而這投影關係矩陣是由相機內部參數所構成，假設 X_w 為現實

空間中之一點，而 X_m 為其投影至二維影像上之投影點， X_w 與 X_m 之間的關係矩陣可由式 2-16 表示， S 為比例常數。

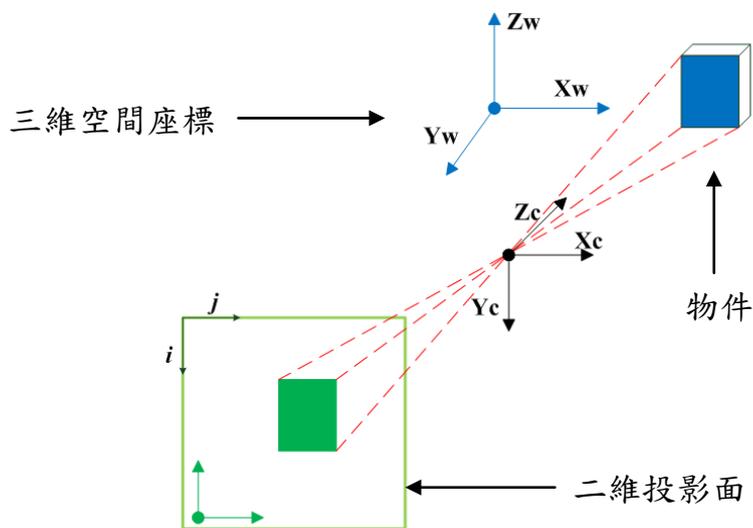


圖 2-13：針孔成像示意圖

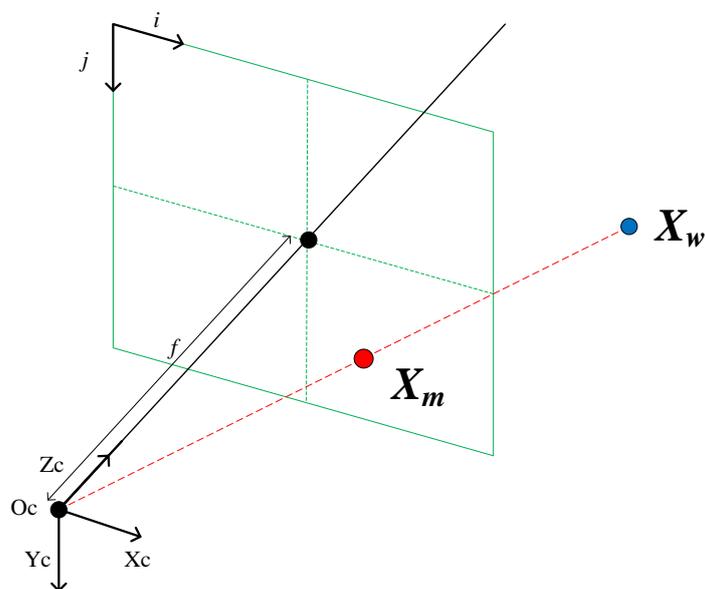


圖 2-14：透視投影法

$$SX_m = \begin{bmatrix} f_x & \gamma & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} X_w \quad (2-16)$$

相機參數中 f_x 、 f_y 為相機水平與垂直焦距， γ 為相機變形參數， u_0 、 v_0 為影像中心點。

2.8 立體視覺法(Stereo vision)

立體視覺法其原理為模擬人類的雙眼來產生深度資訊，假設現實空間裡有一物體分別透過兩台攝影機分別投影至二維的影像座標上如圖 2-15 所式，現實空間中 X 為物體表面上之一點，分別透過攝影機 Camera1 與 Camera2 投影至二維的空間座標並且產生二維影像， x 與 x' 為其投影點，而 c_1 為 Camera1 之投影中心點， c_2 為 Camera2 之投影中心點，而 X 、 x 、 x' 、 c_1 以及 c_2 所位於的平面稱為極線平面(epipolar plane)，而 c_1 與 c_2 的連線稱為 baseline，位於 baseline 上的 e 與 e' 則為稱為極點(epipolar pole)，空間中任一點均會與兩個投影面形成不同的極線平面，但是不變的是 baseline 與極點，而極線平面與投影面的交線 l_1 與 l_2 稱為極線(epipolar line)，極線平面上任一點投影至上 l_1 其相對應的點也只會投影至 l_2 上，如此一來可以縮減相對應投影點的搜尋範圍。

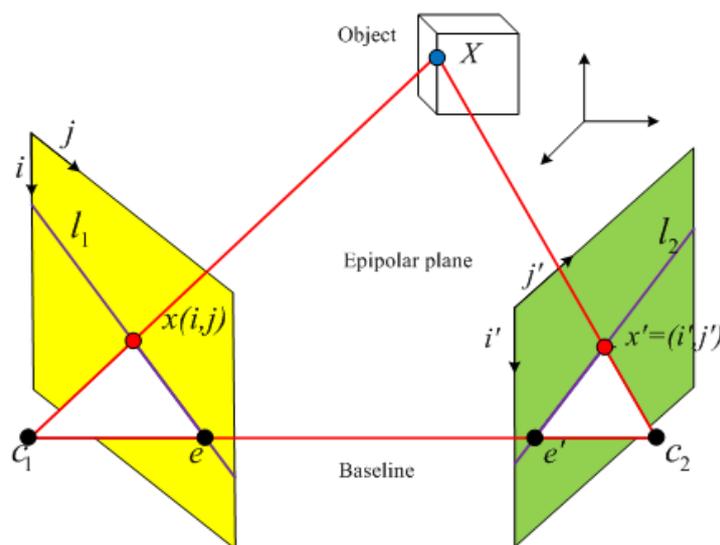


圖 2-15：Epipolar geometry

2.9 Fundamental matrix

欲求得深度資訊則需先找出現實空間中點 X 投影至兩平面的投影點 x 與 x' ，而 x 與 x' 兩投影點間存在一關係矩陣而此關係矩陣稱為 Fundamental matrix，藉由 Fundamental matrix 即可從投影面上的投影點 x 找出另一投影面上相對應的投影點 x' ，如圖 2-16 所示現實空間中的點 X 與投影點 x 、 x' 之間的向量分別為 \vec{v}_1 、 \vec{v}_2 ，而 c_1 與 c_2 之間的向量為 \vec{t}_1 ， \vec{n} 為垂直 epipolar plane 之向量， \vec{v}_1 、 \vec{v}_2 之間的關係是可由式 2-17 表示， R 為旋轉矩陣。 $\vec{v}_1 - \vec{t}_1$ 為 epipolar plane 上之向量因此我們可以藉由外積定理列出式 2-18，將式 2-17 代入式 2-18 得到式 2-19，令 $\vec{n} = S\vec{v}_1$ 代入式 2-19 得到式 2-20，令 $(R^{-1})^T S = E$ 則得到式 2-21，矩陣 E 為 \vec{v}_1 、 \vec{v}_2 的關係矩陣。

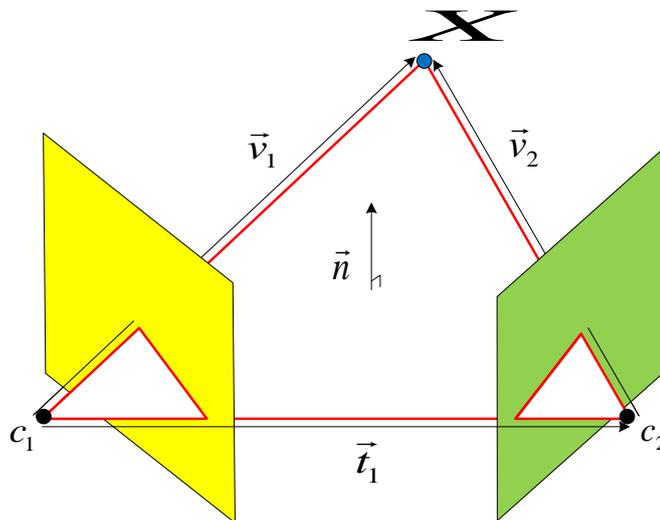


圖 2-16：Epipolar geometry

$$\vec{v}_2 = R(\vec{v}_1 - \vec{t}_1) \quad (2-17)$$

$$(\vec{v}_1 - \vec{t}_1)^T \vec{n} = 0 \quad (2-18)$$

$$(R^{-1}\vec{v}_2)^T \vec{n} = 0 \quad (2-19)$$

$$(R^{-1}\vec{v}_2)^T S\vec{v}_1 = 0 \quad (2-20)$$

$$\vec{v}_2^T E\vec{v}_1 = 0 \quad (2-21)$$

M_1 、 M_2 為攝影機的內部參數、則向量與投影點的關係式如 2-22 與 2-23 式所示，將上述兩式整理代入式 2-21 中可得到 2-24 式，整理 2-24 式令 $(M_2^{-1})^T E(M_1^{-1}) = F$ 則得到 2-25 式、 F 即為兩投影點間之關係矩陣 Fundamental matrix，藉由 Fundamental matrix 即可從投影點 x 找出另一相對應的投影點 x' 。Fundamental matrix 為奇異矩陣且 rank 數為 2。

$$\vec{v}_1 = M_1^{-1}x \quad (2-22)$$

$$\vec{v}_2 = M_2^{-1}x' \quad (2-23)$$

$$(M_2^{-1}x')^T E(M_1^{-1}x) = 0 \quad (2-24)$$

$$x'^T Fx = 0 \quad (2-25)$$

矩陣 F 為 3×3 矩陣，只要兩張影像上任一組相對應的點 x 、 x' 均滿足 $x'^T Fx = 0$ ，因此在未知攝影機內部參數的情形下，只是要先建立起 Fundamental matrix 即可在另一張影像上找出相對應的點，如此一來藉由 x 投影至兩張影像上的兩個投影點就可計算出 X 的深度資訊。

假設已知兩個分別位於左右影像的投影點 $x=(i,j,1)^T$ 、 $x'=(i',j',1)^T$ ， F 以式 2-26 表示， i 、 j 為投影點於左影像平面上之座標，單位為 pixel， i' 、 j' 為投影點於右影像平面上之座標如圖 2-15 所示，其中 F 為 3×3 矩陣，故

將兩投影點的座標矩陣擴充為 3×1 矩陣，將兩投影點代入 2-25 式可得 2-27 式、將 2-27 式以矩陣表示如 2-28、2-29 與 2-30 式，欲求出 F 裡的所有參數我們可以事先將八個已知的點分別投影至左右平面，如此一來即可得到八組相對應的投影點，如 2-28 式所示一組投影點可以得到 1×9 矩陣 a ，若有八組投影點則可得到 8×9 矩陣 $A=(a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8)$ ，而矩陣 f 為 9×1 矩陣，最後只要求解 2-31 式即可求得 Fundamental matrix F 中的各個係數。

$$F = \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix} \quad (2-26)$$

$$ii'f_{11} + jj'f_{12} + i'f_{13} + ij'f_{21} + jj'f_{22} + j'f_{23} + if_{31} + jf_{32} + f_{33} = 0 \quad (2-27)$$

$$a = (ii', jj', i', ij', jj', j', i, j, 1) \quad (2-28)$$

$$f = (f_{11}, f_{12}, f_{13}, f_{21}, f_{22}, f_{23}, f_{31}, f_{32}, f_{33})^T \quad (2-29)$$

$$af = 0 \quad (2-30)$$

$$Af = 0 \quad (2-31)$$

由於代入的已知的相對應投影點有誤差因此導致求出 F 的 rank 數非 2，或是無法順利求解出 F 矩陣內所有之參數，所以我們找出其相近的 F' 矩陣來代替 F 矩陣，並且使 $x'^T F' x$ 有最小值趨近於 0。

第三章 系統設計概念與配置

本章節依據第二章所引述之理論基礎，依本研究之目標架設所需之系統架構，其中包含有三維重建系統架構、三維重建系統設計概念分析…等。

3.1 三維重建系統設計實現目標

本論文之研究目的為重建物件之三維模型，根據第二章之影像處理、投影原理以及立體視覺法等基礎理論，配合 CCD 擷取系統達成三維自動重建之要求。本研究預計達成下列各項條件

1. 快速重建三維模型
2. 重建系統架構簡單
3. 無需使用額外設備
4. 整體設備成本低廉

3.2 三維重建系統架構

本研究三維重建系統如圖 3-1 所示。主要重建環境的設備包含有兩隻 CCD 攝影機、L 型壓克力與光學板，可調式支架，使用個人電腦進行影像處理以及重建程式的運算，以這些結構發展出三維重建系統，重建目標物件的三維模型。

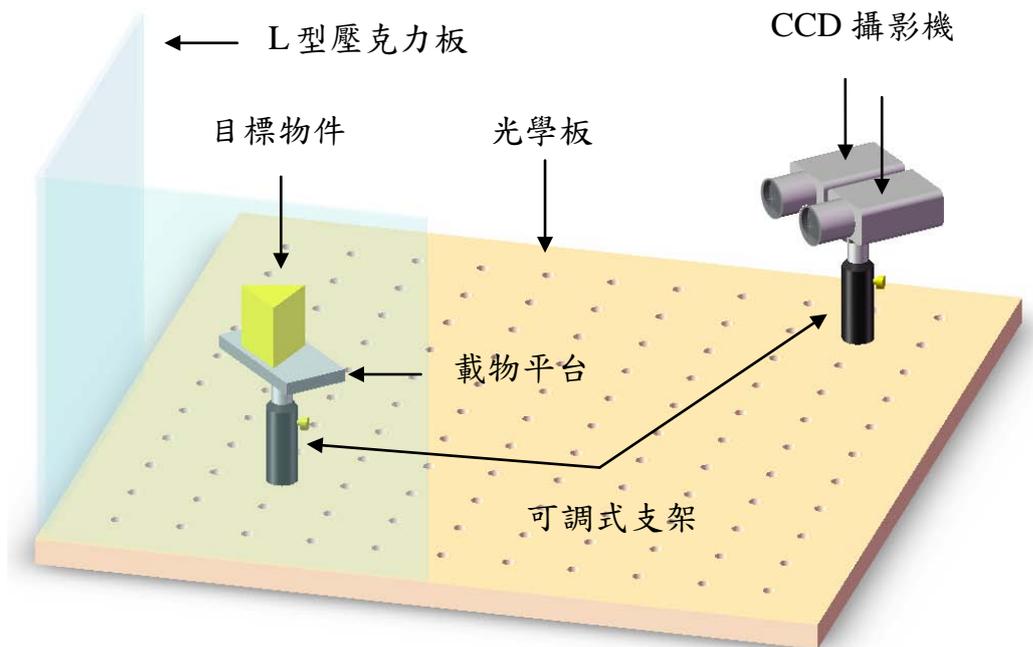


圖 3-1：三維重建系統配置圖

3.3 三維重建系統架設概念

在第二章的投影原理中，透過透式投影的方式，可將空間中物體的影像透過鏡頭投影至投影面，利用立體視覺法根據左右兩 CCD 攝影機所擷取的影像，透過透式投影以及相機參數我們可以建立起現實空間與二維投影面的轉換矩陣，將投影面轉換至重建之三維空間中，以左 CCD 攝影機之投影中心點為重建原點，以左 CCD 攝影機之投影面為基準於右 CCD 攝影機之投影面搜尋相對應投影點，找出所有相對應投影點與兩 CCD 攝影機之投影中心點及位置關係，利用極線幾何原理於三維座標空間重建起物件之三維模型，並且計算出物件模型與 CCD 攝影機之間的距離。

3.4 三維重建系統配置

A. 相機配置

根據參考文獻[19]中關於相機配置與重建模型的實驗結果顯示，兩相

機之間平行設置其誤差為最低會有最佳的重建結果，故在本研究中我們將兩台 CCD 攝影機設為相互平行，且人類雙眼瞳孔間距平均約為 6.5cm，因此參考人類雙眼瞳孔間距以及考慮到攝影機機身之大小，我們設置攝影機中心間距約為 5.5cm，為有利於從左右兩影像中找出相對應之投影點，我們將兩台 CCD 攝影機設置為等高，如此一來即可透過水平搜尋極線上的點定位出相對應的投影點，欲達到此目的我們將兩 CCD 鎖至同一平台上希望從硬體設置降低其重建誤差如圖 3-2 所式。

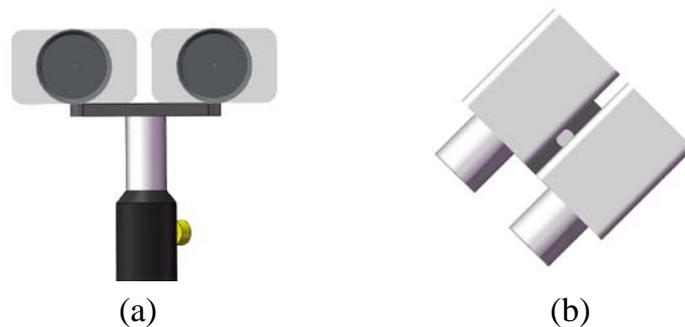


圖 3-2：相機配置圖(a)相機等高設置；(b)相機相互平行

B. 偏光鏡

本系統未對光源進行限制，實驗環境為一般室內光源，事實上光源的種類以及光源的方向均會影響 CCD 攝影機擷取影像的品質，物件的材質以及表面情況會產生程度不一的反光，反光會增加之後影像處理的困難度，反光會造成物件表面的失真，進行邊緣偵測也會產生不必要的輪廓，欲消除物件表面的反光可從硬體或是軟體著手，硬體方面可從光源進行改善，利用背光模組或是調整光源的方向均可降低反光，但是為加速重建的速度，本實驗盡量降低電腦的運算量，因此不使用需龐大計算的影像處理進行反光的消除，所以本實驗於鏡頭前端裝偏光鏡，光線經平滑表面的反射後，部分反射的光線會產生偏振光，偏振光會造成影像上的反光，加裝偏光鏡的主要目的就是將這些破壞影像畫面的偏振光去除，進而得到良好影像的效果，其原理是利用偏光鏡中的柵狀薄膜，將反射的偏振光擋在外面，

