

# 國立交通大學

電機與控制工程研究所

## 碩士論文

利用深度資訊對複雜場景中的三維物體  
進行分割與辨識

3D Object Segmentation and Recognition  
in Cluttered Scene Based on Range Data

研究生：徐煜維

指導教授：林昇甫 教授

中華民國九十八年七月

利用深度資訊對複雜場景中的三維物體進行分  
割與辨識

3D Object Segmentation and Recognition  
in Cluttered Scene Based on Range Data

研 究 生：徐煜維

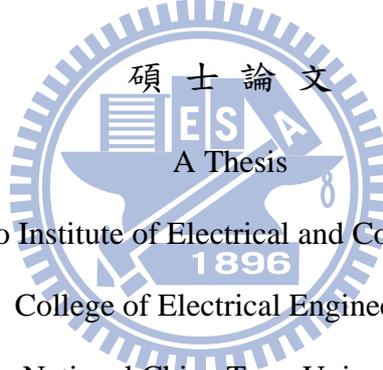
Student：Yu-Wei Hsu

指 導 教 授：林昇甫

Advisor：Sheng-Fuu Lin

國 立 交 通 大 學

電機與控制工程研究所



Submitted to Institute of Electrical and Control Engineering  
College of Electrical Engineering  
National Chiao Tung University

in partial Fulfillment of the Requirements

for the Degree of

Master

in

Electrical and Control Engineering

July 2009

Hsinchu, Taiwan, Republic of China

中華民國九十八年七月

# 利用深度資訊對複雜場景中的三維物體 進行分割與辨識

學生：徐煜維

指導教授：林昇甫 博士

國立交通大學電機與控制工程學系 碩士班

## 摘要

本論文提出一個快速且有效的三維物體辨識系統，以辨識複雜場景下的三維物體。此系統可解決深度資訊(range data)中，因為測量誤差所產生的雜訊，同時提高物體被遮蔽時的辨識率，未來可將此系統應用在機器人視覺上，來進行辨識與引導動作。

首先，本論文結合了適應性中值濾波器(adaptive median filter)與移動式最小平方法(moving least square)來修復因測量誤差產生的三維雜訊，以獲得正確的物體部分表面，並且提出了一個多重臨界值演算法(multilevel thresholding)，使得複雜場景變成多個單一場景，再利用深度影像中像素的連通性，來分離每個單一場景中的物體，以作為辨識之輸入。

其次，為了使得目標物體在遮蔽環境下，也可以有效地被辨識出來，本論文使用了邊緣圖像(edge map)的概念，使單一物體依照其表面變化，被分割成許多不同的封閉區塊，以提高目標物體被遮蔽時的辨識率。首先，利用 Canny 邊緣偵測器(Canny edge detector)來偵測出深度影像中物體的步階邊緣，然後本論文提出了計算物體表面法向量變化以形成梯度影像，再偵測出物體的屋脊邊緣(roof edge)，就形成了邊緣影像(edge image)，最後對邊緣影像使用形態學運算(morphological operator)，使邊緣影像變成邊緣圖像。

然後，對每個物體的邊緣圖像中的每個封閉區塊使用區域成長法(region growing)來抽出該區塊的特徵後，並使用多維直方圖(multidimensional histogram)統計整體特徵與區塊特徵，形成了整體直方圖(unity histogram)與部分直方圖(partial histogram)；其中，在本論文中，使用曲率之形狀指標(shape index)及法向量分量之夾角，這兩個區域特徵來表示三維物體部分表面之特徵。

最後使用  $\chi^2$ -divergence 計算直方圖相異程度，並且結合了幾個常用的直方圖比對方法，以計算部分直方圖的相異程度，同時提出了兩階段的辨識系統來縮短辨識所需的時間。

# 3D Object Segmentation and Recognition in Cluttered Scene Based on Range Data

Student : Yu-Wei Hsu

Advisor: Dr. Sheng-Fuu Lin

Institute of Electric and Control Engineering  
National Chiao Tung University

## ABSTRACT

In this thesis, a highly efficient 3D view-based object recognition system, which is to recognize 3D objects in cluttered scenes, is proposed. This system can handle the 3D noise in the range data because of measure error margin in the range finder, and increase the recognition accuracy when object is covered in cluttered scene. In the future, I hope that this recognition system will apply to the robotic vision.

First of all, in order to handle the 3D noise in the range data, an algorithm which combines adaptive median filter and moving least square (MLS) is proposed in the beginning of the recognition system. After that, a multilevel thresholding method is proposed which segments a cluttered scene into several monotonous scenes, and then separates each object in the scene by using the 8-connected component of pixels in range image. These objects will be the input of the recognition system.

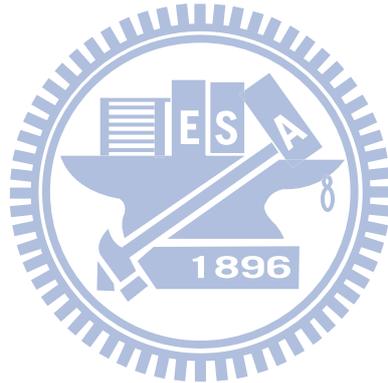
More importantly, in order to recognize objects which are covered in cluttered scene, a concept of edge map is applied in this thesis. Then, extract the feature belongs to each closed region in the edge map by using region growing method, and calculate the features to create unity histogram and partial histogram by using multidimensional histogram; moreover, the local feature is presented as features of 3D object's surface; however, in order to increase the speed during the recognition, a two-step recognition system is presented in this thesis.

## 誌 謝

首先，我要感謝我的母親，提供給我衣食無缺的環境，讓我可以有機會在交通大學就讀碩士班，其次我要感謝我的姐姐，讓我有學習的榜樣，可以一同競爭、一同進步，再來我要感謝我的指導教授林昇甫博士，在這兩年就讀研究所期間，不只在研究上，也在做人處事上，給與我許多的啟蒙與啟發。

再來，我要感謝我的實驗室學長們，晉嘉、國育、士哲、弦澤、啟耀以及實驗室的同學們、學弟們在研究上給我許多的指導與幫助，由於他們熱心的協助我，並時常跟我討論研究內容，所以才有這篇論文的誕生，。

最後，我要感謝我身邊的所有人，由於他們的支持與鼓勵，才得以讓我順利完成碩士學位，致上我最誠摯的感謝，。



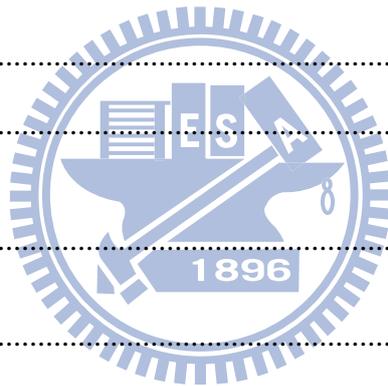
# 目 錄

中文摘要	i
英文摘要	ii
銘謝	iii
目錄	iv
圖目錄	v
表目錄	vi
第一章 緒論 .....	1
1.1 三維辨識介紹 .....	1
1.2 研究動機與背景 .....	2
1.3 相關研究之探討 .....	3
1.3.1 深度影像分割 .....	4
1.3.2 三維物體辨識 .....	5
1.4 論文主體之貢獻 .....	7
第二章 相關技術原理 .....	9
2.1 深度資訊處理 .....	9
2.1.1 三維深度資訊 .....	9
2.1.2 法向量計算 .....	11
2.2 三維雜訊處理 .....	12
2.2.1 適應性中值濾波器 .....	12
2.2.2 移動式最小平方法 .....	14

2.3 複雜場景分割 .....	19
2.3.1 多重臨界值法 .....	19
2.3.2 連通量分析 .....	22
2.4 邊緣資訊處理 .....	23
2.4.1 邊緣偵測 .....	24
2.4.2 法向量之梯度變化 .....	27
2.4.3 形態學運算 .....	31
2.4.3 區域成長法 .....	33
2.5 區域特徵抽取 .....	35
2.5.1 表面之深度變化 .....	35
2.5.2 法向量分量之夾角 .....	36
2.5.3 曲率之形狀指標 .....	37
2.6 區域特徵比對 .....	42
2.6.1 多維度直方圖 .....	43
2.6.2 直方圖比對 .....	45
<b>第三章 系統流程說明 .....</b>	<b>48</b>
3.1 辨識系統架構 .....	48
3.2 三維雜訊處理 .....	49
3.2.1 雜訊前處理 .....	49
3.2.2 雜訊後處理 .....	50
3.3 深度影像切割 .....	52
3.3.1 場景分離 .....	53
3.3.2 物體分離 .....	55
3.4 封閉邊緣處理 .....	57
3.4.1 邊緣影像 .....	58
3.4.2 邊緣圖像 .....	60