只容許與柵狀結構相同方向的光線通過達到濾除反光的效果，其效果如圖 3-3 所示。

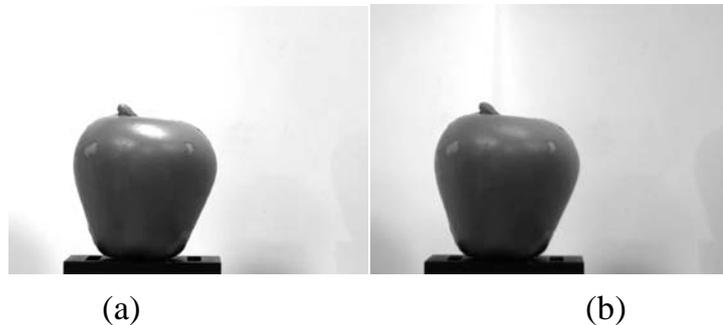


圖 3-3：偏光鏡(a)反光造成的干擾；(b)加裝偏光鏡濾除反光

C. CCD 攝影機與待測物之間之校正

本實驗所使用之 CCD 攝影機最高解析度為 1200×1600 ，鏡頭的焦距約為 60mm，其中 CCD 攝影機解析度的調整會影響畫面的大小進而影響物件擺放的自由度，若是將兩攝影機間距的中心點對準物件中心，物件會在左攝影機之畫面偏右之位置，物件會在右攝影機之畫面偏左之位置，物件擺放位置的第一要點左右兩攝影機均能擷取到完整的物件影像，因此高解析度能較自由的擺放物件，而鏡頭焦距的調整會影響物件擺放的距離，鏡頭光圈的調整會影響鏡頭的透光率，本實驗將 CCD 攝影機之解析度調整為最高目的是希望能有較高物件的擺放自由度，將物件擺放距離兩 CCD 攝影機約 840mm 之處並且調整鏡頭焦距至能清楚擷取物件影像，光圈調整至能清楚的擷取物件輪廓又將反光之干擾降至最低，CCD 攝影機與待測物之設置如圖 3-4 所示，左右兩 CCD 攝影機擷取之影像如圖 3-5 所示。

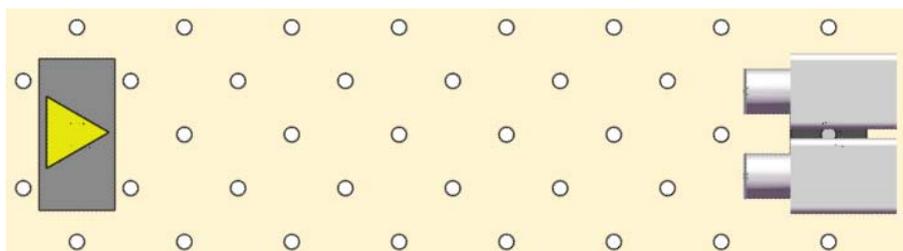


圖 3-4：兩 CCD 攝影機與待測物之設置

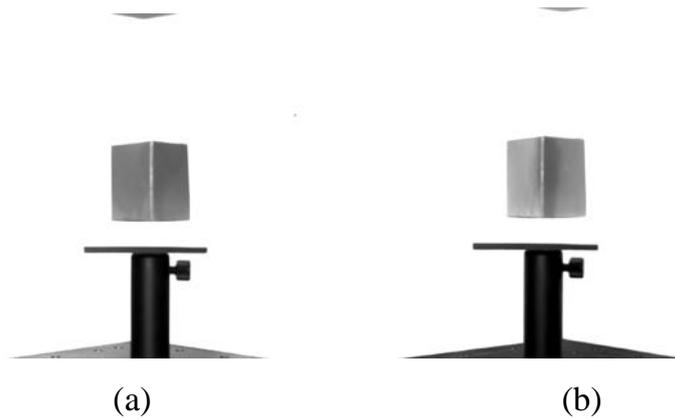


圖 3-5：左右兩 CCD 攝影機擷取之影像(a)左影像；(b)右影像

3.5 本重建系統之特色

本研究主要目的為重建物件之三維模型，其次要能降低電腦之運算量加速重建之速度達到即時的重建目的，因此本系統針對簡單之幾何物件以及較難重建之曲面物件分別提出不同之方法進行迅速的重建，針對簡單之幾何物件其特徵點會出現於物件輪廓之頂點以及邊緣上，因此本系統透過邊緣偵測找出物件之特徵點進行三維模型的重建，曲面物件難以重建的原因即是曲面上並無明顯之特徵點，故本研究提出建構虛擬網格的方法來重建物件之曲面，藉由掃描物件之輪廓複雜度以不同解析度之網格切割左影像物件之表面輪廓，接著利用水平搜尋以及極線定位出左影像網格點於右影像之相對應投影點，利用立體視覺法重建出曲面之三維模型，藉由最低數量之重建點達到迅速重建物件三維模型之目的。

3.6 重建系統流程描述

本研究三維重建系統整體流程如圖 3-6 所示，共分為六階段，如下所述。

第一階段(相機校正)：使用相機校正程式找出相機之參數例如焦距，投影中心點，變形參數等。

第二階段(計算 Fundamental matrix)：帶入大量已知相對應之投影點計算 Fundamental matrix。

第三階段(影像擷取)：透過左右兩 CCD 攝影機擷取物件影像並傳至電腦進行處理。

第四階段(影像處理)：透過各種影像處理例如灰階化，將物件影像從背景提取，邊緣偵測，斷線閉合等，對物件輪廓進行擷取以及強化。

第五階段(找尋 match points)：此階段最本研究之核心，在此我們將物件分為簡單幾何物件以及曲面物件，針對簡單幾何物件透過轉角偵測輔以極限搜尋 match points，對於曲面物件則使用虛擬網格以及極線找尋 match points。

第五階段(重建物件三維模型)：透過立體視覺法，藉由相機參數，空間之轉換，match points 重建出物件之三維模型。

第六階段(旋轉物件)：此階段為非必要，受限於攝影機一次只能擷取一方向之物件影像因此一次重建只能重建出物件一面之三維模型，若是要重建出完整物件三維模型則整合旋轉平台，將物件轉一圈進行拍攝如此一來即可重建出完整物件三維模型，若是欲應用於機械手臂之夾取為講求方便性以及速度，只需重建出物件一面之三維模型即可進行夾取。

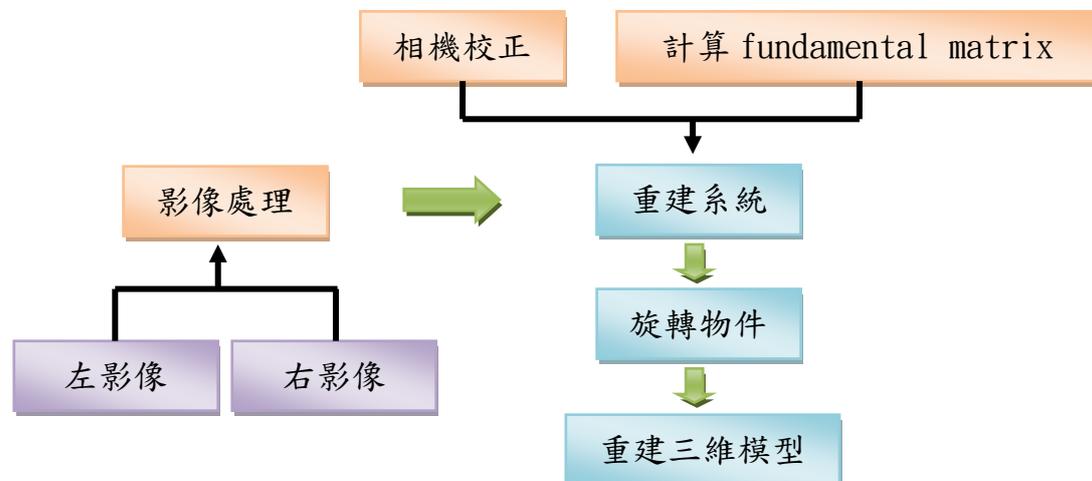


圖 3-6：三維重建系統整體流程

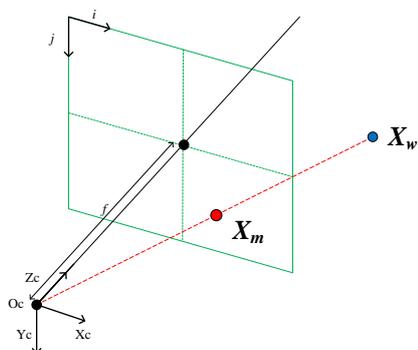
第四章 三維模型重建系統設計原理

本章節依據第二章所引述之理論基礎，依系統流程架構逐一詳述每一階段之實驗目的與方法，其中包含有事前相機校正、Fundamental matrix 的求解、影像前處理、match points 之搜尋、三維重建系統架構、三維重建系統設計概念分析…等。

4.1 系統前處理

A. 相機校正

本實驗主要之實驗為透過 CCD 攝影機擷取物件之二維影像，透過二維影像重建物件之三維模型，因此相機校正準確與否會影響重建模型之準確度，相機校正的主要目的為求出其中的相機參數，CCD 攝影機成像原理為針孔成像原理，我們可透過透視投影法表現時空間一點與其投影點之間的投影關係，而其中的投影關係可由一關係矩陣所表示，此關係矩陣內之參數主要由相機參數所構成如式 4-1，其物理意義為現實空間之一點透過 CCD 攝影機投影至二維投影面上產生投影點，投影原理已於第二張詳細描述，故在進行重建以前必須先透過相機校正求得各相機參數，例如水平、垂直焦距、投影中心點，投影面中心點以及變形參數等以供之後三維重建之使用，透視投影以及相機參數如圖 4-1 所式。



$$SX_m = \begin{bmatrix} f_x & \gamma & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} X_w \quad (4-1)$$

圖 4-1：透視投影法

B. 求出 Fundamental matrix

根據 epipolar geometry 理論，現實空間中一點透過左右兩 CCD 攝影機產生左右兩投影點，而兩左右投影點之間存在著平移以及旋轉幾何關係，而左右投影點之間的幾何關係我們可由 Fundamental matrix 進行表示，因此透過 Fundamental matrix，我們可將 match point 的搜尋範圍由一整個投影面縮小至一直線上，其原理已於第二張詳細描述，為求出此關係矩陣 Fundamental matrix 我們帶入大量已知之相對應投影點進行求解，之後搭配邊緣偵測或是網格切割以及水平搜尋進一步定位出 match point 之位置。

4.2 影像前處理

A. 灰階化

本實驗之重建系統主要目的是依據物件輪廓重建出物件之三維模型，故在影像前處理之部分首重擷取物件輪廓，一般彩色影像之資料量為灰階影像的三倍大，為方便之後的影像處理也為降低運算量一開始先將 CCD 攝影機傳入之影像進行灰階化，之後進行後續的影像處理。

B. 將物件從背景提取出來

本實驗重建之目標僅物件之影像，故須先建影像從背景中提取出來，此一方法很久，可利用影像像素之灰階值建立門檻值將前景與背景進行分類，或是使用適應性且較為強健之奧圖演算法進行分類，但本實驗其背景環境簡單且乾淨，故使用運算成本最低的方法，先建立起環境背景影像，在將擺入物件之影像與事先建立的環境影像進行比較，找出兩影像之間差異之部分其差異的部分即為物件之影像，利用此方法將物件影像提取出來。

C. 邊緣偵測(edge detection)

從物件之影像掃描偵測出其輪廓，目前有很多種邊緣偵測演算法，可針對不同之目的以及需求選用適當的邊緣偵測演算法，例如 Sobel、Laplacian、Canny 等邊緣偵測演算法，本實驗選用適應性高且強健之 Canny 演算法進行邊緣偵測，其演算法可依使用者自行調整其門檻值參數達到最佳之邊緣偵測效果。

D. 二值化

二值化的目的主要是方便且易於之後的影像處理，首先設立門檻值，其門檻值建立之主要依據為整張影像中各灰階值數量的分布情形，如圖 4-2 所式，其中一峰群為背景之像素集合，令一峰群為物件之像素集合，因此我們可用兩峰群之間之波谷作為門檻值將物件與背景分為兩區，之後另屬於物件之像素其值為 1 反之則為 0，之後即可從其像素值判斷其像素是否屬於物件。

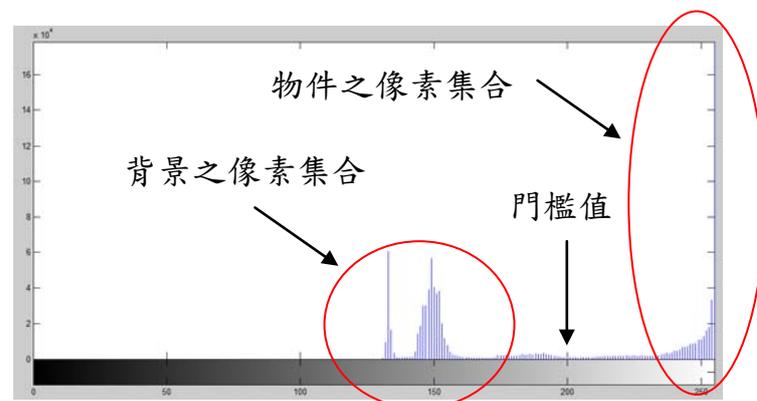


圖 4-2：影像中各灰階值數量的分布情形

D. 型態學影像處理(Morphology)

影像型態學可根據使用者設計之遮罩將物件之影像進行膨脹或是侵蝕等處理，使用影像型態學一方面可以達到濾除雜訊的效果，另一方面也可強化邊緣偵測之效果，藉由膨脹與侵蝕的搭配作用亦可有其他效果，而

本實驗使用影像型態學之主要目的是將前面邊緣偵測所得到之輪廓物件還原成封閉輪廓，邊緣偵測時若是受到雜訊的干擾例如表面光線的散射或是反光等影響，導致輪廓的失真或是未偵測到部分的輪廓，此時我們所得到的輪廓即為開放輪廓，一定程度以下的失真我們可用影像型態學進行還原，先將邊緣影像進行膨脹處理而膨脹時所用的遮罩大小以及形狀使用者可自行設計，膨脹之後斷線之部分即會因線條之膨脹而接合起來，之後再用相同之遮罩進行侵蝕將物件輪廓還原至原始大小，之後我們即可得到封閉區域之物件輪廓。

4.3 搜尋 match points

本研究使用立體視覺法重建物件之三維模型，根據立體視覺法的 epipolar geometry 欲在三維座標中還原一點之三維座標需要兩攝影機之投影中心點以及其投影至左右兩投影面之相對應投影點，攝影機之投影中心點可透過相機校正求得，欲從左投影面於右投影面找出相對應之投影點則是我們接下來的目標也是本研究提出改善之部分，在此本研究將物件初步區分為簡單幾何物件以及曲面物件，我們將提出不同方法針對兩種物件進行 match points 的搜尋。

4.3.1 Epipolar line

第二章已針對 epipolar geometry 詳細介紹過其理論基礎，其中我們可以利用左投影面之投影點以及事先求出之 fundamental matrix 將 match point 的搜尋範圍從一整個右投影面縮小至右投影面上之一直線，如圖 4-3 所示，現實空間一點 X_1 投影至左投影面產生一左投影點 x_1 ，投影至右投影面產生一相對應右投影點 x'_1 ，若是將點 X_1 移至點 X_2 處，則現實空間一點 X_2 投影至左投影面產生一左投影點 x_2 ，投影至右投影面產生一相對應右投影點 x'_2 ，從圖中所示我們可以發現左投影點 x_1 與

x_2 對於左投影面其實為相同一點，但是卻在右投影面上產生不同的兩投影點 x'_1 與 x'_2 ，根據 epipolar geometry 若是 x_1 與 x'_1 為相對應投影點則式子(4-2)成立，其中 F 為事先求出之 fundamental matrix，但是根據圖 4-3 所式若是 x_1 為定值，符合式子 4-2 的 x'_1 為一集合並非只有唯一解，此集合可用直線 l' 進行表示，因此我們可以將式子 4-2 改寫為式子 4-3，其中 l' 稱為右投影面 epipolar line，同理我們也可得到式子 4-4，其中 l 稱為左投影面 epipolar line。

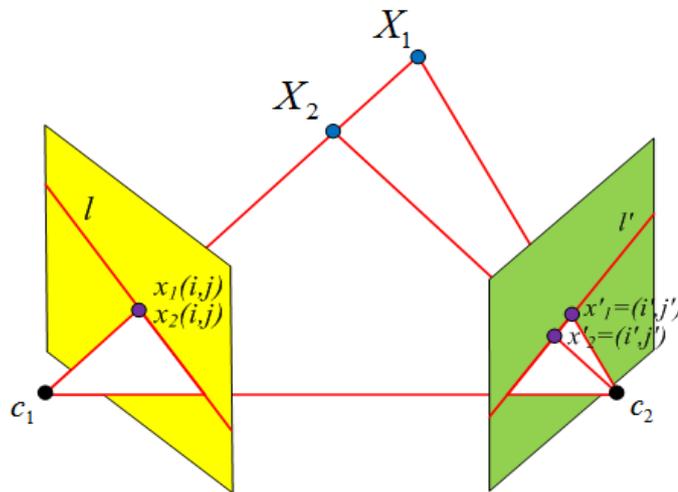


圖 4-3：epipolar geometry

$$x_1'^T F x_1 = 0 \tag{4-2}$$

$$l' = F x_1 \tag{4-3}$$

$$l = F^T x'_1 \tag{4-4}$$

4.3.2 簡單幾何物件搜尋 match point

簡單幾何的特性即是其特徵點均會出現在物件之頂點，如圖 4-4 所示若是可以重建出物件特徵點之三維座標，即可從三維座標空間中之特徵點重建出物件之三維模型，如圖 4-5 所式，因此對於簡單幾何物件的重建，

第一步即是對物件影像進行轉角偵測找出其特徵點之所在位置，由於簡單幾何之頂點特徵強烈，且本實驗設至之左右 CCD 攝影機靠近擷取到的物件影像其實差異並不大，因此對左右兩影像之物件影像進行轉角偵測，會偵測出相同之特徵點。

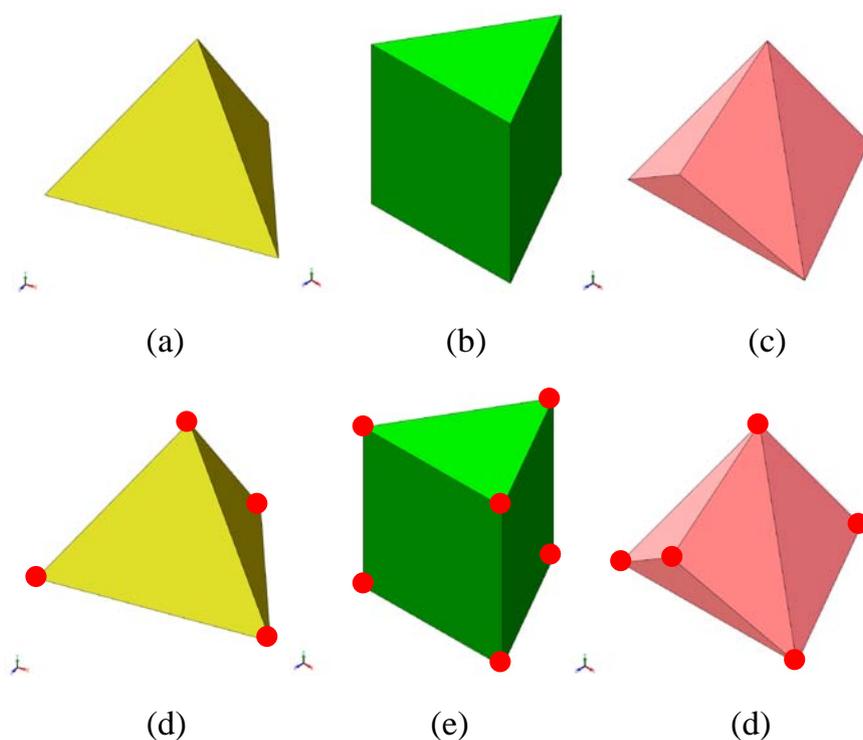


圖 4-4：(a)(b)(c)簡單幾何物件(c)(d)(f) 簡單幾何物件之特徵點

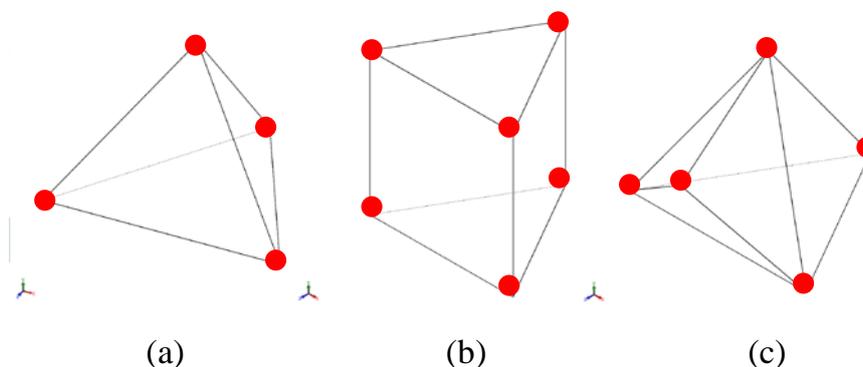


圖 4-5：(a)(b)(c)藉由特徵點重建出物件之三維模型

考慮到物件影像有經過前面階段的影像處理其影像可能會有些微之變異，導致轉角偵測所得到的特徵點偏離正確的投影點，未修正此誤差因此使用 epipolar line 找出最佳之特徵點，藉由 fundamental matrix 以及左投影點 x 根據式子 4-6 我們可於右投影面找出 epipolar line l' ，正確之 match point x' 會出現於此 epipolar line 上 ($x' \in l'$)， X_r 為符合式子 4-5 所有 x' 之集合即 l' 上所有點之集合 ($X_r \in l'$) 式子 4-6、4-7，因此我們根據轉角偵測之轉角點 x_c 為中心擴大偵測的範圍，範圍大小為 $n \times n$ 之遮罩如圖 4-6 所示， $n \times n$ 遮罩與 epipolar line 之交集即為 match point 之位置，將交集之子集合座標做平均設為最佳之 match point 如式子 4-10 所示，將所有特徵點均做以上敘述之處理，即可初步還原出物件部分之三維模型。

$$x'^T Fx = 0 \quad (4-5)$$

$$l' = Fx \quad (4-6)$$

$$X_r = (x'_1, x'_2 \cdots x'_i) \quad i = 1, 2, 3 \dots m \quad X_r \in l' \quad (4-7)$$

$$\text{Mask} = x_{ij} ; i = 1, 2, 3 \dots n ; j = 1, 2, 3 \dots n ; n = 1, 3, 5 \dots \quad (4-8)$$

$$x_{ij} = x_c ; i = j = (n+1)/2 \quad (4-9)$$

$$X' = \text{Mask} \cap X_r ; X' (1 \times S) \quad (4-10)$$

$$x' = \frac{\sum X'}{S} \quad (4-11)$$

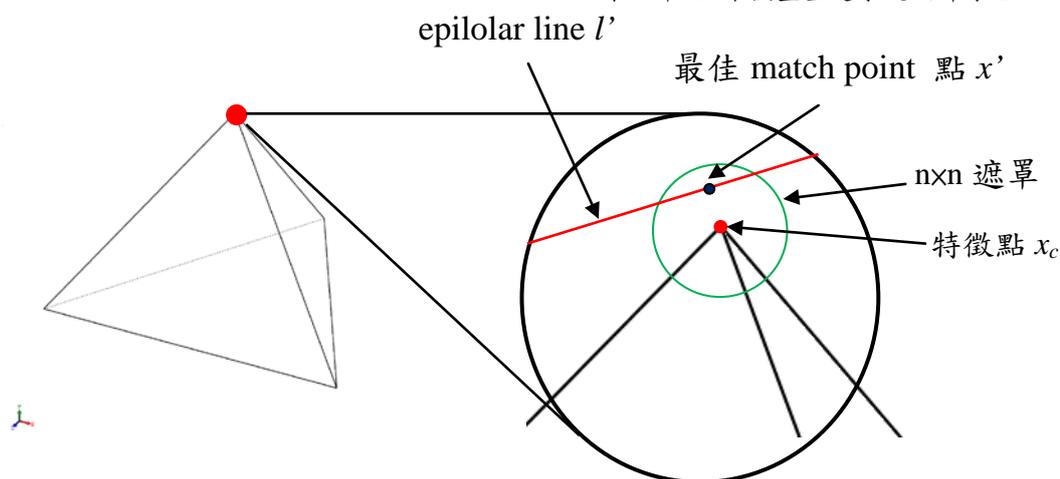


圖 4-6：match point 最佳化

4.3.3 曲面物件定位 match point

曲面物件不同於簡單幾何物件有明顯之特徵點可供重建，以往之文獻大部分是透過外部使用投影機投射結構光至物件表面協助系統重建物件之三維模型，但是透過結構光重建物件模型容易受到物件表面特性的影響，例如結構光會受到物件表面顏色的干擾甚至結構光與物件表面同色則會造成結構光無法順利反射至鏡頭，以至於物件輪廓無法被重建，如此一來可使用結構光重建之物件就會因為表面顏色而有所限制了，除此之外結構光也容易受到外部光源的干擾，如此一來即降低了重建系統之便利性，為解決結構光的缺點本研究提出虛擬網格之建立，利用系統自動產生之虛擬網格代替結構光重建曲面物件之三維模型，此虛擬網格須具備以下幾點特性。

1. 虛擬網格需涵蓋物件完整影像
2. 虛擬網格依表面輪廓之複雜度調整解析度
3. 虛擬網格點需近似物件輪廓
4. 虛擬網格外輪廓需符合物件外部輪廓

接下來將詳述系統如何自動產生最佳之網格，首先虛擬網格需涵蓋物

件完整影像，系統會依虛擬網格重建物件之三維模型，若虛擬網格未涵蓋到完整之物件輪廓，那麼重建的三維模型也不會是完整的物件表面輪廓，故為重建最佳之三維模型也為避免遺失物件表面輪廓資訊，虛擬網格需涵蓋物件完整影像，因此透過邊緣偵測以及影像型態學處理過後之物件封閉輪廓影像，從水平方向以及垂直方向偵測出物件輪廓之起始點以及終點設定成虛擬網格產生之起始點以及終點，如圖 4-7 所示。

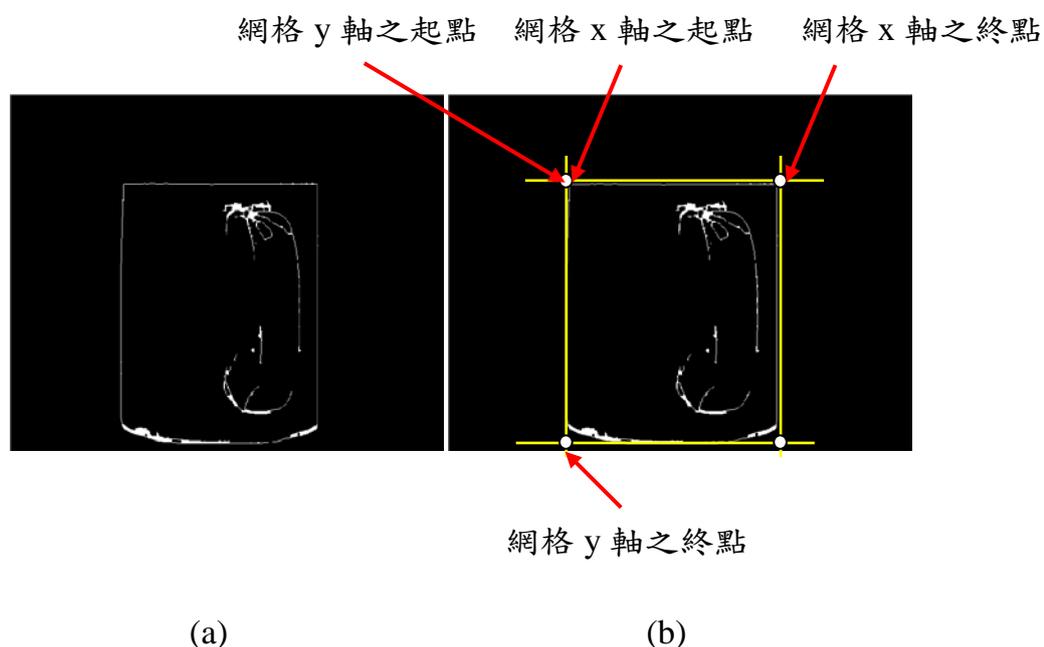


圖 4-7：(a)物件之封閉輪廓；(b)網格產生之起始點與終點

設定好網格產生之範圍後接著對物件輪廓進行切割沿著 x 軸方向與 y 軸方向進行網格的分佈，網格的解析度會影響重建的效果，解析度越高取樣資訊越多其重建效果越好，當相對的就必須花費較多的時間進行運算與重建，而本研究之目的是希望做到即時的重建效果，因此運算量是越少越好，因此我們希望以最少需求之網格點重建能提供足夠物件資訊之三維模型，為達此一目的我們希望虛擬網格能依表面輪廓之複雜度調整解析度，以較高解析度重建物件較為複雜之部分，以較低解析度重建物件較為平坦

或是較為平滑之部分，如此一來在減低取樣資訊時也不會造成重建三維模型的嚴重失真。

首先網格延著 x 軸的方向從起始點進行切割，在切割的同時會延著 y 軸的方向偵測輪廓的複雜度，偵測的方法為沿著 y 軸的方向計算通過邊緣的次數，使用者可設立一門檻值當計算通過邊緣的次數大於門檻值時判定為複雜表面，此時網格解析度設定為小，當計算通過邊緣的次數小於門檻值時判定為複雜表面，此時網格解析度設定為大，而網格解析度之大小也可由使用者決定，因此若是網格沿著 x 軸的方向進行切割，在切割的同時沿著 y 軸的方向偵測輪廓的複雜度，同理若是網格沿著 y 軸的方向進行切割，則在切割的同時沿著 x 軸的方向偵測輪廓的複雜度，如此一來即可用不同解析度的網格分割不同複雜度的物件表面，達到節省運算量之目的，其網格分割以及輪廓之偵測如圖 4-8、4-9 所示。

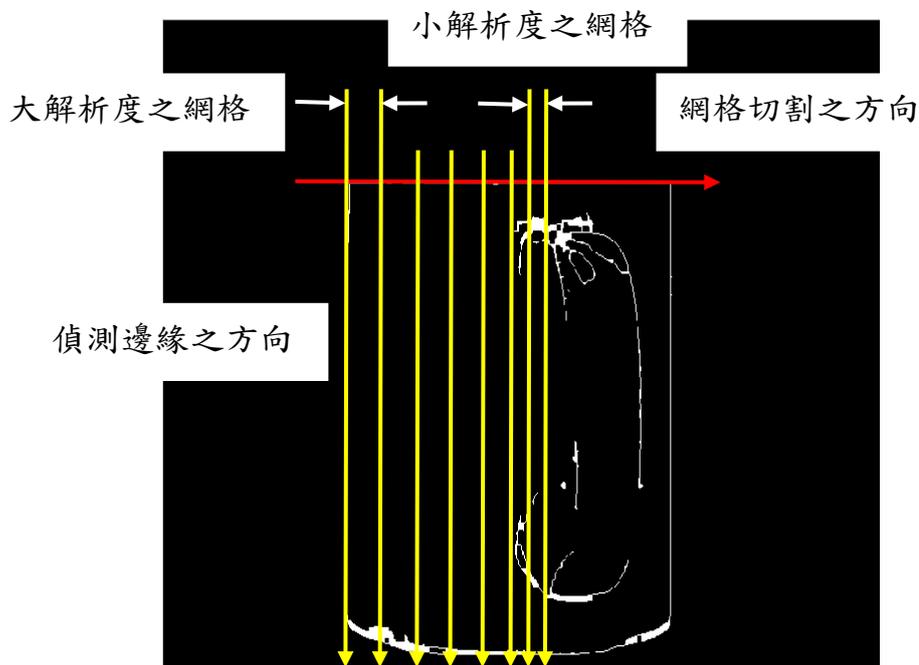


圖 4-8：網格延著 x 軸的方向進行切割

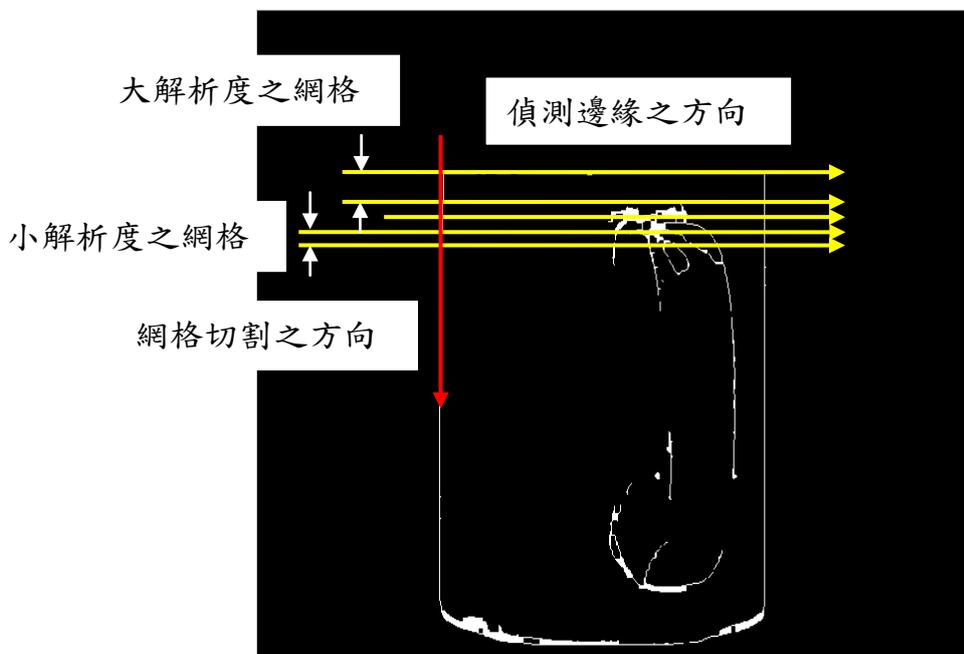


圖 4-9：網格沿著 y 軸的方向進行切割

在網格切割的時候有可能會發生以下這種情況，系統偵測邊緣的結果為使用解析度較大之網格，下一步系統偵測邊緣的結果為使用解析度較小之網格，但此時細小網格的切分並非位於複雜輪廓之起始位置，亦或是系統偵測邊緣的結果為使用解析度較小之網格，下一步系統偵測邊緣的結果為使用解析度較大之網格，但此時細小網格的切分並非位於複雜輪廓之終點位置，為解決此一問題當系統偵測到網格解析度之改變時則以較小解析度之網格掃描前一步驟，目的就是盡量避免遺失物件表面資訊如圖 4-10 所示。

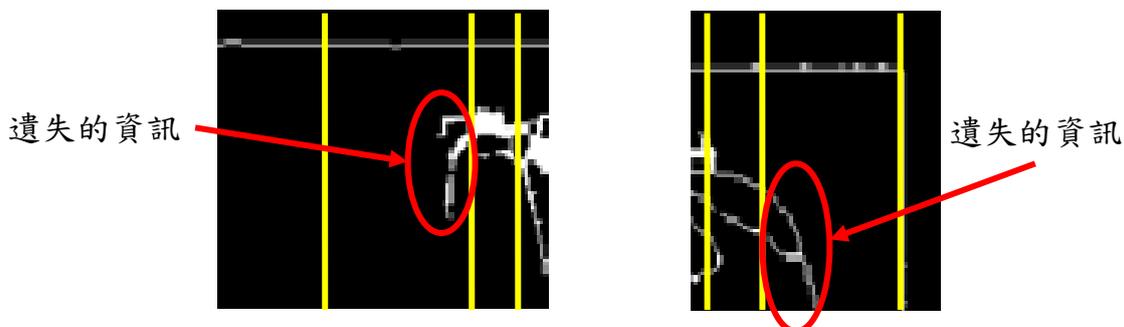


圖 4-10：網格切割所遺失的資訊

由於曲面物件少有特徵點可供重建，因此輪廓資訊對於曲面物件來說是珍貴的資訊，為了優化重建結果，我們使用 3x3 矩陣做為遮罩對每一個網格點進行掃描，以網格點為中心參考網格點周遭八點之像素資訊，判斷網格點是否位於輪廓周圍如圖 4-11 所示。