3.5 三維特徵抽取 .....	61
3.5.1 整體特徵統計 .....	62
3.5.2 區塊特徵統計 .....	64
3.6 辨識系統介紹 .....	65
3.6.1 整體直方圖比對 .....	66
3.6.2 部分直方圖比對 .....	67
<b>第四章 實驗結果與討論 .....</b>	<b>68</b>
4.1 實驗設備與效能評估 .....	68
4.1.1 系統實驗設備 .....	68
4.1.2 系統效能評估 .....	94
4.2 實驗結果 .....	95
4.3 結果分析 .....	103
<b>第五章 總結與未來方向 .....</b>	<b>106</b>
<b>參考文獻 .....</b>	<b>108</b>



## 表目錄

表 4-1 資料庫中的物體 .....	69
表 4-2 已知物體旋轉角度與姿態對照表 .....	70
表 4-3 未知物體的角度範圍與姿態對照表 .....	71
表 4-4 未知物體的測試角度與該角度所對應的姿態 .....	88
表 4-5 實驗組別與實驗變因之關係 .....	95
表 4-6 各方法與各種樣本區間個數列表 .....	96
表 4-7 輸入的測試資料為組別 1 之各種方法的辨識率比較 .....	97
表 4-8 輸入的測試資料為組別 2 之各種方法的辨識率比較 .....	97
表 4-9 輸入的測試資料為組別 3 之各種方法的辨識率比較 .....	97
表 4-10 輸入的測試資料為組別 4 之各種方法的辨識率比較 .....	97



## 圖目錄

圖 2-1 強度影像與深度影像之差異 (a)強度影像 (b)深度影像 .....	10
圖 2-2 深度資訊的擷取方式之示意圖 .....	11
圖 2-3 深度資訊顯示於三維座標中之顯示圖 .....	11
圖 2-4 3×3遮罩所對應的係數及點 (a) 所對應的係數 (b)所對應的點 .....	12
圖 2-5 3×3遮罩所對應的係數及值(a) 遮罩係數 (b) 遮罩下所對應的值 .....	13
圖 2-6 移動式最小平方法流程	
(a)某點集合之代表點 $\mathbf{p}_k$ (b)重新取樣的點集合之代表點 $\mathbf{r}_k$	
(c) $\mathbf{r}_k$ 投影在原本曲面上之投影點 $\mathbf{q}_k$ (d) $S_p$ 為近似原本曲面之曲面 ....	16
圖 2-7 移動式最小平方法之原理 .....	16
圖 2-8 投影點的區域座標 .....	18
圖 2-9 深度影像與深度統計後的直方圖 (a)深度影像 (b)直方圖 .....	20
圖 2-10 不同區間個數的直方圖與臨界值位置	
(a) $b = 16$ (b) $b = 32$ (c) $b = 64$ (d) $b = 128$ .....	20
圖 2-11 深度影像之單一場景 (a)場景一(b)場景二(c)場景三(d)場景四 .....	22
圖 2-12 不同的鄰接方式 (a) 4-鄰接 $N_4(P)$ (b) 8-鄰接 $N_8(P)$ .....	23
圖 2-13 不同的連通方式 (a) 三個 4 連通成份 (b) 兩個 8 連通成份 .....	23
圖 2-14 影像邊緣之灰階與一階導數之變化	
(a)具明顯邊緣影像 (b)水平灰階變化 (c)一階導數變化 .....	24
圖 2-15 Sobel 運算子 (a)計算 $G_x$ (b)計算 $G_y$ .....	25
圖 2-16 高斯濾波器遮罩 .....	26
圖 2-17 數位影像使用不同邊緣偵測器之差異	
(a)數位影像 (b)使用 Sobel 邊緣偵測器 (c)使用 Canny 邊緣偵測器 .....	27
圖 2-18 藉由深度資訊所測量到物體之邊緣	
(a)真實物體 (b)深度影像 (c)步階邊緣 (d)屋脊邊緣 .....	28

圖 2-19	3 × 3 的遮罩內的係數與對應的法向量 (a)係數 (b)法向量	29
圖 2-20	法向量的梯度變化	30
圖 2-21	不同的 $\gamma$ 值使得屋脊邊緣細節改變 (a) $\gamma = 0.5$ (b) $\gamma = 2$ (c) $\gamma = 10$	30
圖 2-22	二值影像與結構元素 (a)原本二值影像 <b>A</b> (b)結構元素 <b>B</b>	31
圖 2-23	二值影像的膨脹過程與結果	
	(a)二值影像進行膨脹(b)膨脹後的二值影像	32
圖 2-24	二值影像的侵蝕過程與結果	
	(a)二值影像進行侵蝕 (b)侵蝕後的二值影像	32
圖 2-25	閉合後的二值影像	33
圖 2-26	收縮 n 次的二值影像 (a) n=20 (b) n=50	33
圖 2-27	不同視角下的深度影像 (a)原本的視角 (b)後來的視角	36
圖 2-28	法向量分量之夾角	37
圖 2-29	物體的區域特徵 (a) $\Phi = [\Phi_{ij}]_{m \times n}$ (b) $\Theta = [\Theta_{ij}]_{m \times n}$	37
圖 2-30	二維曲率的定義之示意圖	38
圖 2-31	馬鞍面的主曲率計算之示意圖	39
圖 2-32	3 × 3 的遮罩內的係數與對應的點集合 (a)係數 (b)點集合	41
圖 2-33	計算曲率方式之示意圖 (a)側面投影圖 (b)三維空間投影圖	41
圖 2-34	曲面形狀指標	42
圖 2-35	不同影像的直方圖 (a)強度影像的直方圖 (b)深度影像的直方圖	43
圖 2-36	強度影像的整體直方圖與部分直方圖關係	
	(a)強度影像 (b)整體直方圖為部分直方圖之總和	45
圖 3-1	實際拍攝深度資訊之整體系統架構	48
圖 3-2	深度資訊之雜訊處理架構	49
圖 3-3	深度影像 (a)椅子 (b)電腦	51
圖 3-4	去除雜訊前的法向量之 z 方向分量 (a)椅子 (b)電腦	52
圖 3-5	去除雜訊後的法向量之 z 方向分量 (a)椅子 (b)電腦	52

圖 3-6 影像切割架構 .....	53
圖 3-7 $T=6$ 時，由真實場景中所分離的物體 (a)真實場景 (b)物體 1 (c)物體 2 (d)物體 3 (e)物體 4 (f)物體 5 (g)物體 6 (h)物體 7.....	54
圖 3-8 $T_h=0.01$ 時，由真實場景中所分離的物體 (a)物體 1 (b)物體 2 (c)物體 3 (d)物體 4 (e)物體 5 (f)物體 6 (g)物體 7 (h)物體 8 (i)物體 9 (j)物體 10 ...	56
圖 3-9 $T_h=0.05$ 時，由真實場景中所分離的物體 (a)物體 1 (b)物體 2 (c)物體 3 (d)物體 4 (e)物體 5 .....	57
圖 3-10 邊緣處理架構 .....	58
圖 3-11 物體的步階邊緣影像 (a)椅子 (b)電腦 .....	59
圖 3-12 物體的屋脊邊緣影像 (a)椅子 (b)電腦 .....	59
圖 3-13 物體的邊緣影像 (a)椅子 (b)電腦 .....	60
圖 3-14 形態學閉合後的邊緣影像 (a)椅子 (b)電腦 .....	60
圖 3-15 形態學收縮後的邊緣影像 (a)椅子 (b)電腦 .....	61
圖 3-16 物體的邊緣圖像 (a)椅子 (b)電腦 .....	61
圖 3-17 特徵抽取架構 .....	62
圖 3-18 物體的整體區域特徵影像 $D$ (a)椅子 (b)電腦 .....	63
圖 3-19 物體的整體區域特徵影像 $\Phi$ (a)椅子 (b)電腦 .....	63
圖 3-20 物體的整體區域特徵影像 $\theta$ (a)椅子 (b)電腦 .....	63
圖 3-21 物體的整體區域特徵影像 $S_1$ (a)椅子 (b)電腦 .....	64
圖 3-22 物體的區塊圖像 (a)椅子 (b)電腦 .....	64
圖 3-23 辨識系統架構 .....	66
圖 4-1 雷射測距儀 .....	69
圖 4-2 資料庫中的八個物體 .....	71
圖 4-3 CRT 電腦螢幕的強度影像與 39 張姿態的深度影像 (a)俯角 45 度，z 軸固定 (b)仰角 45 度，z 軸固定 (c)水平 0 度，z 軸固定 (d)水平 0 度，x 軸固定 (e)水平 0 度，y 軸固定 .....	73
圖 4-4 玩具人偶的強度影像與 39 張姿態的深度影像	

(a)俯角 45 度，z 軸 e 固定 (b)仰角 45 度，z 軸固定 (c)水平 0 度，z 軸固定	
(d)水平 0 度，x 軸固定 (e)水平 0 度，y 軸固定 .....	75
圖 4-5 摺疊椅的強度影像與 39 張姿態的深度影像	
(a)俯角 45 度，z 軸固定 (b)仰角 45 度，z 軸固定 (c)水平 0 度，z 軸固定	
(d)水平 0 度，x 軸固定 (e)水平 0 度，y 軸固定 .....	77
圖 4-6 桌子(無抽屜)的強度影像與 39 張姿態的深度影像	
(a)俯角 45 度，z 軸固定 (b)仰角 45 度，z 軸固定 (c)水平 0 度，z 軸固定	
(d)水平 0 度，x 軸固定 (e)水平 0 度，y 軸固定 .....	79
圖 4-7 檯燈的強度影像與 39 張姿態的深度影像	
(a)俯角 45 度，z 軸固定 (b)仰角 45 度，z 軸固定 (c)水平 0 度，z 軸固定	
(d)水平 0 度，x 軸固定 (e)水平 0 度，y 軸固定 .....	81
圖 4-8 LCD 電腦螢幕的強度影像與 39 張姿態的深度影像	
(a)俯角 45 度，z 軸固定 (b)仰角 45 度，z 軸固定 (c)水平 0 度，z 軸固定	
(d)水平 0 度，x 軸固定 (e)水平 0 度，y 軸固定 .....	83
圖 4-9 旋轉椅的強度影像與 39 張姿態的深度影像	
(a)俯角 45 度，z 軸固定 (b)仰角 45 度，z 軸固定 (c)水平 0 度，z 軸固定	
(d)水平 0 度，x 軸固定 (e)水平 0 度，y 軸固定 .....	85
圖 4-10 桌子(有抽屜)的強度影像與 39 張姿態的深度影像	
(a)俯角 45 度，z 軸固定 (b)仰角 45 度，z 軸固定 (c)水平 0 度，z 軸固定	
(d)水平 0 度，x 軸固定 (e)水平 0 度，y 軸固定 .....	87
圖 4-11 資料庫中的各別單一物體的深度資訊顯示 .....	
	88
圖 4-12 未遮蔽之場景 1 的強度影像 .....	
	89
圖 4-13 未遮蔽之場景 1 的強度影像 .....	
	89
圖 4-14 遮蔽之場景 1 的強度影像 .....	
	89
圖 4-15 遮蔽之場景 1 的強度影像 .....	
	89
圖 4-16 未遮蔽狀況下之場景 1 的不同姿態之深度影像 .....	
	90
圖 4-17 未遮蔽狀況下之場景 2 的不同姿態之深度影像 .....	
	91
圖 4-18 遮蔽狀況下之場景 1 的不同姿態之深度影像 .....	
	92

圖 4-19 遮蔽狀況下之場景 2 的不同姿態之深度影像 .....	93
圖 4-20 組別 1 與組別 2 的各種方法之物體辨識率比較 .....	98
圖 4-21 組別 1 與組別 2 的各種方法之姿態辨識率比較 .....	98
圖 4-22 組別 1 與組別 2 的各種方法之前三名物體辨識率比較 .....	98
圖 4-23 組別 3 與組別 4 的各種方法之物體辨識率比較 .....	99
圖 4-24 組別 3 與組別 4 的各種方法之姿態辨識率比較 .....	99
圖 4-25 組別 3 與組別 4 的各種方法之前三名物體辨識率比較 .....	100
圖 4-26 組別 1 與組別 3 的各種方法之物體辨識率比較 .....	101
圖 4-27 組別 1 與組別 3 的各種方法之姿態辨識率比較 .....	101
圖 4-28 組別 1 與組別 3 的各種方法之前三名物體辨識率比較 .....	101
圖 4-29 組別 2 與組別 4 的各種方法之物體辨識率比較 .....	102
圖 4-30 組別 2 與組別 4 的各種方法之姿態辨識率比較 .....	102
圖 4-31 組別 2 與組別 4 的各種方法之前三名物體辨識率比較 .....	103
圖 5-1 雷射測距儀所測得的大型複雜深度影像 .....	107



# 第一章 緒論

本章分成四節，將三維物體辨識這個領域作詳盡的介紹。首先，1.1節將介紹三維物體辨識(3D object recognition)相對於以往二維影像辨識的優點，然後說明三維物體資料的取得方式以及三維物體辨識的未來發展潛力；1.2節將介紹研究三維物體辨識動機，以及如何將三維物體辨識應用到機器人視覺中；1.3節將說明在三維物體辨識中，切割複雜場景中的三維物體之演算法以及抽取三維物體特徵之演算法；1.4節將說明本論文在三維物體辨識的貢獻，提出一個三維物體切割與辨識系統以處理複雜場景的多物體辨識，並且在不犧牲計算速度的前提下，提高物體的辨識率。

## 1.1 三維辨識介紹

近十年以來，由於取像技術的發展是取得二維影像為主，例如：灰階影像或是彩色影像，使得影像辨識大多是從二維資訊中抽取特徵並發展其辨識演算法，像是文字辨識、人臉辨識、指紋辨識、車牌辨識、生物特徵辨識等方面，但限於二維影像提供的資訊會因為光源照射角度的差異以及光源強度的不同，使得辨識的效果不好或是無法辨識出影像中的目標物體；然而，近年來隨著擷取物體三維資訊的取像技術越來越快速及準確，使得以物體表面的三維資訊為基礎來進行物體辨識，已經成為電腦視覺及圖形識別等領域中，一個極具發展潛力的研究主題。

一般二維影像只能提供以真實世界投影在影像平面上的資訊，基於物體表面的反射特性，來獲得物體的輪廓、邊界、紋理等特徵，但是以二維資訊為基礎的辨識演算法，來辨識出實際的三維物體時，會產生許多問題，而大多的問題是因為缺乏資訊而產生的，因為光源照射角度的差異以及光源強度的不同，以致於針對同一場景，會獲得不同的二維資訊，使得影像辨識複雜場景的辨識率偏低；然而，相較之下，如果能夠直接取得場景的三維資料來進行物體的辨識，相對於二維影像所提供的資訊，三維資訊更能準確地描述複雜場景中每個物體的外觀，並且提供更多辨識用的資訊，也具有更佳處理複雜場

景的辨識能力。

然而，要進行三維物體辨識(3D object recognition)，首先要取得三維物體表面的資訊。三維物體表面資訊的取得，根據許多研究取得資料的方式，大部分都經由雷射測距儀(laser range finder)所取得的深度資訊(range data)為主[1][2][3]。雷射測距的技術廣泛地被使用在許多不同的領域，像是距離測量、地形地貌的監測，建築物的測量、三維立體物體的測量等。藉由三維雷射測距儀獲得的深度資訊，可以得知物體表面完整的三維空間分布資訊，也可獲得物體的輪廓、邊界、紋理等；這些影像資訊可以廣泛地應用於機器人視覺[4]，地形變化的測量，機械精密定位以及生物立體影像的重建等；但是在過去這些裝置不僅昂貴且體積過大以致於不易攜帶，所以取得三維立體影像技術十分不普及化；然而，隨著近年來三維取像儀器的成本降低以及精確度大幅提高，故直接利用雷射測距儀來獲得物體的三維資訊，快速且便利許多，相較於以往是利用立體視覺影像產生的三維資訊，三維雷射測距儀的影像，不需要經過複雜的三維重建過程，而且獲得的量測結果相當準確且快速。現今，由於三維雷射測距儀的越來越普及化，藉由三維雷射測距儀獲得物體表面的三維資訊分布，並利用此資訊來提升三維影像識別技術的可能性大幅提升。