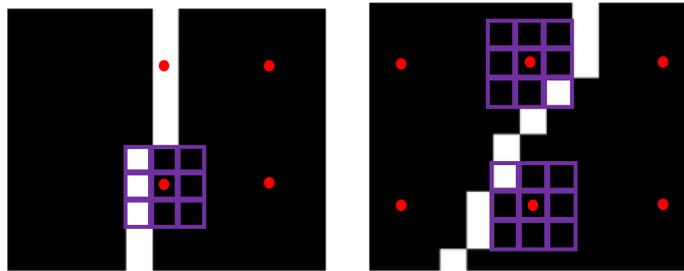


圖 4-11：使用 3x3 矩陣做為遮罩對每一個網格點進行掃描

輪廓出現於遮罩內會有多種情形，因此我們將針對不同情形進行判斷，然後依判斷結果微調網格點之位置，首先我們將輪廓線與網格點之相對位置的情況做以下初步分類，分類如圖 4-12 所示。

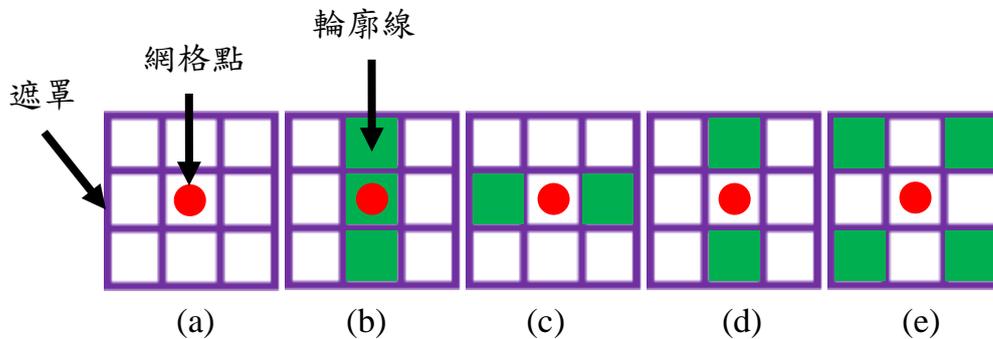


圖 4-12：(a)未偵測到輪廓線；(b)網格點位於輪廓線上；(c)輪廓線位於遮罩左右兩側；(d)輪廓線位於遮罩上下兩側；(e)輪廓線位於遮罩四端

情況一：網格點周遭未偵測到輪廓線

情況二：網格點位於輪廓線上

情況三：輪廓線出現於網格點左右兩側

情況四：輪廓線出現於網格點上下兩側

情況五：輪廓線位於遮罩四端

輪廓線於遮罩內的分佈情況主要分為五上情形，但是也有機率情況二至五同時出現，為了順利將網格點移至輪廓線上避免系統產生錯誤，因此我們將遮罩內的九個點均給予不同之權重值，因此當情況二至五同時出現時系統會依照事先給予的權重值情形判別，設遮罩之權重值為 $W=w_{ij}$ ， $i=1,2,3$ ； $j=1,2,3$ ， $w_{22}>w_{21}>w_{23}>w_{12}>w_{32}>w_{11}>w_{13}>w_{31}>w_{33}$ ， i,j 表示此權重值位於遮罩之位置，若遮罩某一位置出現輪廓線則將此位置之權重值 $w_{ij}\times 1$ ，否則 $w_{ij}\times 0$ ，將遮罩之九個點均做此處理，最後將網格點優先移至最大值之位置，網格點優化步驟流程如圖 4-13 所示，圖 4-14 為示意圖。

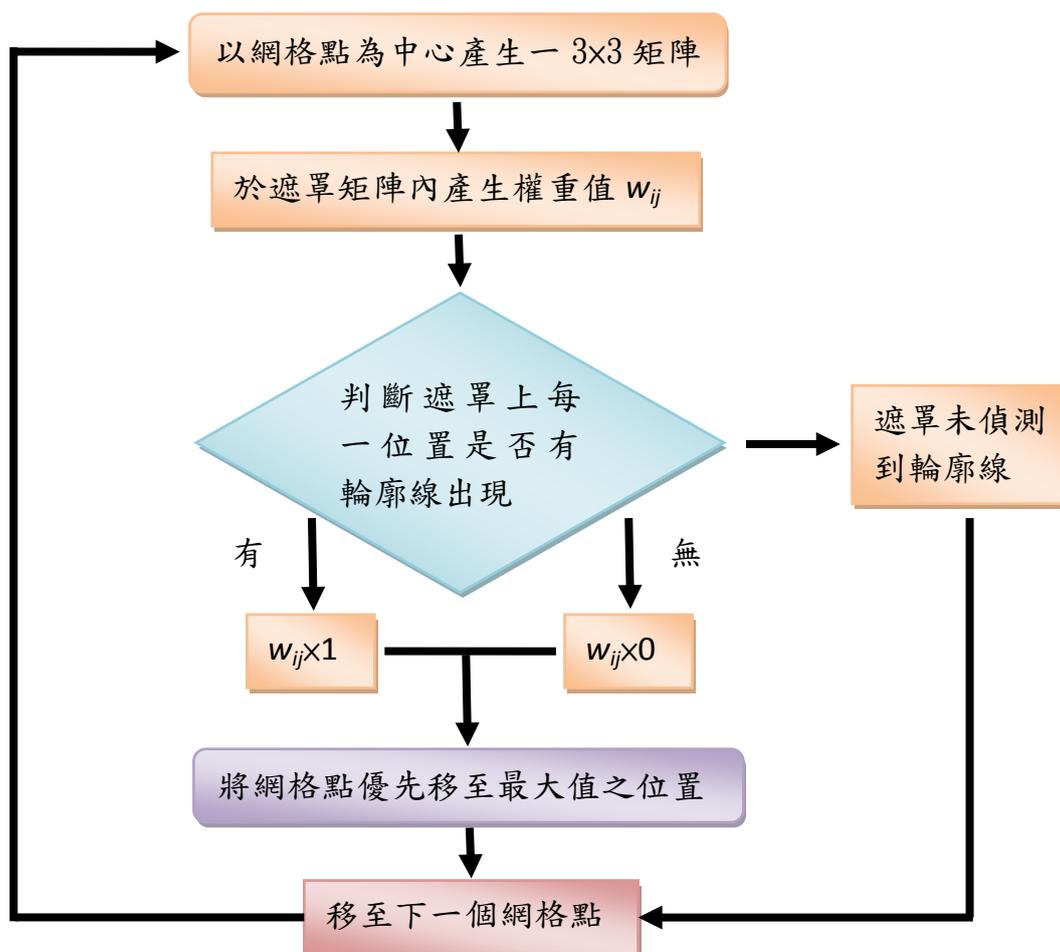


圖 4-13 網格點優化步驟流程圖

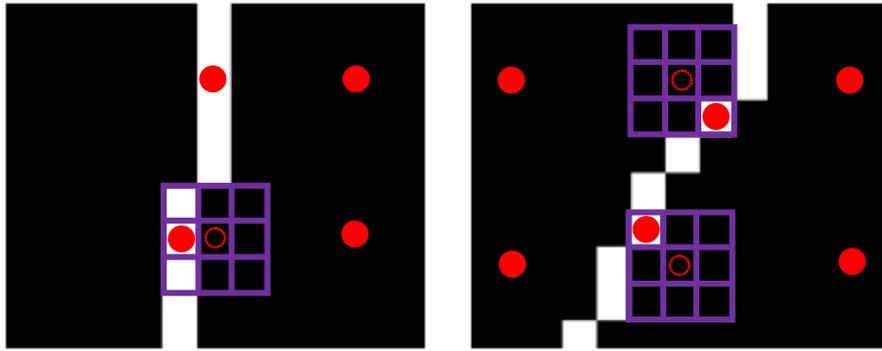


圖 4-14：網格點優化示意圖

將網格進行優化之後接著將產生的網格與物體輪廓進行比對，比對之主要目的是要將網格修剪為符合物件之輪廓，雖然網格點已做了初步之優化動作，但是仍有許多不屬於物件輪廓之網格點，因此我們將整個網格與物件輪廓之剪影進行比較，將位於輪廓外部不屬於物件表面之網格點去除，因此將之前處理好的物件封閉輪廓進行處理並且將此封閉輪廓進行填滿形成物件之剪影，整個網格與物件輪廓之剪影進行比對修剪得到符合物件外圍輪廓之網格，如圖 4-15 所示。

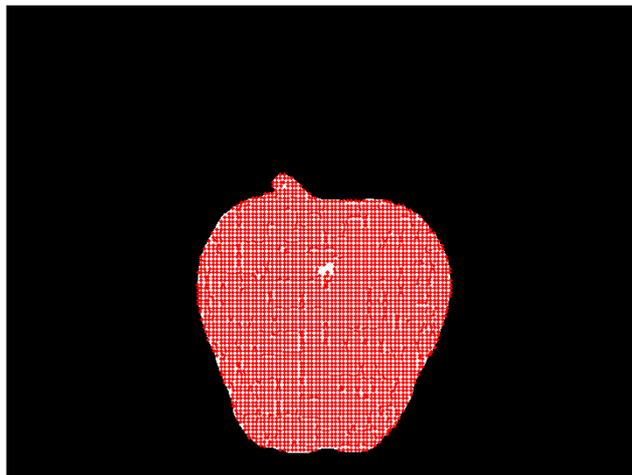


圖 4-15：將網格與物件剪影進行比對修剪

於左邊影像產生網格覆蓋物件影像後，下一步就是根據左邊影像之網格點於右邊影像找出其相對應之 match points，由於曲面物件無明顯之特徵點可協助系統進行 match points 之定位，因此我們使用 epipolar line 輔以水

平搜尋的方式定未出 match points ，第三章中三圍重建系統配置內有提到本系統將兩 CCD 攝影機固定至同一平台上，藉此維持兩 CCD 攝影機為等高之狀態，並且調整兩 CCD 攝影機護為平行，若兩 CCD 攝影機為等高也互為平行之狀態下則其投影中心點也會是等高之狀態，同理投影平面也會是等高之狀態，如此一來空間中之一點投影至左右兩投影面時其左右兩投影點也會位於等高之位置如圖 4-16 所示。

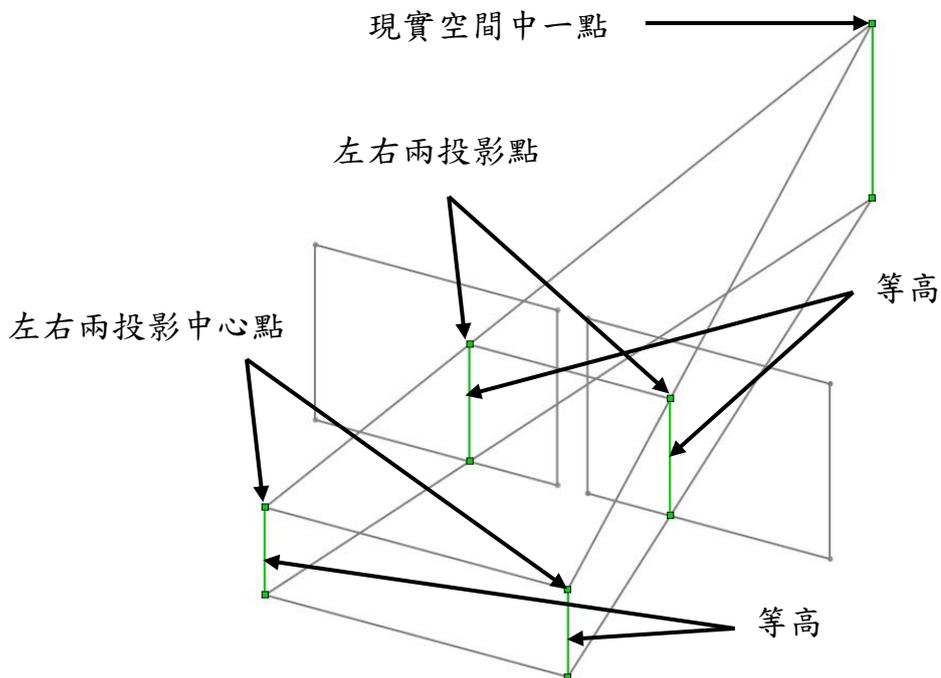


圖 4-16：投影等高示意圖

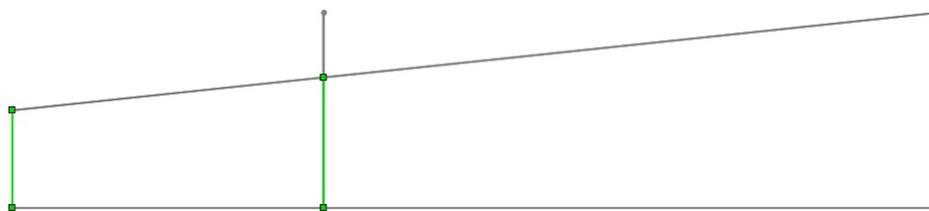


圖 4-17：投影側視圖

欲從右投影面定位出 match points 同樣的先依據 epipolar geometry 藉由左投影點與 fundamental matrix 於右投影面繪出 epipolar line，epipolar line 為 match point 出現之集合詳細情形已於之前章節詳述過了，結合右投影點會出現在與左投影點等高位置之特性，從左投影點做水平搜尋，水平線與 epipolar line 之交點即為 match point 如圖 4-18 所示。

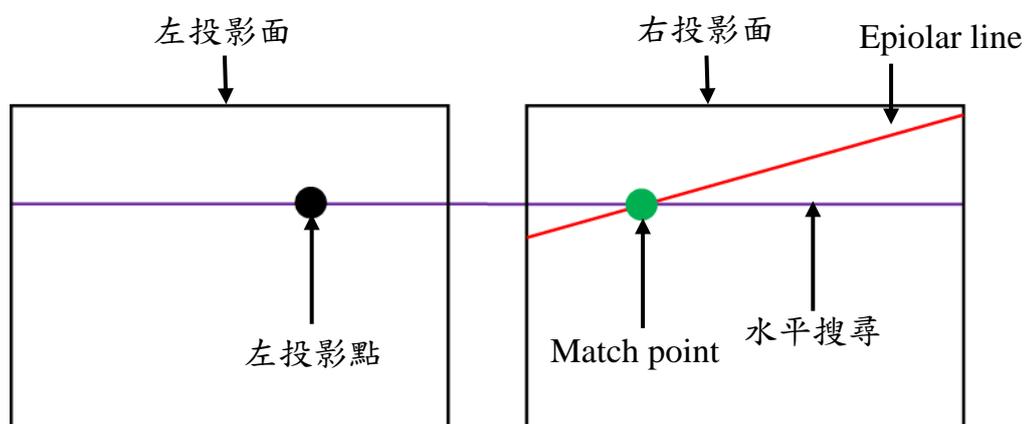


圖 4-18：match point 定位示意圖

將左影像上之所有網格點均做水平搜尋以及與 epipolar lines 作相交定位出所有左影像上之網格點位於右影像之 match points，如此一來即可用左影像之虛擬網格與右影像之 match points 藉由 stereo vision 之原理重建出曲面物件之三維模型。

4.4 建立虛擬三維空間座標

本研究為重建物件之三維模型，因此需事先產生虛擬之三維空間座標，接著定義現實空間座標，二維影像座標，相機座標。之後再找出這些座標系統之間的轉換關係，如此一來系統即可藉由二維影像重建物件之三維模型。先前章節以有事先詳細描述現實空間中之物體，透過攝影機投影產生二維之影像，因此欲從二維影像重建物件之三維模型其順序與成像過程為

相反，重建的步驟為將二維影像藉由相機空間轉成三維座標，之後再利用立體視覺法重建物件模型於虛擬之三維空間座標，此虛擬之三維空間座標即是模擬現實空間座標。

因此首先我們定義二維影像座標如圖 4-19 所示，二維影像座標之原點位於投影面之左上角。

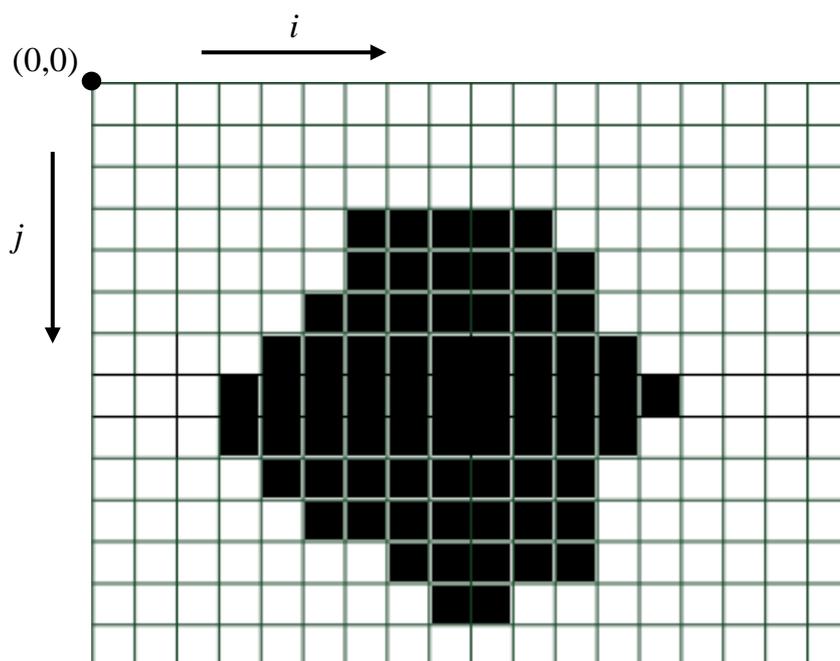


圖 4-19：二維影像座標

接著定義相機空間，設定相機空間以投影中心點為原點以光軸方向為 z 軸，並將二維影像座標轉換至相機空間中如圖 4-20 所示，式子 4-12 為轉換表示式，在相機空間中，投影面之原點由右上角移至投影面之中心，投影面與相機空間之原點距離為焦距 f 。

$$\begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = \begin{bmatrix} 1 & 0 & -u_0 \\ 0 & 1 & -v_0 \\ 0 & 0 & f \end{bmatrix} \begin{bmatrix} i \\ j \\ 1 \end{bmatrix} \quad (4-12)$$