## 1.2 研究動機與背景

在現實生活中，自動化辨識系統應用在居家保全、安全監控以及自動化設備等方面的需求性日益增加；然而，目前的自動化辨識系統仍是以二維影像為基礎來進行辨識為主，應用在人臉辨識、車牌辨識、指紋辨識、生物辨識等，而這些辨識系統應用單一物體的辨識上，均有蠻高的辨識率，主要原因是因為這些辨識系統皆處在良好的環境下來進行辨識，例如外來光源充足的環境中或是被辨識之物體沒被遮蔽的前提下，才可以有這麼高的辨識率，但是這些以二維影像為基礎的辨識系統，若應用在機器人視覺中[4]，辨識率卻很低，主要原因是這些辨識系統無法處理場景中的物體彼此有互相遮蔽之情形或是目標物體處於光源不足的環境中。

機器人視覺是希望用計算機來模擬出跟人一樣的視覺系統，並且藉由儀器來取得真實空間中的三維物體之資訊，進行資訊處理並且加以理解，最後可以用於機器人上。相對於以往的計算機視覺，計算機視覺只要求辨識率高，準確度高，然後可以應用自動化設備的機器上，速度並不是主要的考量，故計算機視覺大多是二維影像辨識系統為基礎，再進一步地應用到其他方面；然而，在機器人視覺中，必須要求計算速度快且可辨識出處於惡劣環境中的物體，因為機器人必須要視覺引導下立即作出判斷及動作，並且取代人類於惡劣的環境下的工作，故執行速度與辨識能力是主要考量。

而人類的視覺系統就是將真實的三維世界中的場景投影到視網膜上，類似投影到一個二維陣列中，陣列中所包含的資訊有真實空間中物體的相對位置，其類似由雷射測距儀所獲得的深度資訊，以及因為外來光源的變化，物體所表現的顏色、輪廓等資訊，其類似由相機所取得的 RGB 二維影像；而人類可以辨識出真實世界中的三維物體，除了利用物體的二維資訊外，也需要利用物體的三維資訊，像是物體的表面形狀變化，才得以辨識出該物體，在機器人視覺中就包括這兩個資訊的處理與辨識，但由於使用二維資訊為基礎的辨識系統，已經有很高的辨識率，然而使用三維資訊進行三維物體辨識仍在發展階段，故本論文希望設計出一個針對三維資訊進行辨識的辨識系統，並且改進以往三維辨識系統的缺點，使得此辨識系統更為優良，並且未來進一步地應用在機器人視覺上。

在本論文中，將設計出一個三維物體辨識系統，可以應用在機器人視覺上，針對一個複雜場景，可以同時將一個複雜場景分割成很多單一場景，並從單一場景中分離不同物體，然後平行進行不同物體之辨識。在只能取得物體之部分表面的情形下，並且考慮當物體被遮蔽的情形發生，此辨識系統仍然有很高的辨識率；此外，即使因測量視角的改變造成物體的姿態略為改變，此辨識系統依舊可以辨識出場景中的物體。

### 1.3 相關研究之探討

在本論文中，將對複雜場景的深度影像進行分割，分割成許多不同深度範圍的單一

場景，然後再從個別單一場景分離出單一物體；然而，為了使每個單一物體在被遮蔽的情況下，可以有效地被辨識出來，將會對每個單一物體進行分割，使得單一物體可以分割成許多封閉的區塊；如此一來即使物體有部分的面積被遮蔽，仍可有效地辨識出該物體，故本論文將會在 1.3.1 節介紹關於深度影像切割的相關研究。

在本論文中，要辨識出複雜場景中的未知物體對應資料庫中哪個物體時，將比對兩物體表面的特徵是否相符，然而，依照抽取物體表面特徵的方式，可以將特徵分為全域特徵(global structural feature)、區域特徵(local feature)以及半區域特徵(semi-local feature)三類，依照不同的表面特徵抽取方式，會發展出不同的三維物體辨識演算法。故本論文將會在 1.3.2 節介紹關於利用深度資訊來抽取特徵以進行三維物體辨識的相關研究。

### 1.3.1 深度影像分割

深度影像分割的主要目的，是將複雜場景中的多個物體各自獨立分割出來，作為三維物體辨識系統的輸入。在深度影像的分割演算法的研究中，深度影像分割可依照分割方法的不同[5]，分成兩類：以區域為基礎的分割演算法(region-based)與以邊緣為基礎的分割演算法(edge-based)。區域基礎的深度影像切割演算法都是以深度影像中的像素的相似性來進行分割，主要根據應用影像分割之原理於深度影像中，例如：區域成長法(region growing)[5]、臨界值法(thresholding)、區域合併與分裂(merge and split)[6]、分群法(clustering)[5][7][8][9]，然而以區域為基礎的分割演算法中，若是預先設定好的幾何曲面特性有無來分割深度影像，則可分為兩類：參數模型基礎(parametric model-based)[10]以及區域成長基礎(region-growing based)[5][6]。以參數模型為基礎的分割演算法，主要是在分割深度影像前，設定好一連串的參數化的曲面模型，像是平面、雙二次曲面、圓柱面等，然後將深度資訊中的點開始進行區域成長，其中這些點將視為曲面上的點來進行成長，最後這些點可以成長為一個原本設定的曲面，缺點是在於需要假設好未知物體可能形狀的曲面模型；此外，當物體可能是有許多複雜曲面所組成的，要決定像素屬於哪種曲面，是一個比較大的問題；而以區域成長為基礎的分割演算法，主要是根據預定

的準則或是策略，使得影像可以被分割不同的區域，缺點是分割區域的過程是以區域為基本單位，故分割出來的區域邊界往往有可能跟原本物體的邊界不能重疊；此外，當遇到不同的分割問題時，就需設定不同的準則來分割影像。

以邊緣基礎的深度影像切割演算法，主要利用深度影像深度值變化的不連續或是深度資訊中取每點曲率變化的不連續性，來分割深度影像，主要的方法為邊緣偵測(edge detection)[11]，將原本影像處理中的偵測強度影像邊緣的方法，應用到深度影像或是深度資訊上，像是利用梯度運算子、灰階矩量搜尋法、鏈狀輪廓追蹤法、空間矩量次像素搜尋法等去找尋深度影像中深度值為不連續的點，而以找尋邊緣為基礎的切割演算法的優點是計算速度快並且在不同的區域交界處可產生明確的邊界，缺點是對於曲面的邊界處較為連續時，會無法偵測出其邊界，以及不一定會產生出封閉的邊界以隔離不同的區域。

本論文希望設計出一個強健性的三維物體分割系統應用到機器人視覺上，可以有效且快速分割場景與物體，所以將使用尋找邊緣為基礎的切割演算法來分割深度影像，並且改良切割演算法，使得針對連續變化的曲面，也可偵測出其邊緣所在。

### 1.3.2 三維物體辨識

三維物體辨識的主要目的，是以物體的三維資訊為基礎來進行辨識，將複雜場景中的未知物體，與資料庫中的已知物體進行比對，進而辨識出該未知物體。主要的比對方式是根據物體表面的特徵是否相符，以決定兩物體是否相同。首先，將由雷射測距儀測量到一堆空間點資料作為輸入，這些點資料來表示複雜場景中的未知物體在某姿態下之表面，然後利用演算法找出屬於該物體的點，並且抽取出該物體的表面特徵，作為代表此三維物體的主要特徵。在許多三維物體辨識演算法的研究中，主要可依照表示特徵方式的不同，將特徵分為全域特徵(global feature)、區域特徵(local feature)以及半區域特徵(semi-local feature)三類。

以全域結構特徵作為特徵表示的辨識演算法，像是M. Kazhdan、F. Stein等人所提出

全域化球諧波特徵(global spherical harmonic feature)[12][13]，主要以球諧波函數(spherical harmonic function)為基礎，不考慮物體表面上的點與鄰近點之相對關係，只考慮整體形狀，將物體的表面上的點用體素(voxel)堆積方式來表示。之後利用類似傅利葉轉換法(Fourier transform)，用一組正交函數去近似每個有著固定半徑的球面其所涵蓋的體素數量。這樣表示法的優點是考慮到物體部分表面的整體形狀，但是會因為測量視角的些微差異，使得物體的部分表面的整體形狀會有很大的差異，會受到物體姿態變化的影響很大。

以半區域特徵作為特徵表示方式的辨識演算法，像是A. E. Johnson等人提出的旋轉影像(spin image)[14][17]，相對於以全域特徵的特徵表示法，半區域特徵在特徵的描述上更具有彈性。因為只要物體表面上的點與鄰近點之間相對位置沒有改變，即使對於不同尺寸的同樣物體，也會具有相同的特徵。有不少辨識法與應用[15][16]，採用半區域特徵抽取法作為描述物體表面的基礎。其中以旋轉影像來描述三維物體，是最具代表性的方法。旋轉影像是由Johnson等人所提出，是一種與目標物體姿態無關的特徵表示法。利用許多方向點(oriented point)作為基準點，來計算物體整體的表面特徵。由於旋轉影像是使用局部座標，對該物體表面上的所有點來進行編碼，因此取得的特徵，不會因為物體在三維空間中姿態的不同而改變。也就是說，旋轉影像的特徵是與姿態無關的特徵；但是，由於旋轉影像是針對物體表面上每個頂點都做特徵描述，最後每個點都對應一張旋轉影像，故資料量相對地十分龐大，也造成計算上龐大負擔，大多應用在電腦視覺中的三維物體檢索；此外，深度資訊只能獲得物體部分表面資訊，即物體在某視角下的表面，因此半區域特徵不太適合應用基於深度資訊的三維物體辨識上；此外，使用旋轉影像作為物體部分表面的特徵描述，有許多的限制與參數，故本論文不使用半區域特徵作為三維物體辨識的特徵。

以區域特徵作為特徵表示方式的辨識演算法，例如：C. Dorai、R. Jarvis以及 J. Thirion所提出的shape index[18][19][20]，是以物體表面變化作為抽取特徵的基礎(shape-based)，將一個三維物體之組成採用以自由曲面(free form)的方式來描述，能真實的描述物體的表面組成，然後抽取出物體表面的幾何特徵，作為特徵比對用，例如：物體表面之曲率。

此幾何特徵跟因為物體的大小無關，以及受到物體姿態變化的影響較小，故即使測量視角些微改變，但是仍測量到物體相同表面時，代表此物體的表面特徵並不會有太大改變；然而，此演算法只能針對單一物體的辨識有很好的效果，對於處於在雜亂場景中的目標物體或是物體的表面以非連續方式來組成，就無法提供良好的辨識能力。

本論文將使用區域特徵作為三維物體辨識系統中的特徵比對基礎，區域特徵相較於全域特徵，比較不受到測量視角些微變化的影響，即物體姿態些微變化的影響，相較於半區域特徵，特徵比對的計算量卻小非常多，故將採用區域特徵作三維物體辨識適用之特徵。

## 1.4 論文主體與貢獻

以往在電腦圖學以及圖形識別這兩個領域，在國外已經有許多文獻提出物體的三維物體辨識法則，其中包括使用基因演算法、類神經網路、模糊系統以及基於樣版模式的三維物體辨識系統，但是這些法則都未考慮在使用雷射測距儀測量物體時，外在環境或式儀器本身等因素之影響，導致所測量到的物體表面是凹凸不平的，例如：物體的表面材質為透明或是反光，導致無法測量到物體表面上的點，或是儀器本身的誤差，導致使用雷射測距儀所獲得的深度資訊會有誤差。以上兩個原因導致深度資訊中的物體之表面凹凸不平，在本論文中稱這種現象為三維雜訊，所以在進行三維資訊辨識前，必須對於獲得的三維資訊進行前處理，以消除三維雜訊並獲得近似於原本物體之表面。本論文貢獻一：利用兩階段的雜訊處理，使得測量時所獲得物體的表面資訊更可以回復成原本物體的表面資訊，以取得正確的物體表面特徵，並且提高物體在雜訊影響嚴重下之辨識率。

當未知物體在複雜場景中，先將物體個別分離出來，然後可以平行地進行下一步的辨識，以加快辨識速度。本論文貢獻二：提出一個切割複雜場景的架構，利用深度變化將一個複雜場景分割成幾個不同的單一場景，然後有效地將每個物體從個別單一場景個別分離出來，作為接下來辨識系統的輸入。

當未知物體在複雜場景中，可能會有被其他物體遮蔽的情形發生，為了解決物體在被遮蔽的環境下，也可以有效地辨識出該物體。本論文貢獻三：利用邊緣圖像(edge map)的概念，將物體切割成許多不同的區塊，利用偵測物體的步階邊緣(step edge)以及屋脊邊緣(roof edge)之演算法，並且使用形態學運算將物體的所有邊緣變成連結成封閉邊緣，形成該物體的邊緣圖像，之後可以針對邊緣圖像中每個區塊抽取特徵，然後針對未知物體的每個區塊特徵作辨識，即使該物體的某區塊被遮蔽住了，也可以有效地辨識出來該未知物體。

為了同時提高辨識三維物體辨識系統的辨識率以及不犧牲其辨識速度的前提下，本論文使用了兩階段辨識系統，將原本的一階段辨識系統，該系統只比對未知物體的整體特徵 [32]，後方加入了辨識物體區塊的辨識系統，形成了二階段辨識系統，在第一階段辨識系統完後，列出可能是該未知物體的候選清單，之後再經由第二階段辨識系統仔細比對未知物體與後選清單上的物體，兩者的所有區塊特徵，進而辨識出該未知物體。



## 第二章 相關技術原理

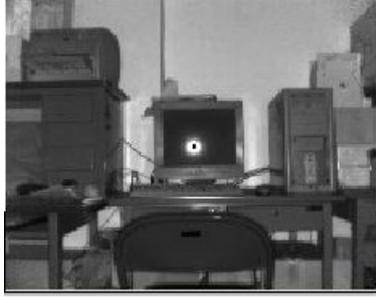
本章分成六節，依照本論文第三章所提出的三維物體分割與辨識系統中的每個步驟之演算法，依其順序分成六小節討論，每節依序介紹演算法的原理及流程。2.1 節將介紹深度資訊的取得與表示方式，以及如何計算深度資訊中每一點的法向量；2.2 節將介紹在取得深度資訊時會受到測量誤差的影響，藉此提出雜訊模型來模擬深度資訊中的測量誤差，並且介紹處理雜訊的方法；2.3 節介紹分割所量測到的深度影像之演算法，並適用於測量的場景為一個複雜場景時；2.4 節介紹如何偵測三維物體的邊緣，並且使邊緣形成封閉的邊界；2.5 節將介紹如何抽取出三維物體的區域特徵；2.6 節將介紹將區域特徵統計成多維度直方圖，然後如何其比對其直方圖。

### 2.1 深度資訊處理

本節將介紹由雷射測距儀取得一個場景的深度資訊，其深度資訊為多維度之陣列，每個二維維度陣列中的值，為所量測的點投影到卡氏座標系中三個正交向量上的值，若是只取投影在雷射測距儀之切平面上的陣列，即為深度影像，其中 2.1.1 節將說明深度資訊的取像原理與表示方式；2.1.2 節將說明如何由深度資訊計算出深度資訊中每點的法向量，並形成一個多維度的法向量陣列。

#### 2.1.1 三維深度資訊

強度影像(intensity image)跟深度影像(range image)的差別是：強度影像的像素值表示的是物體表面經由光反射後的強度資訊，即經由取像儀器所接收的亮度值(gray level)而強度影像如圖 2-1(a)所示；深度影像的像素值表示的是物體表面與取像儀器的所在切平面之距離，即經由取像儀器所接收的深度值(depth)，深度影像如圖 2-1(b)所示。



(a)



(b)

圖 2-1 強度影像與深度影像之差異 (a)強度影像 (b)深度影像

一般的二維強度影像可用下式表示：

$$\mathbf{A} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}_{m \times n}, \quad (2.1)$$

其中  $1 \leq i \leq m$ ,  $1 \leq j \leq n$ , 上式也可表示為：

$$\mathbf{g} = \begin{bmatrix} g(1,1) & \cdots & g(1,n) \\ \vdots & \ddots & \vdots \\ g(m,1) & \cdots & g(m,n) \end{bmatrix}, \quad (2.2)$$