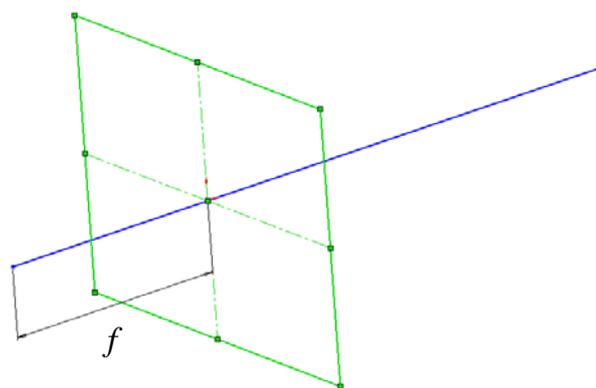


圖 4-20：相機座標

最後定義系統所創造之虛擬三維空間，虛擬三維空間與相機空間存在一比例關係，因此在將相機空間轉換至虛擬三維空間前須先將相機座標空間正規化，並且考慮水平與垂直焦距，因此我們將式子(4-12)改寫為式子(4-13)。

$$\begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = \begin{bmatrix} 1/f_x & 0 & 0 \\ 0 & 1/f_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -u_0 \\ 0 & 1 & -v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} i \\ j \\ 1 \end{bmatrix} \quad (4-13)$$

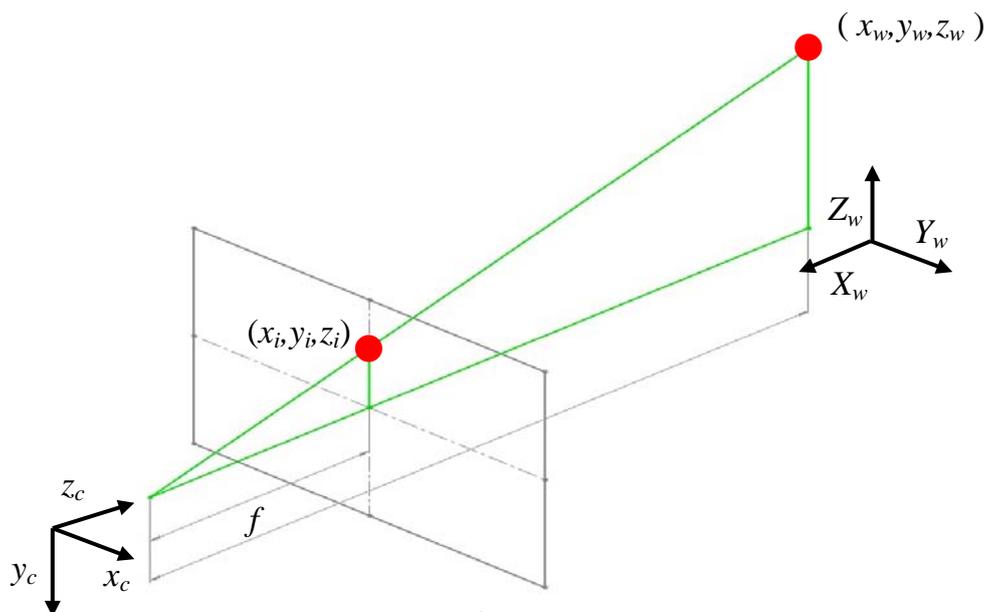


圖 4-21：相機座標與虛擬三維空間座標

虛擬三維空間中一點 (x_w, y_w, z_w) 投過相機空間投影至投影面上產生投影點，此投影點座標為 (x_i, y_i, z_i) ，且從圖 4-21 中可看出，相機空間坐標軸 (x_c, y_c, z_c) 之方向與虛擬三維空間座標軸 (X_w, Y_w, Z_w) 之方向不同，因此我們利用一轉換矩陣將相機空間轉換至虛擬三維空間如式子 4-14 所式，最後在乘上一比例值 S 即可完成轉換，將二維影像座標轉換至虛擬三維空間座標如式子 4-15 所示。

$$\begin{bmatrix} X_w \\ Y_w \\ Z_w \end{bmatrix} = \begin{bmatrix} 0 & 0 & -1 \\ 1 & 0 & 0 \\ 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} \quad (4-14)$$

$$\begin{bmatrix} X_w \\ Y_w \\ Z_w \end{bmatrix} = S \begin{bmatrix} 0 & 0 & -1 \\ 1 & 0 & 0 \\ 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} 1/f_x & 0 & 0 \\ 0 & 1/f_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -u_0 \\ 0 & 1 & -v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} i \\ j \\ 1 \end{bmatrix} \quad (4-15)$$

4.5 重建三維模型

將投影面轉換至虛擬三維空間中之後即可根據立體視覺法 epipolar geometry 進行重建，將左投影面之特徵點或是網格點以及右投影面的 match points 轉換成虛擬三維空間座標，在虛擬三維空間中根據 epipolar geometry 之原理，藉由兩投影中心點以及左投影點與右相對應投影點 match point 重建空間中之一點， l_1 為左投影點與左 CCD 攝影機投影中心點之連線， l_2 為右投影面相對應投影點 match point 右 CCD 攝影機投影中心點之連線， l_1 與 l_2 之交點即為重建點，如圖 4-22。

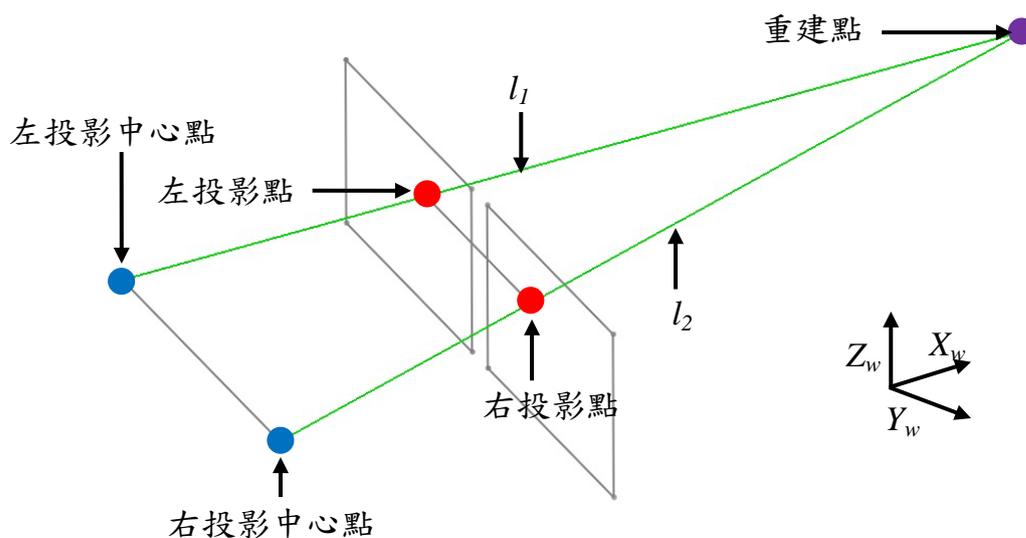


圖 4-22：虛擬三維空間重建示意圖

4.6 結合運動模型

受限於攝影機一次只能擷取一方向之物件表面資訊，因此一次只能重建局部之三維模型，若是欲重建完整之三維模型可藉由結合運動模型重建完整之三維模型，將物件進行旋轉進行拍攝或是將攝影機裝置在機械手臂上沿著物件周圍拍攝一圈，若是將物件擺至平台上旋轉則戴入平台旋轉之運動模型，若示將設影機裝置在機械手臂上沿著物件周圍拍攝一圈則結合機械手臂之運動模型，本實驗選擇將物件放至平台上進行旋轉，每次旋轉平台 90° ，將已重建好之局部物件三維模型結合運動模型如式子 4-16 所示， P_p 為旋轉前重建點所構成之矩陣， R 為旋轉矩陣， T 為平移矩陣， P_r 為轉移後之重建點，平台圍繞 Z 軸旋轉 90° 可用式子 4-17 表示，圖 4-23 為其旋轉示意圖，物件旋轉一圈後即可重建物件之完整模型。

$$[P_r] = R \times T \times [P_p] \quad (4-16)$$

$$R_z = \begin{bmatrix} \cos \theta_z & -\sin \theta_z & 0 \\ \sin \theta_z & \cos \theta_z & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \theta_z = \frac{\pi}{2} \quad (4-17)$$

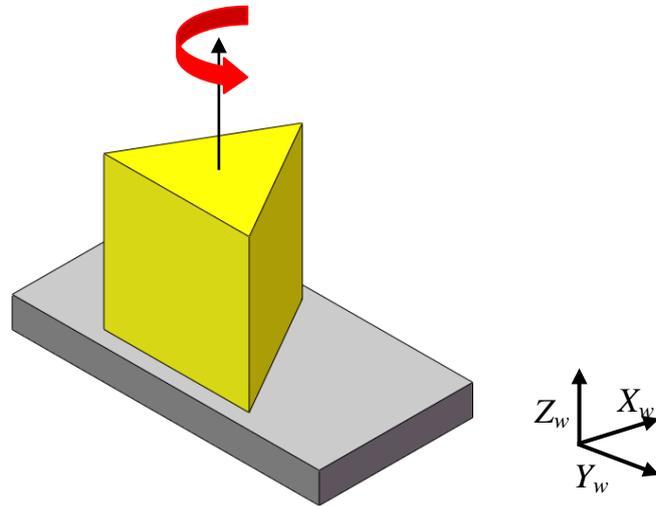


圖 4-23：沿 Z 軸旋轉示意圖

第五章 實驗結果與討論

為驗證此三維模型重建系統之重建效果，本實驗主要採用 matlab 程式語言撰寫重建物件三維模型系統，撰寫一完整流程從影像輸入，影像處理，最後輸出重建之物件三維模型，以及位置資訊。

5.1 系統設備描述

圖 5-1 為本研究之系統硬體設備之架構，其硬體設備如以下所述：

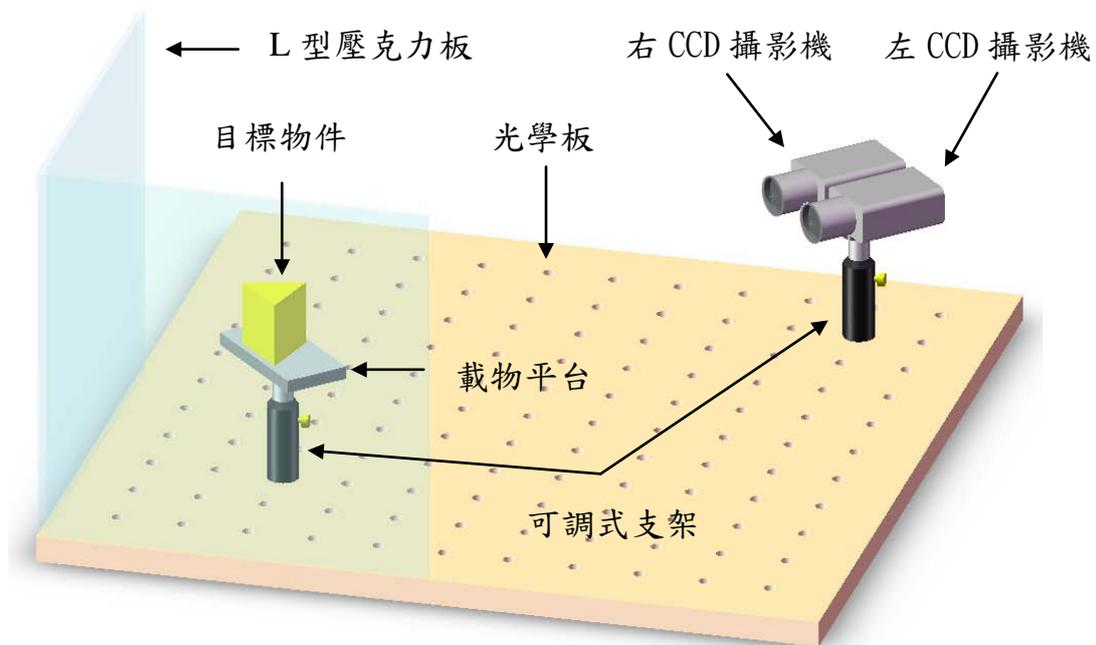


圖 5-1：三維模型重建系統配置圖

1. 個人電腦：

使用 cpu 為 Pentium[R]M 2.00GHz，RAM 為 512MB 的個人電腦，並且使用 matlab 程式語言撰寫重建物件三維模型系統。

2. 影像擷取卡：

本研究所使用之影像擷取卡為美商國家儀器公司所代理，型號為 NI

PXI-8252，傳輸通道為 IEEE 1394，此影像擷取卡支援三個 IEEE 1394 通道。



圖 5-2：影像擷取卡

3.攝影機：

本研究所使用之攝影機為美商國家儀器公司所代理，相機型號為 Basler SCA1600，傳輸通道為 IEEE 1394，相機解析度 1600x1200，最大解析度下 14FPS。



圖 5-3：攝影機

4.鏡頭：

本實驗預設重建之工作範圍物件距離鏡頭 0.5m 以上，物件大小為 30 立方公分以內，因此根據此預設需求選用焦距為 16mm 之鏡頭，因此本研

究所使用之鏡頭為 Computar 公司所代理，型號為 M1614-MP2，焦距 16mm。



圖 5-4：鏡頭

5.偏光鏡：

為降低反光所帶來之雜訊因此於鏡頭前端裝置符合尺寸之偏光鏡。

6.純白 L 型壓克力板：

藉由純白 L 型壓克力板，隔離不必要之背景資訊，方便系統將物件從背景之中擷取出來。

7.光學桌板：

光學桌板為鋁合金材質，經陽極處理強化其剛性，均勻佈滿 16x16 之 M6 螺孔，水平與垂直孔距為 50mm，可將可調式支架藉由全螺牙螺絲固定至光學桌板上。

8.可調式支架：

可調式支架由立柱與立柱支架所構成，可自由調整其長度，可將攝影機或是載物平台固定於立柱上，再將支架固定於光學桌板上，即可將整個系統固定於光學桌板上。

5.2 三維模型重建系統架設操作流程

1. 將左右兩攝影機固定於同一平台上使兩攝影機等高且互為平行，將平台固定於支架上，將支架固定於光學桌上。
2. 將載物平台固定於支架上，將支架固定於光學桌上，將待測物件擺至載物平台上。
3. 將兩攝影機與電腦連線，藉由擷取畫面調整鏡頭之光圈與微調焦距，微調焦距至能清楚擷取物件影像，調整光圈使鏡頭能擷取到清晰之物件影像，調整偏光鏡降低物件表面反光干擾。
4. 利用攝影機擷取之畫面調整攝影機之姿勢與載物平台之高度，調整左右兩攝影機均能擷取到完整之待測物影像。

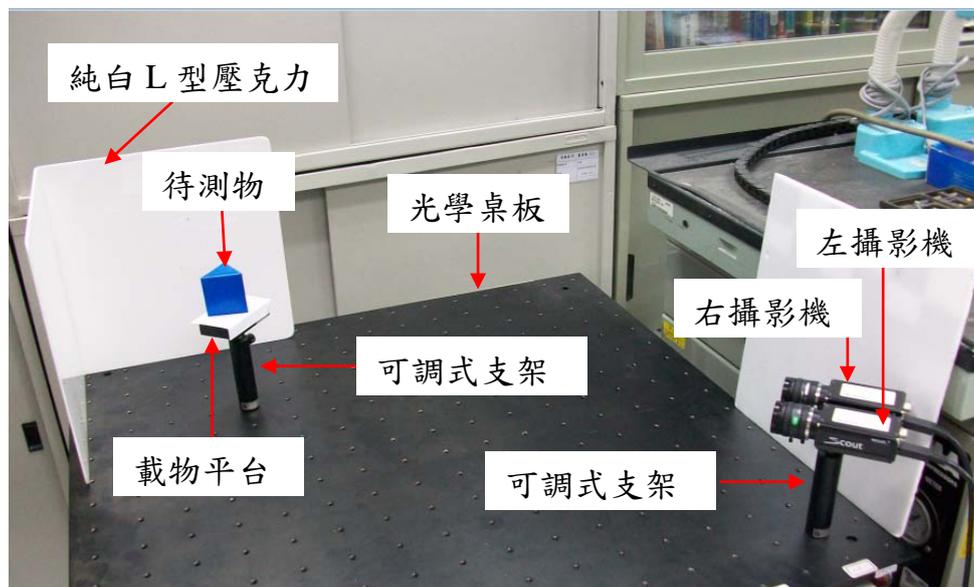


圖 5-5：三維重建系統實體配置圖

5.3 三維模型重建流程

5.3.1 得到相機參數與 Fundamental matrix

對於未知相機參數得到之方法可由軟體進行校正，但是本實驗之相機參數可從產品資訊得之，故不用廢時的進行校正。

計算系統之 Fundamental matrix，欲計算 Fundamental matrix 須帶入 8 組以上已知的相對應投影點進行計算，帶入越多組的已知相對應投影點，求解之 Fundamental matrix 越是準確，而且帶入的已知相對應投影點還原至空間中最好是均勻分布，事先設計好校正板，校正板為 6x6 黑白相間之方塊，每方塊大小為 50x50mm² 如圖 5-6 所示，將校正板以不同角度傾斜以得到空間中各位置之資訊，由左右兩 CCD 攝影機擷取校正板之影像，記錄左右兩影像校正板上方塊之交點位置作為已知之相對應投影點如圖 5-7、圖 5-8 與圖 5-9 所示，已知點於空間之分佈如圖 5-10 所示。