其中  $1 \leq x \leq m$ ,  $1 \leq y \leq n$ , 每個像素值  $a_{ij}$  或是  $g(x,y)$  代表的是物體上該點在影像中的亮度值，然而在深度資訊(range data)表示中，也可用下式來表示：

$$\mathbf{P} = \begin{bmatrix} \mathbf{p}_{11} & \cdots & \mathbf{p}_{1n} \\ \vdots & \ddots & \vdots \\ \mathbf{p}_{m1} & \cdots & \mathbf{p}_{mn} \end{bmatrix}_{m \times n}, \quad (2.3)$$

其中  $\mathbf{p}_{ij} = [x, y, z]^T$ ,  $1 \leq i \leq m$ ,  $1 \leq j \leq n$ , 將屬於物體部分表面上的點，對於一個平面陣列作投影，示意圖如圖 2-2 所示，以獲得該物體或是該場景的三維資訊，每個物體部分表面上的點對應其陣列中的像素，但像素在陣列中的相對位置，並非是物體上的點在空間中真正的相對位置，每個陣列的元素  $\mathbf{p}_{ij}$  表示的是物體上該點在空間中的位置，立體圖如圖 2-3 所示，若只取得深度資訊在  $z$  方向投影的影像，就形成了深度影像(range image)，而深度影像會因為測量視角的不同，擷取到物體不同的姿態。

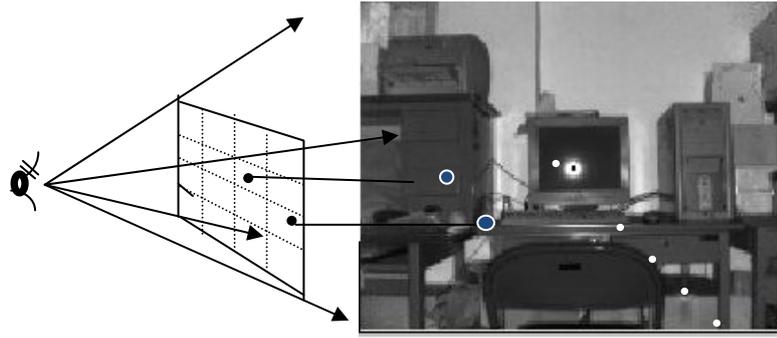


圖 2-2 深度資訊的擷取方式之示意圖

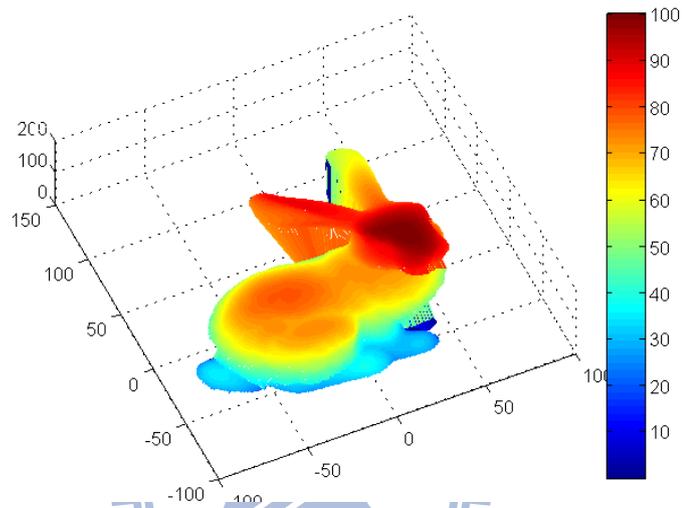


圖 2-3 深度資訊顯示於三維座標中之顯示圖

### 2.1.1 法向量計算

深度資訊可以描述出物體表面輪廓的資訊，同時也可以計算出該點在空間中的法向量，深度資訊中法向量的計算是利用 Hoffman [21]所提出的方法來計算某點  $\mathbf{p}$  的法向量  $\mathbf{n}_p$ 。首先，要定義出一個遮罩(mask)  $S_{xy}$ ，遮罩大小為  $S \times S$ ，遮罩中的係數為  $w_{st}$ ，然後移動遮罩於深度資訊陣列中，去計算遮罩下所對應的點集合  $P = \{\mathbf{p}_{11}, \mathbf{p}_{12}, \dots, \mathbf{p}_{st}, \dots, \mathbf{p}_{ss}\}$  所形成曲面的法向量  $\mathbf{n}_{p_{st}}$ ，其中遮罩中心所對應的點為  $\mathbf{p}_{st}$ ，也是此點集合的代表點。以  $3 \times 3$  為例，其中  $\mathbf{p}_{st} = \mathbf{p}_{22}$ ，如圖 2-4 所示。

$w_{11}$	$w_{12}$	$w_{13}$
$w_{21}$	$w_{22}$	$w_{23}$
$w_{31}$	$w_{32}$	$w_{33}$

(a)

$\mathbf{p}_{11}$	$\mathbf{p}_{12}$	$\mathbf{p}_{13}$
$\mathbf{p}_{21}$	$\mathbf{p}_{22}$	$\mathbf{p}_{23}$
$\mathbf{p}_{31}$	$\mathbf{p}_{32}$	$\mathbf{p}_{33}$

(b)

圖 2-4  $3 \times 3$  遮罩所對應的係數及點 (a) 所對應的係數 (b) 所對應的點

目的希望取得下列  $D$  的最小值，如下式所示：

$$D = \sum_{(s,t) \in S_{xy}} (w_{st} (\mathbf{p}_{st} - \mathbf{p}_{\bar{s}\bar{t}})^T \mathbf{n}_{\mathbf{p}_{\bar{s}\bar{t}}})^2, \quad (2.4)$$

其中，令協方差矩陣  $\mathbf{C}$ ，如下式所示：

$$\mathbf{C} = \sum_{(s,t) \in S_{xy}} w_{st} (\mathbf{p}_{st} - \mathbf{p}_{\bar{s}\bar{t}}) (\mathbf{p}_{st} - \mathbf{p}_{\bar{s}\bar{t}})^T, \quad (2.5)$$

則可以得到  $D = \mathbf{n}_{\bar{s}\bar{t}}^T \mathbf{C} \mathbf{n}_{\mathbf{p}_{\bar{s}\bar{t}}}$ ，由於  $\mathbf{n}_{\mathbf{p}_{\bar{s}\bar{t}}}$  是法向量，所以令  $\mathbf{n}_{\mathbf{p}_{\bar{s}\bar{t}}}$  為單位法向量，即  $\mathbf{n}_{\mathbf{p}_{\bar{s}\bar{t}}}^T \mathbf{n}_{\mathbf{p}_{\bar{s}\bar{t}}} = 1$ ，若欲使  $D$  最小值，使用拉格朗日乘子法可得下式：

$$\frac{\partial}{\partial \mathbf{n}} (D + \lambda (1 - \mathbf{n}_{\mathbf{p}_{\bar{s}\bar{t}}}^T \mathbf{n}_{\mathbf{p}_{\bar{s}\bar{t}}})) = 0, \quad (2.6)$$

由此可得  $\mathbf{C} \mathbf{n}_{\mathbf{p}_{\bar{s}\bar{t}}} = \lambda \mathbf{n}_{\mathbf{p}_{\bar{s}\bar{t}}}$ ，代入(2.6)式後，可得下式：

$$D = \mathbf{n}_{\mathbf{p}_{\bar{s}\bar{t}}}^T \lambda \mathbf{n}_{\mathbf{p}_{\bar{s}\bar{t}}} = \lambda \mathbf{n}_{\mathbf{p}_{\bar{s}\bar{t}}}^T \mathbf{n}_{\mathbf{p}_{\bar{s}\bar{t}}} = \lambda, \quad (2.7)$$

由上式可知，所以要使  $D$  有  $D_{\min}$ ，恰好為所對應的是協方差矩陣的最小特徵值  $\lambda = \lambda_{\min}$ ，而對應到特徵向量就是該點  $\mathbf{p}_{\bar{s}\bar{t}}$  的單位法向量  $\mathbf{n}_{\mathbf{p}_{\bar{s}\bar{t}}} = [n_x, n_y, n_z]^T$ 。

## 2.2 三維雜訊處理

本節將介紹之後用於處理三維雜訊的演算法，進而有效地消除三維雜訊。2.2.1 將介紹應用適應性中值濾波器；2.2.2 將介紹移動式最小平方法。

### 2.2.1 適應性中值濾波器

適應性中值濾波器(adaptive median filter)是中值濾波器(median filter)的改良，是一個非線性的濾波器，原本中值濾波器對於某些特定隨機雜訊的類型，像是脈衝雜訊，提供了絕佳抗雜訊的能力，並且對於同樣遮罩大小的線性平滑濾波器所產生的模糊少了許多。然而，適應性中值濾波器相對於原本的中值濾波器，在雜訊的處理上也可有跟中值濾波器一樣的效果；但是相對於中值濾波器處理後的影像，卻更可以保留影像內物體的邊緣與細節，降低物體的失真與模糊。

根據 Gonzalez 與 Woods[22]所定義的，主要是利用一個濾波器遮罩  $S_{xy}$ ，大小為  $S \times S$ ，其中  $S$  可以允許的最大值為  $S_{\max}$ ， $w_{st}$  為遮罩所對應的係數，其中  $w_{st} = 1$ ， $g(s, t)$  為遮罩  $S_{xy}$  下所對應的值， $1 \leq s, t \leq S$ ，以  $3 \times 3$  的遮罩為例，如圖 2-5 所示

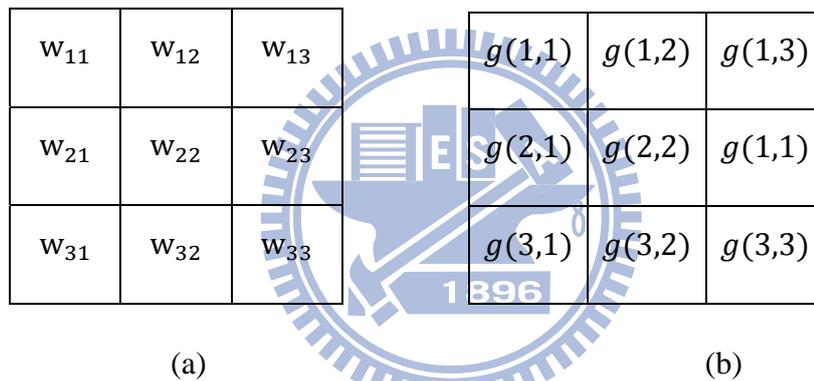


圖 2-5  $3 \times 3$  遮罩所對應的係數及值 (a) 遮罩係數 (b) 遮罩下所對應的值

此遮罩的大小會因為濾波器區域內影像的統計特性而改變其遮罩大小，此演算法有主要三個目的：(1) 移除脈衝雜訊 (2) 提供其他可能不是脈衝雜訊的平滑化 (3) 減少物體邊界過度細小化與增厚化等失真。

其演算法如下：

符號表示：

$$z_{\min} = \min_{g(s,t) \in S_{xy}} \{w_{st} \tilde{g}(s, t)\}$$

$$z_{\max} = \max_{g(s,t) \in S_{xy}} \{w_{st} \tilde{g}(s, t)\}$$

$$z_{\text{med}} = \text{median}\{w_{st} \tilde{g}(s, t)\}$$

$z_{xy} = w_{st} g(\bar{s}, \bar{t})$ , 其中  $(\bar{s}, \bar{t})$  為該遮罩  $S_{xy}$  之中心位置

適應性中值濾波器，依照下列二步驟來運作。

步驟一：

$$A_1 = z_{med} - z_{min}, A_2 = z_{med} - z_{max}$$

假如  $A_1 > 0$  並且  $A_2 < 0$ ，跳到步驟三

否則增加  $S$  的大小

假如  $S < S_{max}$ ，重複步驟二，否則輸出  $z_{xy}$

步驟二：

$$B_1 = z_{xy} - z_{min}, B_2 = z_{xy} - z_{max}$$

假如  $B_1 > 0$  並且  $B_2 < 0$ ，輸出  $z_{xy}$

否則輸出  $z_{med}$

## 2.2.2 移動式最小平方法

由 Lancaster 與 Salkauskas[23]所提出的移動式最小平方法(moving least square)，

是將原本加權最小平方法(weighted least square)，如下式所示：

$$\min_{f \in \Pi_m^d} \sum_i \theta(\|\bar{\mathbf{x}} - \mathbf{x}_i\|) \|f(\mathbf{x}_i) - f_i\|^2, \quad (2.8)$$

作改良，最後如下式所示：

$$f(\mathbf{x}) = f_{\mathbf{x}}(\mathbf{x}), \min_{f_{\mathbf{x}} \in \Pi_m^d} \sum_i \theta(\|\mathbf{x} - \mathbf{x}_i\|) \|f_{\mathbf{x}}(\mathbf{x}_i) - f_i\|^2, \quad (2.9)$$

主要用於使多維資料中的點分布較為平滑，是將原本作為輸入的任意固定點  $\mathbf{x}_i$ ， $\mathbf{x}_i \in \mathcal{R}^d$ ，改成此固定點  $\mathbf{x}_i$  可移動於整個參數域  $\mathcal{R}^d$  中，利用多個區域函數  $f_{\mathbf{x}}(\mathbf{x})$  去近似一個全域函數  $f(\mathbf{x})$ ，其中  $\theta(\mathbf{x})$  是一個加權函數(weighted function)。

移動式最小平方法也被應用在電腦圖學領域中，Marc Alexa[24]等人應用移動式最

小平方法來處理三維空間的資料點，是專門針對三維物體之曲面表示(representation)或是三維場景之重建(reconstruction)等問題，這個方法也可以解決在使用雷射測距儀或是其他儀器去取得物體表面的點資料時，出現取樣不均(up-sampling or down-sampling)或是測量距離誤差的問題，利用移動式最小平方方法去對這些三維資料點重新取樣以改變原本三維資料點的位置，可計算出近似原本物體曲面上的點；此外，也應用在處理電腦中三維模型，改變曲面的平滑程度，使物體表面更為平滑且連續，這些問題在電腦圖學領域中是個重要的議題。在本論文中，將使用移動式最小平方方法使三維資料點所建立的曲面平滑化，以達成消除雜訊的效果。

Marc Alexa 等人所提出的概念如下所示：

假設某區域曲面  $S_p$  是物體表面  $S$  的一部分，在使用雷射測距儀取得物體表面上的所有點資料集合  $P_t$  時，因為測量誤差的關係，導致所量測到之每一個區域曲面上的點集合  $P = \{p_i\}$  可能不在曲面  $S_p$  上，其中  $p_k$  為代表此集合  $P$  的點，如圖 2-6(a)所示，為了要獲得一個區域曲面  $S_{\bar{p}}$  以近似原本曲面  $S_p$ ，首先，將此空間點集合  $P = \{p_i\}$  重新取樣，找出一群更能代表曲面  $S_p$  的點集合  $R = \{r_i\}$ ，其中  $r_k$  為代表此集合  $R$  的點，如圖 2-6(b)所示，再利用加權最小平方方法去求出點集合  $R$  的近似區域曲面  $S_{\bar{p}}$  (MLS surface)，以近似真正的曲面  $S_p$ ，最後再將  $r_k$  投影至近似區域曲面  $S_{\bar{p}}$  上，以獲得投影點  $q_k$ ，重複此步驟使得  $P_t$  中的所有點都進行更新，如圖 2-6(c) 所示，這樣被稱為移動式最小平方方法，這個方法可以確保近似的區域曲面  $S_{\bar{p}}$  與原本的區域曲面  $S_p$  的誤差會達到最小，如圖 2-6(d)所示。