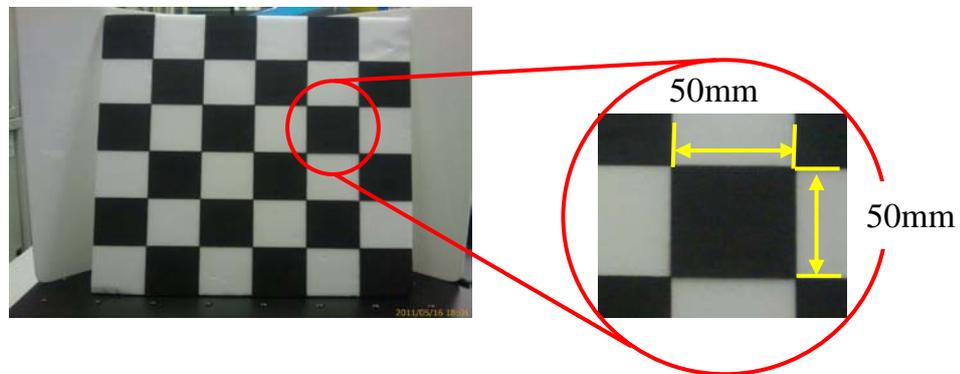
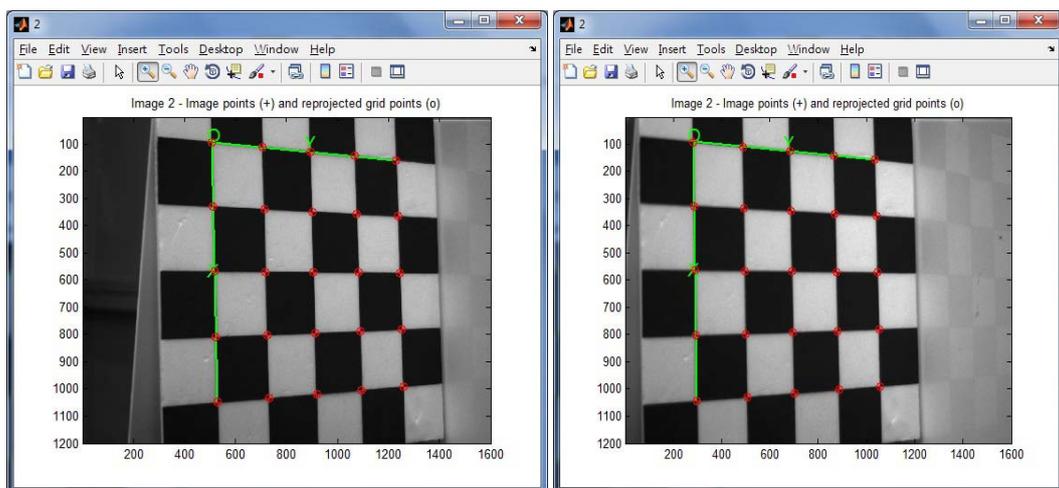


圖 5-6：校正板



(a)

(b)

圖 5-7：(a)左校正板影像；(b)右校正板影像

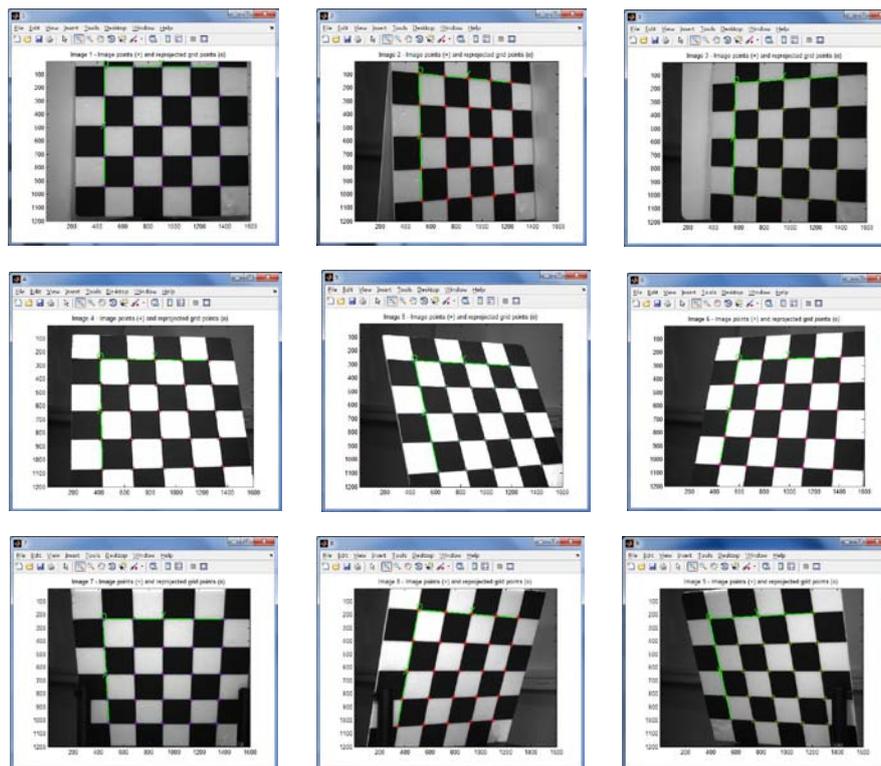


圖 5-8：傾斜各角度之左校正板影像

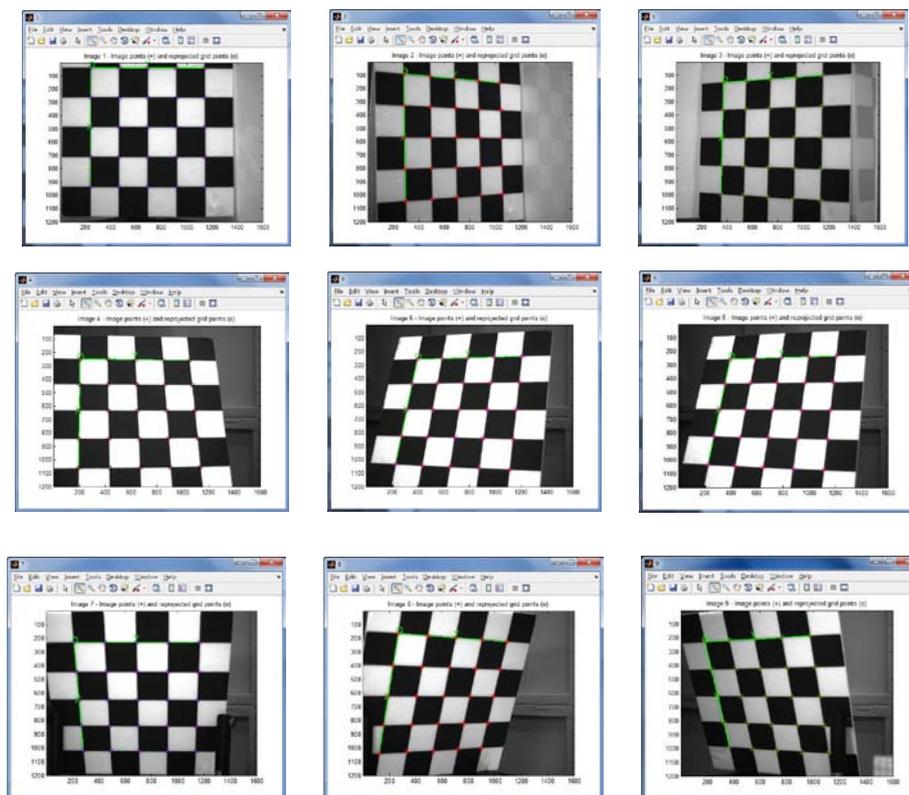


圖 5-9：傾斜各角度之右校正板影像

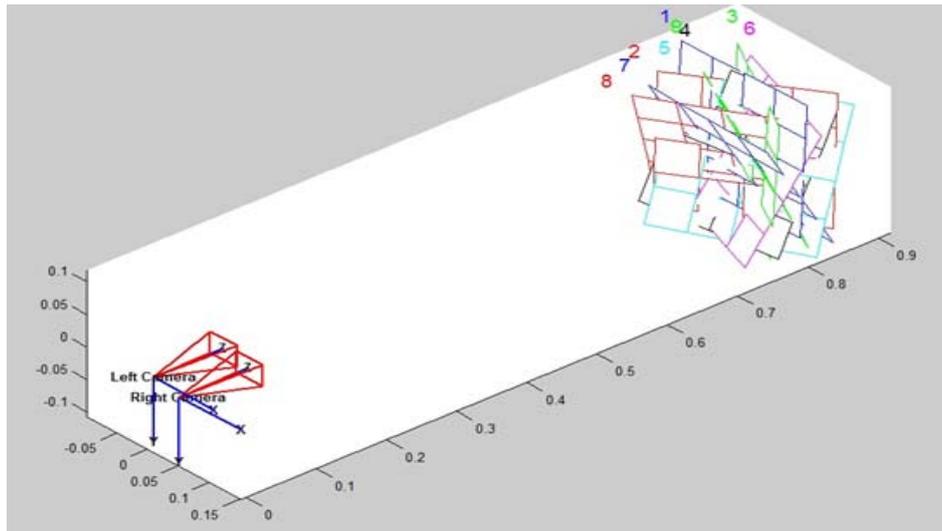


圖 5-10：已知點於空間之分佈圖

5.3.2 重建簡單之幾何物件

A. 影像處理

左右 CCD 攝影機擷取物件之影像輸至系統內進行影像處理。

步驟一：將左右兩影像灰階化，並事先建立背景影像，將有物件之影像與背景影像比較，從背景中擷取出物件影像如圖 5-11 所示。

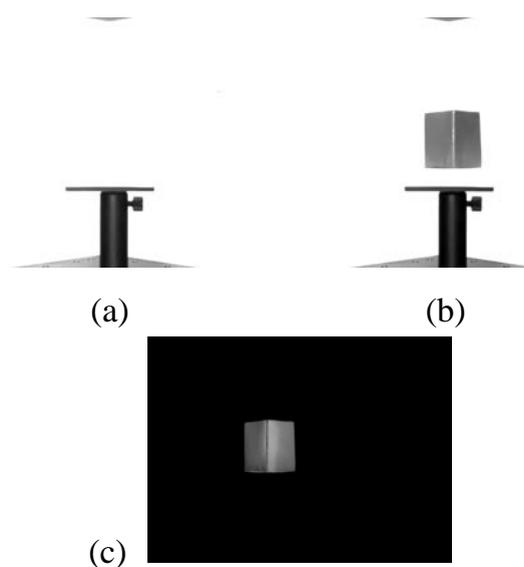


圖 5-11：(a)背景影像；(b)輸入之左影像；(b)物件之影像

步驟二：將左右兩影像作邊緣偵測並使用影像型態學濾除雜訊並強化影像輪廓結果如圖 5-12 所示。

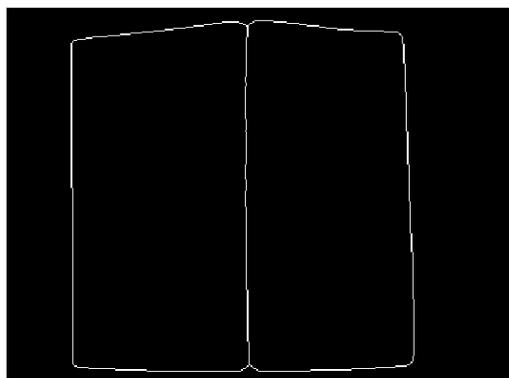


圖 5-12：物件影像之輪廓

步驟三：對左右兩影像作轉角偵測如圖 5-13 所示。

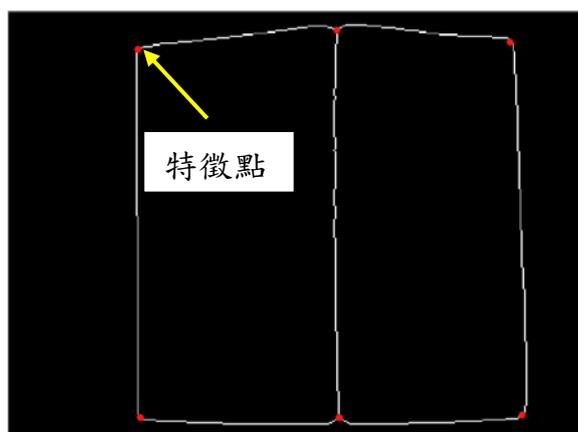


圖 5-13：轉角偵測

B. 搜尋 match point

將左影像上已偵測到之特徵點使用 fundamental matrix 於右影像繪出其相對應之 epipolar line，以右影像上已偵測到之相同位置的特徵點為圓心產生一 $n \times n$ 遮罩，以此遮罩與 epipolar line 作交集，以此交集之均值作為最佳 match point，最佳 match point 之找尋如圖 5-14 所示。

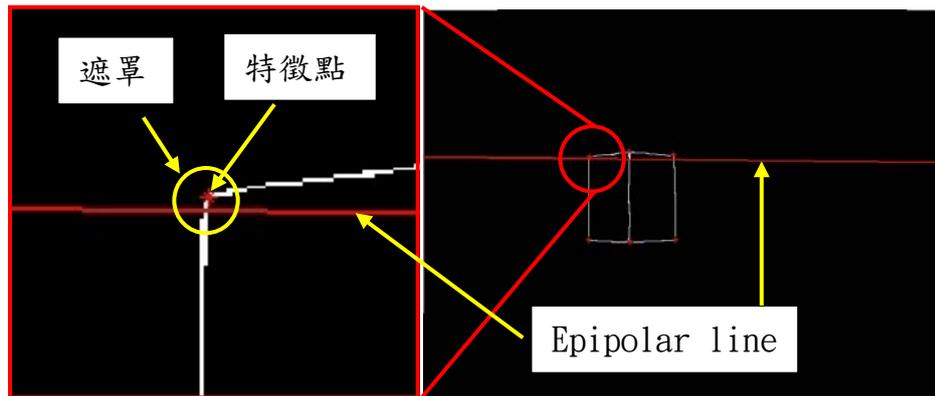


圖 5-14：最佳 match point 之找尋

C. 重建物件之三維模型

使用立體視覺法帶入相機參數以及 match points 重建物件局部之三維模型，結合運動模型重建出物件之完整三維模型，重建結果包含物件之位置資訊，重建結果如圖 5-15~20 所示。

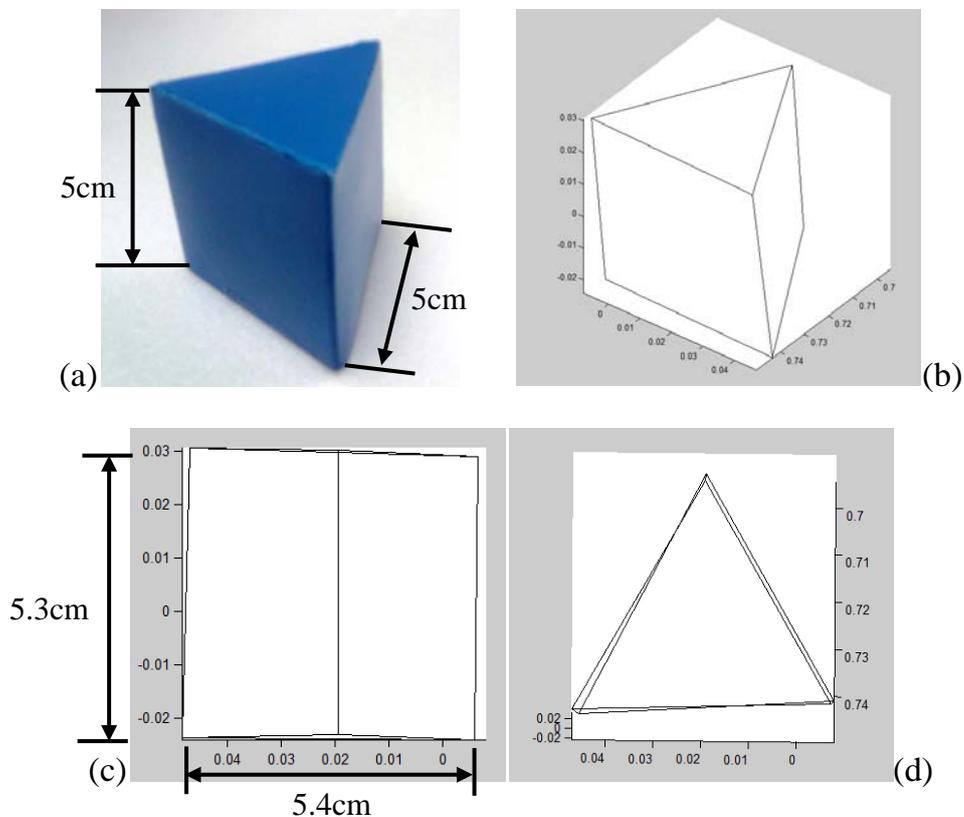


圖 5-15：(a)實體物件；(b)重建結果；(c)模型前視圖；(d)模型上視圖

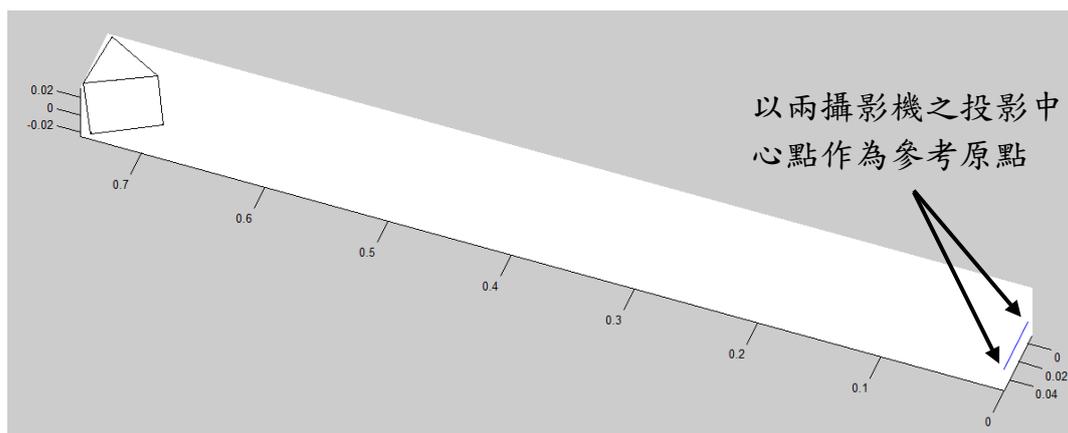


圖 5-16：模型之位置資訊

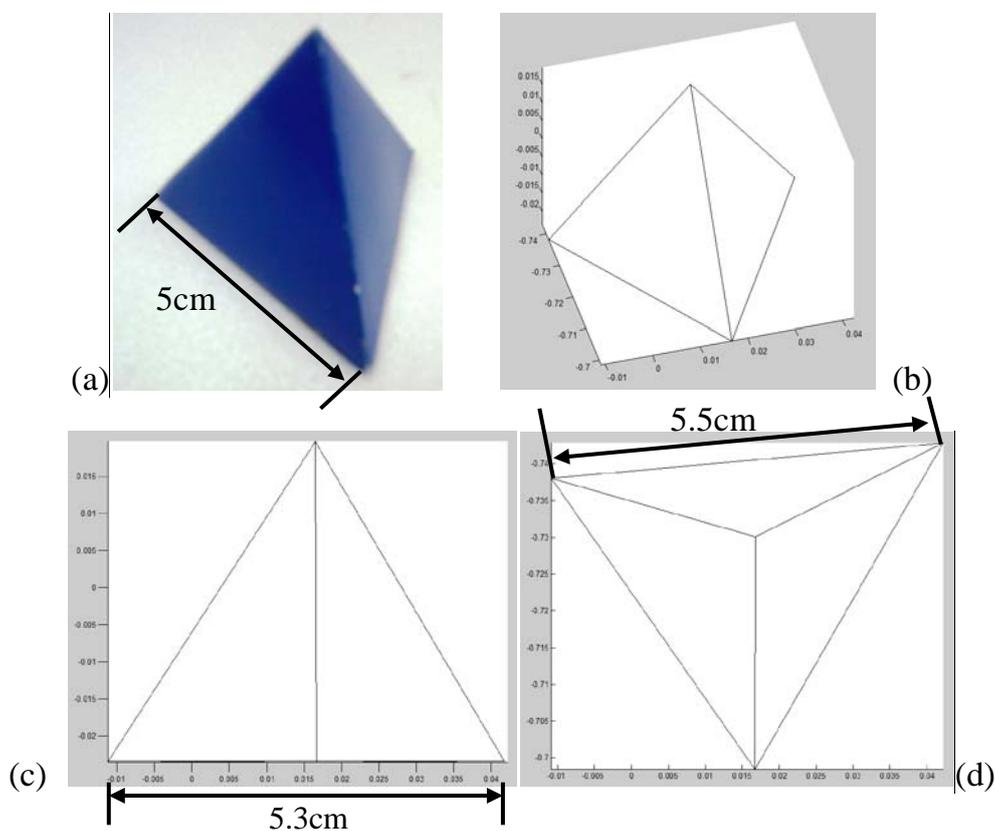


圖 5-17：(a)實體物件；(b)重建結果；(c)模型前視圖；(d)模型上視圖

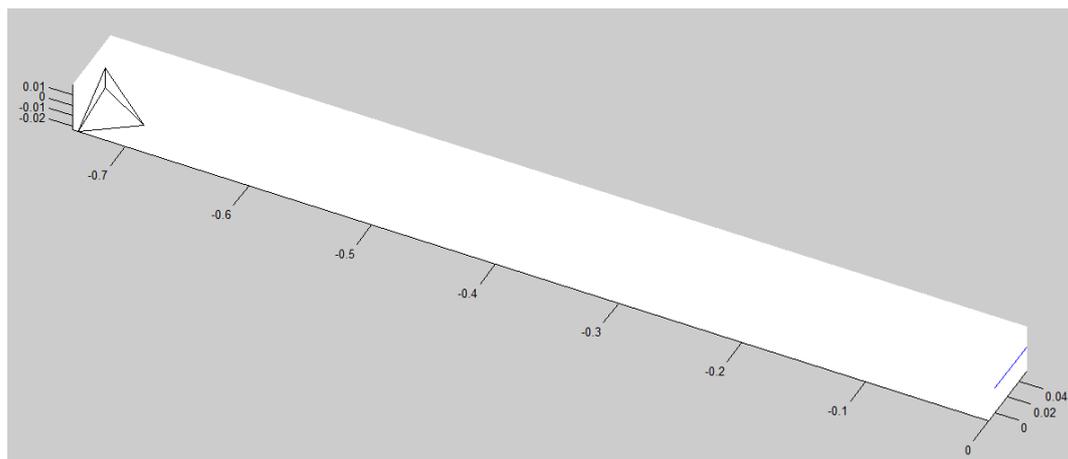


圖 5-18：模型之位置資訊

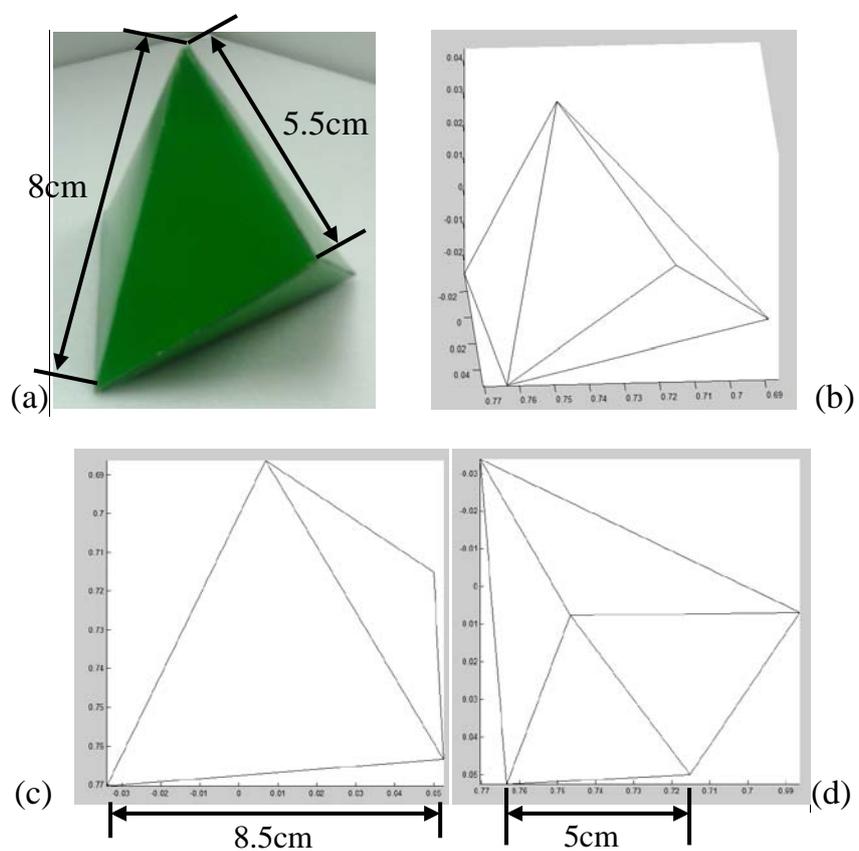


圖 5-19：(a)實體物件；(b)重建結果；(c)模型側視圖；(d)模型上視圖

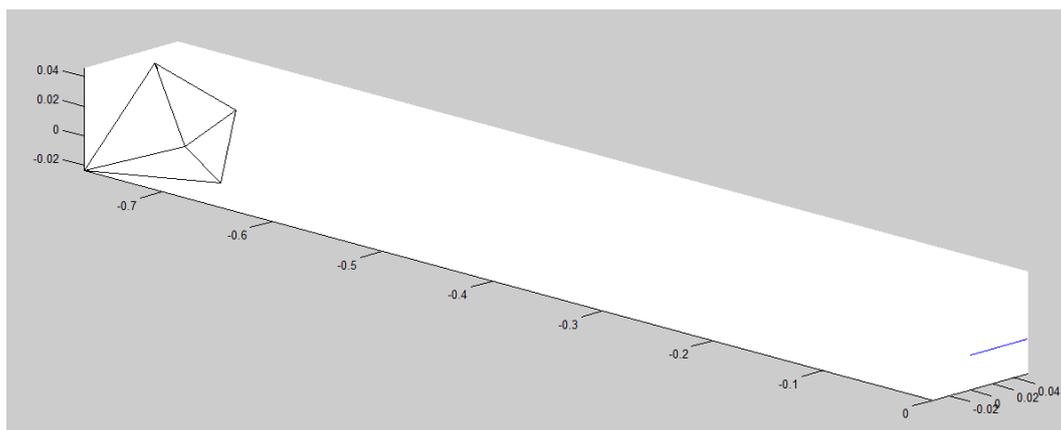


圖 5-20：模型之位置資訊

5.3.3 重建曲面物件

A. 影像處理

影像處理之步驟已於前面敘述過了，對於重建曲面物件則多了產生物件剪影此步驟，將處理好之物件封閉輪廓填滿產生剪影，如圖 5-18 所示。



(a)

(b)

圖 5-21：(a)物件原始影像；(b)物件之剪影

B. 產生虛擬網格

系統自行掃描物件表面輪廓自動產生虛擬網格如圖 5-22 所示。

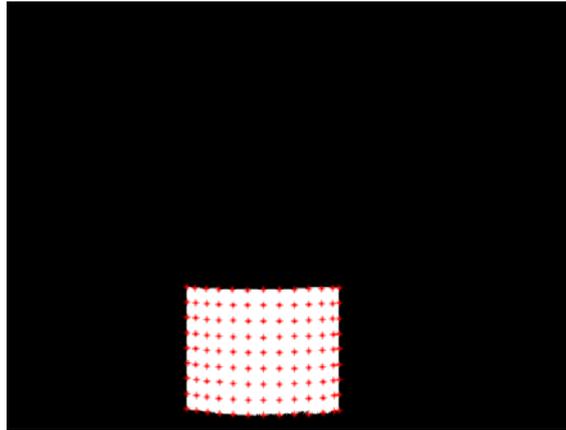


圖 5-22：網格分佈

C. 搜尋 match point

從左影像之網格點做水平搜尋，以左影像之網格點作為搜尋之基準找出等高水平線與 epipolar line 之交點，定位出 match points，將所有網格點均依此步驟找出所有的 match points，結果如圖 5-23 所示。

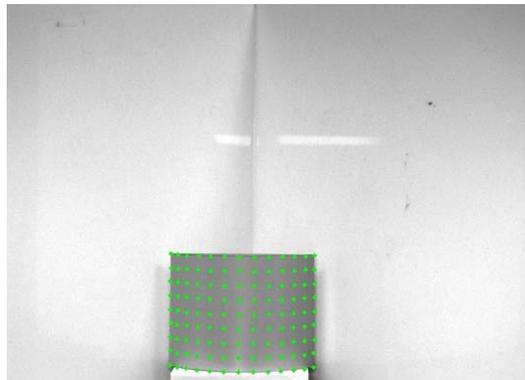


圖 5-23：右影像 match points 之分佈

D. 重建物件之三維模型

使用立體視覺法帶入相機參數以及 match points 重建曲面物件之三維模型如圖 5-23 所示。

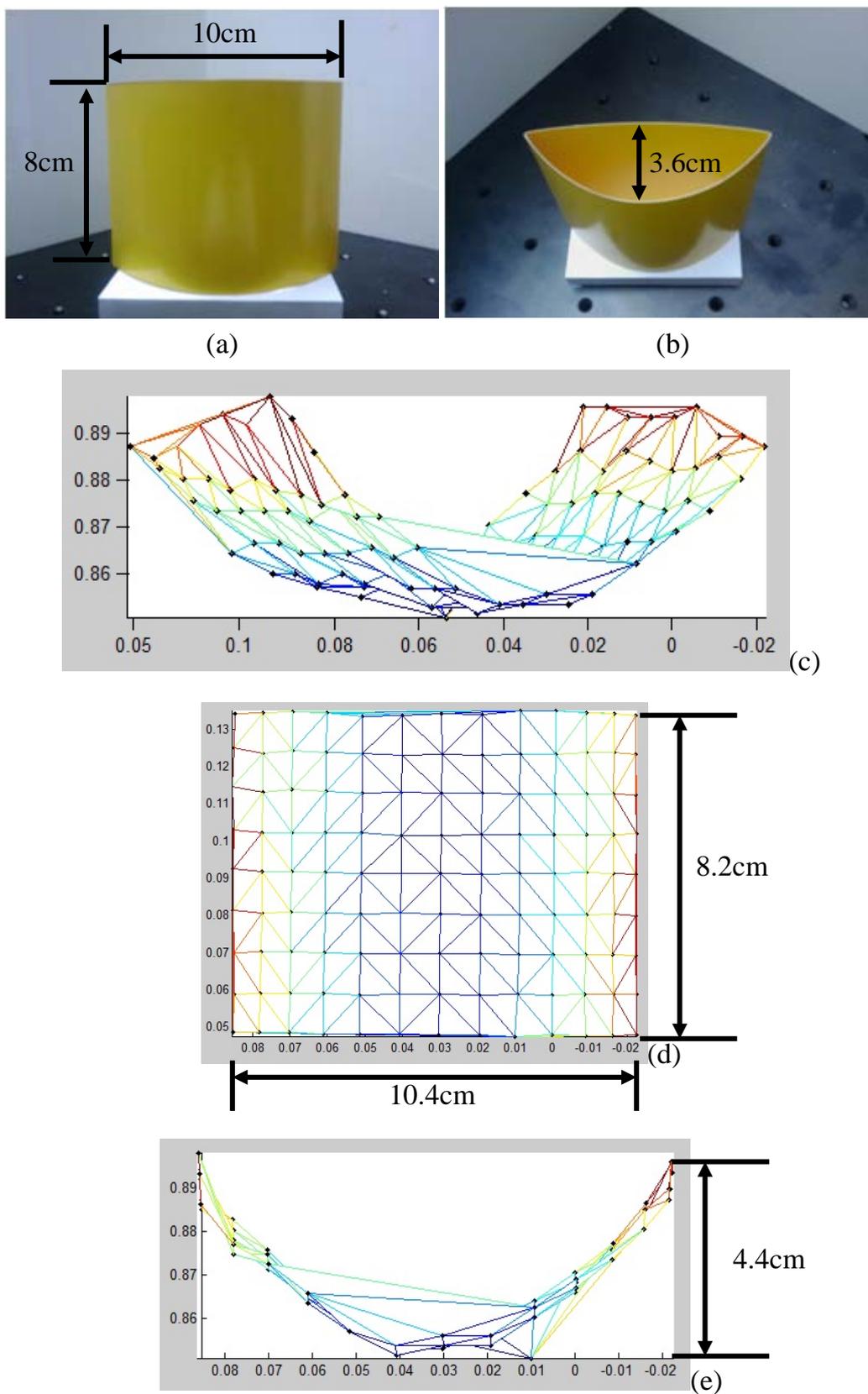


圖 5-24：(a)原始物件前視圖；(b)原始物件上視圖；(c)重建結果；(d)模型前視圖；(e)模型上視圖

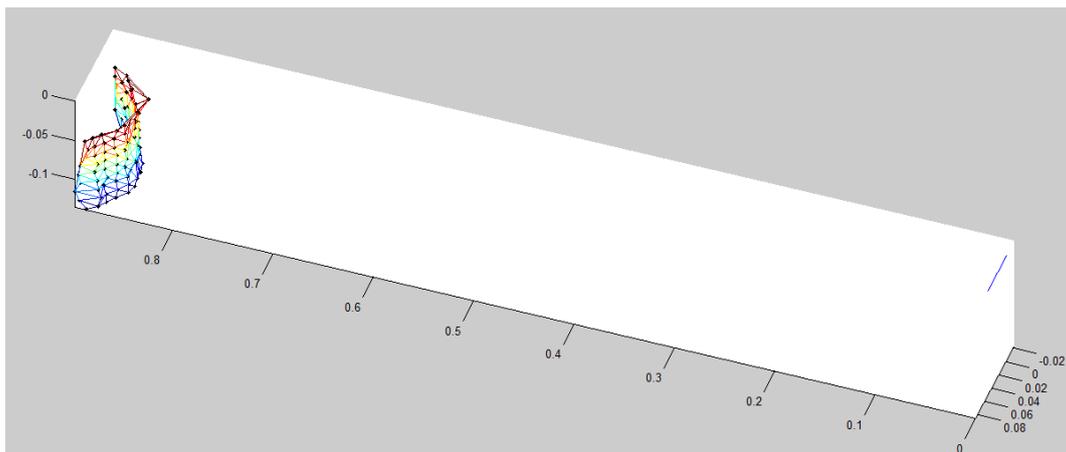
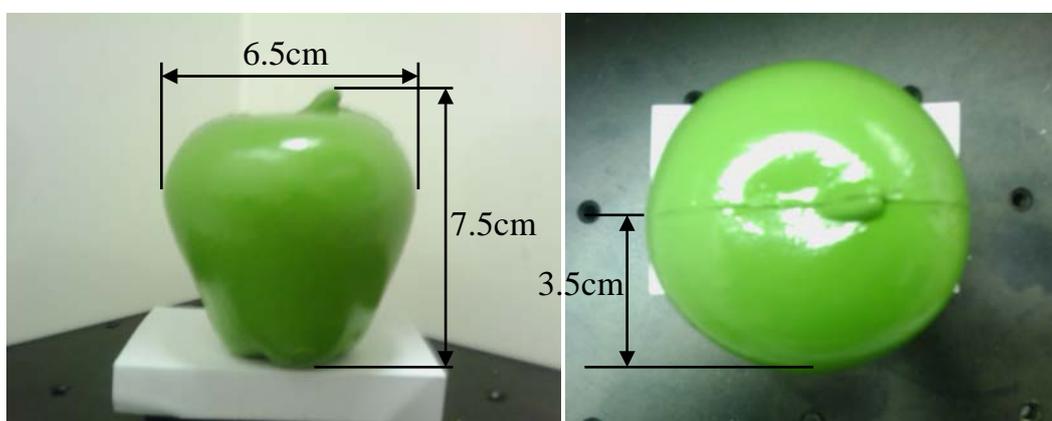
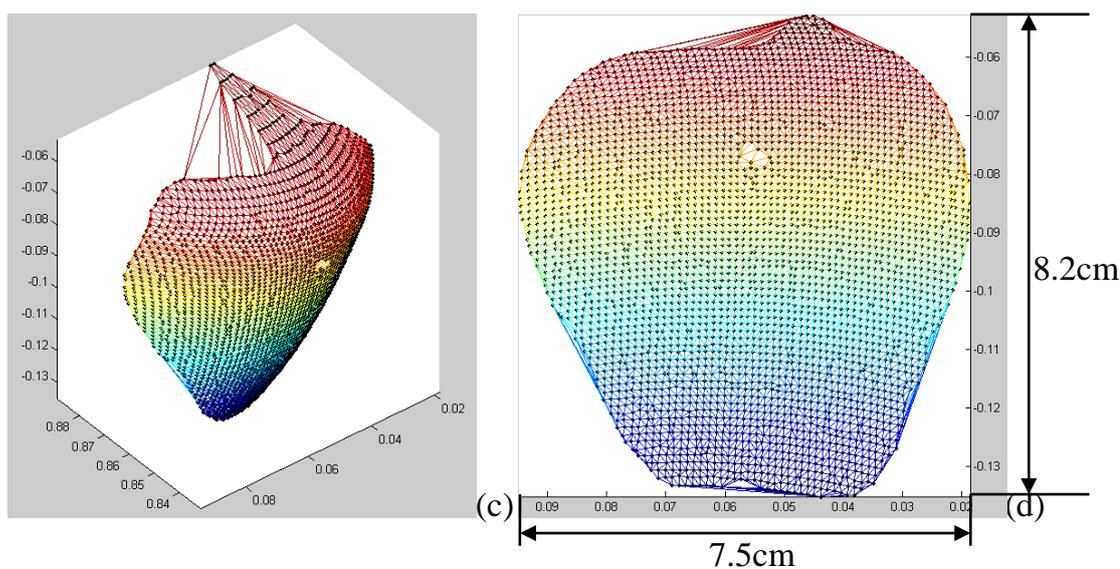


圖 5-25：模型之位置資訊



(b)

(b)



(c)

(d)

圖 5-26：(a)原始物件前視圖；(b)原始物件上視圖；(c)重建結果；(d)模型前視圖

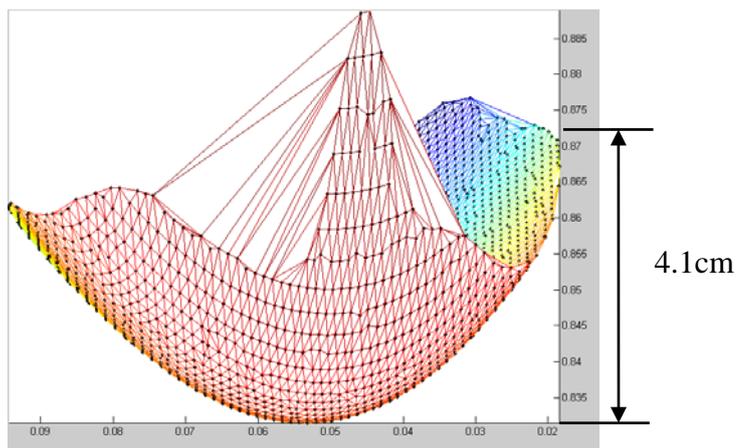


圖 5-27：模型上視圖

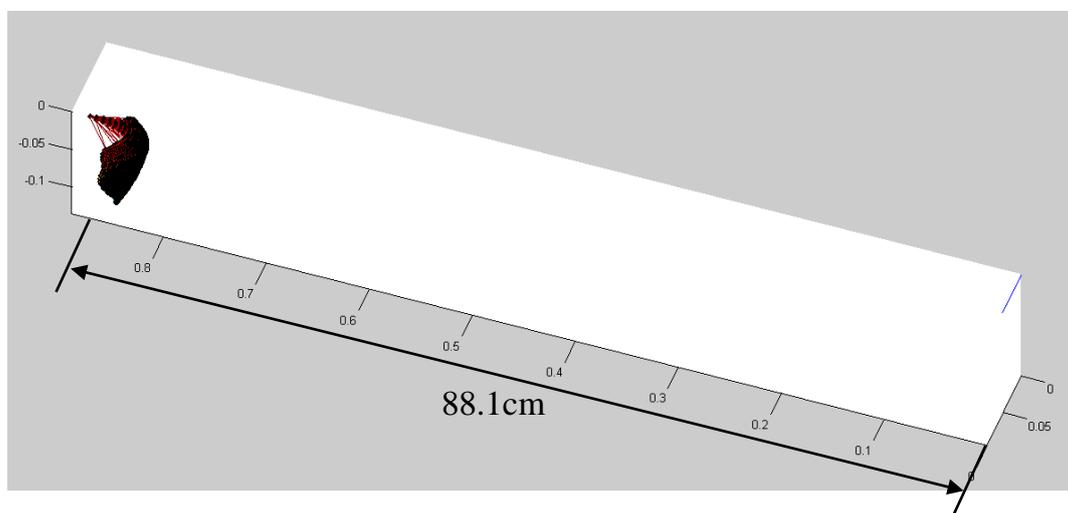


圖 5-28：模型之位置資訊(1)

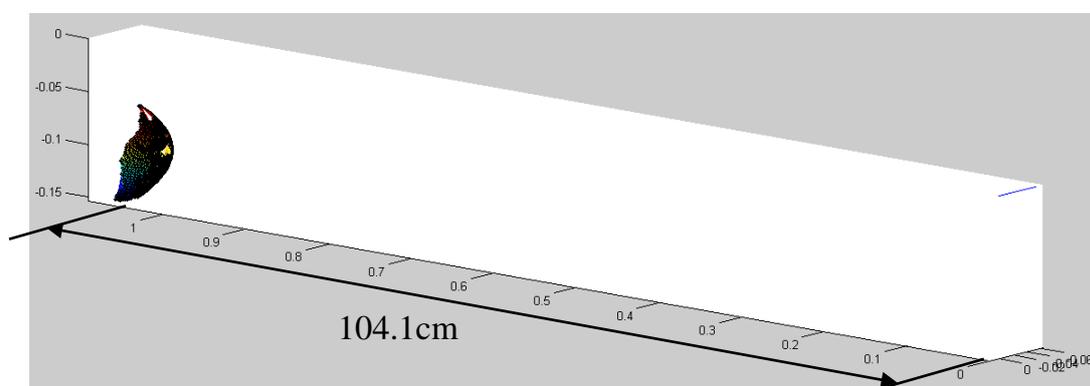


圖 5-29：模型之位置資訊(2)

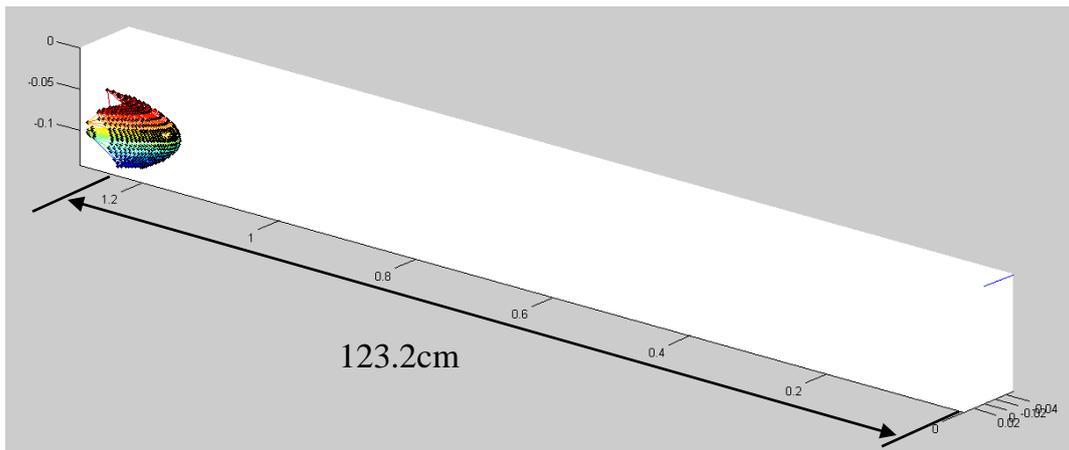


圖 5-30：模型之位置資訊(3)

5.4 三維模型重建結果分析

由重建結果證實本實驗系統能確實的重建起物件之三維模型，重建的三維模型其外觀形狀與原始物體相似，曲面之重建只能以線段進行重建故其外觀較為失真，且曲面物件之 match point 定位較簡單幾何物件要難，故其 match points 之定位有較大之誤差進而影響重建之果，對於物件尺寸重建之準確度可由表 5-1~4 所示，從表中可以看出此次重建之物件大小其邊長均小等於 10cm，重建之尺寸誤差約 0.5cm 左右，曲面物件之曲面深度誤差較大約 0.8cm，但本重建系統之重建方法為還原出空間中一點之位置，進而重建出物件之三維模型，故以重建距離 71cm 來看尺寸之誤差，其誤差比例已經很小，71cm 距離之計算誤差約為 2cm，85cm 距離之誤差約為 3cm，誤差量約佔整體之 3% 左右。

除此之外也對重建距離進行測試，將同一物件分別擺放至 85cm、100cm 與 120cm 進行重建，實驗結果顯示系統能確實的計算出物件之位置距離，其誤差均在一定的範圍內，但實驗結果也顯示隨著物件擺放的距離越遠其模型的重建效果會越差，模型會產生型變，推測造成此一現象係因

為用來計算 Fundamental matrix 之已知點均位於 70 至 90cm 這個區間，故在這個區間內模型的重建效果最好。

接著對重建所需時間進行探討，由表 5-4 所示簡單幾何物件與曲面物件重建所需時間相比較，簡單幾何物件重建所需的時間較多，曲面物件重建時間較短，主要的原因在於簡單幾何物件多了轉角偵測此影像處理步驟，承如前面所述影像處理涉及到大量的資料運算因此需花費較多的時間進行計算，簡單幾何物件左右兩影像均需進行轉角偵測，而曲面物件無須進行轉角偵測故其重建時間減少許多。

表 5-1：尺寸、距離比較表

	邊長一	邊長二	距離
簡單幾何物件一	5 cm	5 cm	71 cm
三維模型	5.3 cm	5.4 cm	73 cm
誤差	0.3 cm	0.4 cm	2 cm
	邊長一	邊長二	距離
簡單幾何物件二	5 cm	5 cm	71 cm
三維模型	5.3 cm	5.5 cm	73 cm
誤差	0.3 cm	0.5 cm	2 cm
	邊長一	邊長二	距離

簡單幾何物件三	8 cm	5.5 cm	71 cm
三維模型	8.5 cm	5 cm	73 cm
誤差	0.5 cm	0.5 cm	2 cm

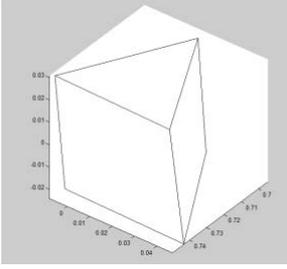
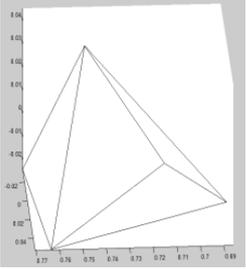
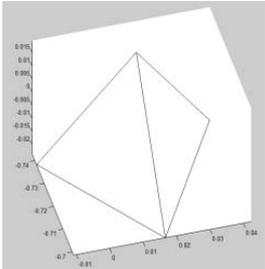
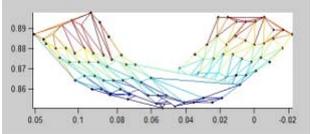
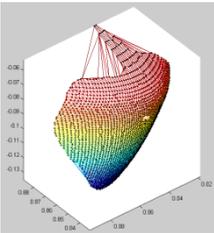
表 5-2：尺寸、距離比較表

	邊長一	邊長二	邊長三	距離
曲面物件(一)	10 cm	8 cm	3.6 cm	85 cm
三維模型	10.4 cm	8.2 cm	4.4 cm	88 cm
誤差	0.4 cm	0.2 cm	0.8 cm	3 cm
	邊長一	邊長二	邊長三	距離
曲面物件(二)	6.5 cm	7.5 cm	3.5 cm	85 cm
三維模型	7.2 cm	8.5 cm	4.1 cm	88.1 cm
誤差	0.7 cm	1 cm	0.6 cm	3.1 cm

表 5-3：距離誤差表

曲面物件(二)	距離一	距離二	距離三
實際距離	85 cm	100 cm	120 cm
重建模型距離	88.1 cm	104.1 cm	123.2 cm
誤差	3.1 cm	4.1cm	3.2cm

表 5-4：重建時間表

物件	時間	物件	時間
	6.71 second		7.47 second
	6.13 second		0.73 second
	0.91 second		

第六章 結論與未來展望

本研究希望能發展一套三維重建系統，此系統之特點為無需從外部投影一結構光至物體表面上即可重建起物件之三維模型，因此本研究以兩台 CCD 攝影機基於立體視覺法，以轉角偵測與虛擬網格之建立進一步的重建起物件之三維模型，本研究目前將重建之物件區分為簡單幾何物件以及曲面物件，針對簡單幾何物件採用轉角偵測法，藉由特徵點之擷取重建起物件之三維模型，針對曲面物件採用虛擬網格法，使用虛擬網格之建立估測物件之表面曲率重建起物件三維模型，除此之外無論簡單幾何物件或是曲面物件系統均能確實計算出物件之距離，且由於本系統無需投影一結構光至物件表面，因此重建系統不受物件表面顏色之影響，如此一來此系統即具有能力提供足夠訊息給機械手臂進行物件之夾取。

由於本系統目前將物件區分為簡單幾何與曲面物件兩種類型，其中曲面物件之重建為建立虛擬網格來估測曲面進而重建起三維模型，因此曲面之重建誤差較大且系統目前能重建之曲面物件有限，因此未來系統的改善希望能成功的將演算法進行結合，重建起更為複雜之物件之三維模型並且有更佳之適應性與強健性。

參考文獻

- [1] Q. Zhao, D. Zhang, L. Zhang and N. Luo, “High resolution partial fingerprint alignment using pore–valley descriptors,” *Pattern Recognition*, vol. 43, pp. 1050-1061, March 2010.
- [2] C. N. E. Anagnostopoulos, V. Loumos and E. Kayafas, “License Plate Recognition From Still Images and Video Sequences: A Survey,” *IEEE, Transactions on Intelligent Transportation Systems*, vol. 9, pp. 377-391, September 2008.
- [3] J. Y. Choi, Y. M. Ro and K. N. Plataniotis, “Boosting Color Feature Selection for Color Face Recognition,” *IEEE, Transactions on Image Processing*, vol. 20, pp. 1425-1434, May 2011.
- [4] W. Wei, G. Wang and H. Chen, “3D reconstruction of a femur shaft using a model and two 2D X-ray images, ” *IEEE, International Conference on Computer Science & Education*, pp. 720-722, 2009.
- [5] P. Gamage, S. Q. Xie , P. Delmas, P. Xu, “ 3D reconstruction of patient specific bone models from 2D radiographs for image guided orthopedic surgery, ” *IEEE, 2009 Digital Image Computing: Techniques and Applications*, pp. 212 – 216, 2009.
- [6] M.J. Clarkson, D. Rueckert, D.L.G. Hill and D.J. Hawkes, “ Using photo-consistency to register 2D optical images of the human face to a 3D surface model, ” *IEEE, Transactions on pattern analysis and machine intelligence*, vol. 23, no. 11, pp. 1266 -1280, 2001.

- [7] T. Hassner and R. Basri, “ Example based 3D reconstruction from single 2D images, ” *IEEE, Computer vision and pattern recognition workshop*, pp. 15 -15, 2006.
- [8] M. Strand and R. Dillmann, “ Using an attributed 2D-grid for next-best-view planning on 3D environment data for an autonomous robot, ” *IEEE, International conference on information and automation*, pp. 314-319, 2008.
- [9] G. M. Bone, A. Lambert and M. Edwards, “Automated Modeling and Robotic Grasping of Unknown Three-Dimensional Objects,” *IEEE International Conference on Robotics and Automation Pasadena*, pp. 292-298, May 2008.
- [10] T. Tasic and B. Acko, “Integration of a laser interferometer and a CMM into a measurement system for measuring internal dimensions ” *Measurement*, vol. 44, pp. 426-433, February 2011.
- [11] M. Strand and R. Dillmann, “Using an attributed 2D-grid for next-best-view planning on 3D environment data for an autonomous robot,” *IEEE, International Conference on Information and Automation*, pp. 314-319, June 2008.
- [12] M. Kimura and H. Saito, “3D reconstruction based on epipolar geometry,” *IEICE, Transactions on Information and Systems*, pp. 1690-1697, 2001.
- [13] S. Prakoonwit and R. Benjamin, “Optimal 3D surface reconstruction from multiview photographic images,” *International Conference on Cyber Worlds*, pp. 126-131, Sep 2009.

- [14] R. T. Frankot and R. Chellappa, “A Method for Enforcing Integrability in Shape from Shading Algorithms,” *IEEE, Transactions on Pattern Analysis and Machine Intelligence*, vol. 10, pp. 439-451, Jul 1988.
- [15] D.G. Aliaga and Y. Xu, “A Self-Calibrating Method for Photogeometric Acquisition of 3D Objects,” *IEEE, Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, pp. 747-754, April 2010.
- [16] P. Lavoie, D. Ionescu and E.M. Petriu, “3-D Object Model Recovery From 2-D Images Using Structured Light,” *IEEE, Transactions on Instrumentation and Measurement*, vol. 53 pp. 377- 382, 2004.
- [17] L.M. Song and D.N. Wang, “A novel grating matching method for 3D reconstruction,” *NDT & E International*, vol. 39, pp. 282-288, Jun 2006.
- [18] J. D. Zheng, L. Y. Zhang, X. Y. Du and Z. A. Ding, “3D curve structure reconstruction from a sparse set of unordered images,” *Computers in Industry*, vol. 60, pp. 126-134, Feb 2009.
- [19] J. Canny, “A computational approach to edge detection, ” *IEEE, Transaction on Pattern Analysis and Machine Intelligence*, vol. PAMI-8, pp. 679-698, Nov 1986.
- [20] C. Harris and M. Stephens, “A combined corner and edge detector, ” *Alvey Vision Conference*, pp. 147-152, 1988.
- [21] Z. Wang and B. Boufama, “Using stereo geometry towards accurate 3D reconstruction, ” *IEEE International Conference on Electro/Information Technology*, pp. 134-140, June 2009.

- [22] G. L. Mariottini and D. Prattichizzo, “EGT for multiple view geometry and visual servoing: robotics vision with pinhole and panoramic cameras,” *IEEE, Robotics & Automation Magazine*, vol. 12, pp. 26-39, Dec 2005.
- [23] P. K. Ho and R. Chung, “Stereo-motion with stereo and motion in complement,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, pp. 215-220, Feb 2000.
- [24] 洪維恩 編著，「Matlab 7 程式設計」，旗標出版股份有限公司，中華民國95年。
- [25] 繆紹綱 編著，「數位影像處理：運用matlab」，台灣東華書局股份有限公司，中華民國94年。
- [26] http://www.vision.caltech.edu/bouguetj/calib_doc/index.html---2010.07.09