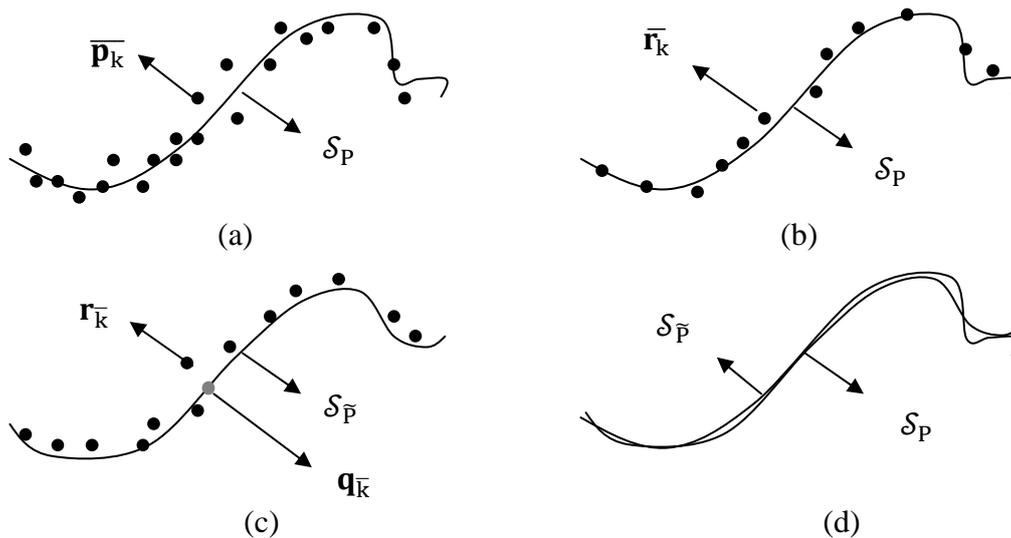


圖 2-6 移動式最小平方法流程

- (a) 某點集合之代表點  $\bar{p}_k$                       (b) 重新取樣的點集合之代表點  $\bar{r}_k$   
 (c)  $\bar{r}_k$  投影在原本曲面上之投影點  $\bar{q}_k$     (d)  $S_{\bar{p}}$  為近似原本曲面之曲面

將移動式最小平方法中所提的部分點集合  $P = \{p_i\}$  移動於整體點集合  $P_t$  之中，去求得每個近似區域曲面  $S_{\bar{p}}$  後，然後更新每個點集合的代表點，就完成了去除雜訊的動作了，其中移動式最小平方法的演算法步驟如下，如圖 2-7 所示。

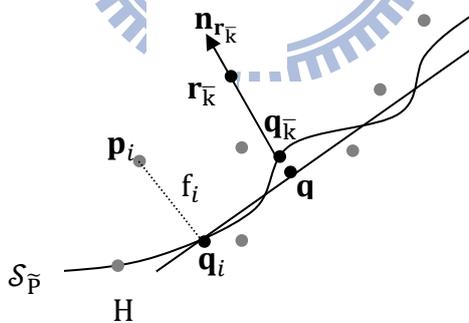


圖 2-7 移動式最小平方法之原理

步驟一：求得點  $r_{\bar{k}}$  的單位法向量  $n_{r_{\bar{k}}}$

定義一群空間點之集合  $P = \{p_i\}$ ，然後對此空間點集合重新取樣以獲得  $R = \{r_i\}$ ，使得  $R$  相對於  $P$  更能描述原來物體的表面  $S_P$ ，並計算出代表其中此集合  $R$  的點  $r_{\bar{k}}$ 。首先，為了要計算出  $r_{\bar{k}}$  此點的單位法向量  $n_{r_{\bar{k}}}$ ，用 2.1.2 節的方法來計算法向量，先計

算該群空間點集中的每點  $\mathbf{p}_i$  與  $\mathbf{r}_{\bar{k}}$  的距離，然後加上加權函數  $w_i = \theta(\|\mathbf{p}_i - \mathbf{r}_{\bar{k}}\|)$ ，最後形成的協方差矩陣  $\mathbf{C} \in \mathcal{R}^{3 \times 3}$ ，如下式所示：

$$\mathbf{C} = \sum_i w_i (\mathbf{p}_i - \mathbf{r}_{\bar{k}}) (\mathbf{p}_i - \mathbf{r}_{\bar{k}})^T, \quad (2.10)$$

計算  $\mathbf{C}$  的最小特徵值所對應的特徵向量，即單位法向量  $\mathbf{n}_{\mathbf{r}_{\bar{k}}}$ 。

步驟二：求得  $\mathbf{r}_{\bar{k}}$  投影至切平面  $\mathcal{H}$  的投影點  $\mathbf{q}$

如圖 2-7 所示，切平面  $\mathcal{H}$  會滿足下式：

$$\mathcal{H} = \{\mathbf{p}_i \mid \mathbf{p}_i^T \mathbf{n}_{\mathbf{r}_{\bar{k}}} - D = 0, \mathbf{p}_i \in P \subset P_t\}, \quad (2.11)$$

其中  $D$  是常數， $\mathbf{q}$  是  $\mathbf{r}_{\bar{k}}$  投影在切平面  $\mathcal{H}$  上的點，可以使得下式有最小值：

$$\min \sum_i (\mathbf{p}_i^T \mathbf{n}_{\mathbf{r}_{\bar{k}}} - D)^2 \theta(\|\mathbf{p}_i - \mathbf{q}\|), \quad (2.12)$$

令  $\theta(\|\mathbf{p}_i - \mathbf{q}\|) = e^{-\frac{\|\mathbf{p}_i - \mathbf{q}\|^2}{h^2}}$ ，是一個高斯函數，也是一個加權函數，其中  $h$  是一個固定的參數，然後設  $\mathbf{q} = \mathbf{r}_{\bar{k}} + t \mathbf{n}_{\mathbf{r}_{\bar{k}}}$  代入式(2.12)，並對(2.12)式進行偏微分，可獲得下式：

$$2 \sum_i \left( (\mathbf{p}_i - \mathbf{r}_{\bar{k}} - t \mathbf{n}_{\mathbf{r}_{\bar{k}}})^T \mathbf{n}_{\mathbf{r}_{\bar{k}}} \right) \left( 1 + \frac{((\mathbf{p}_i - \mathbf{r}_{\bar{k}} - t \mathbf{n}_{\mathbf{r}_{\bar{k}}})^T \mathbf{n}_{\mathbf{r}_{\bar{k}}})^2}{h^2} \right) e^{-\frac{\|\mathbf{p}_i - \mathbf{r}_{\bar{k}} + t \mathbf{n}_{\mathbf{r}_{\bar{k}}}\|^2}{h^2}}, \quad (2.13)$$

其中  $t \neq 0, t \in \left[-\frac{h}{2}, \frac{h}{2}\right]$ ，求得  $t$  以得到  $\mathbf{q}$ 。

步驟三：求  $\mathbf{p}_i$  至切平面  $\mathcal{H}$  之距離  $f_i$  與  $\mathbf{p}_i$  投影至切平面上之投影點  $\mathbf{q}_i$

由圖 2-7 可得知  $f_i$ ，如下式所示：

$$f_i = \mathbf{n}_{\mathbf{r}_{\bar{k}}}^T (\mathbf{p}_i - \mathbf{q}), \quad (2.14)$$

並設立局部座標系統(local coordinate system)於切平面  $\mathcal{H}$  上， $(u_i, v_i)$  為該座標系統的座標值，在此局部座標下，令  $\mathbf{q}_H$  為該局部座標系統之原點，即  $\mathbf{q}_H = (0, 0)$ ，其中  $\mathbf{q}$  和  $\mathbf{q}_H$  在三維空間中是同一點，只是  $\mathbf{q}$  是相對於世界座標系統(global coordinate system)的點， $\mathbf{q}_H$  是相對於局部座標系統的點，接下來要計算其他點  $\mathbf{q}_i$  相對於此局部座標系統下的點  $\mathbf{q}_{iH}$ 。

首先，令任意兩向量  $\mathbf{u}, \mathbf{v} \in H$  且  $\mathbf{u}, \mathbf{v} \in \mathcal{R}^3$  為此局部座標之兩正交單位向量，其中

$\mathbf{u} \times \mathbf{v} = \mathbf{n}_{\bar{r}_k}$  且  $\mathbf{u}^T \mathbf{v} = 1$ ，欲求出投影點  $\mathbf{q}_i$ ，如圖 2-8 所示。

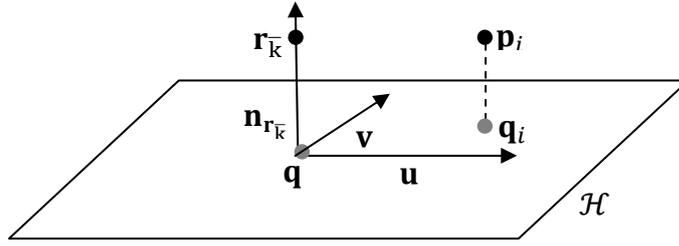


圖 2-8 投影點的區域座標

由圖 2-8 可得  $\mathbf{q}_{iH}$ ，如下式所示：

$$\mathbf{q}_{iH} = (u_i, v_i) = ( (\mathbf{q}_i - \mathbf{q})^T \mathbf{u}, (\mathbf{p}_i - \mathbf{q})^T \mathbf{v} ). \quad (2.15)$$

步驟四：求得近似區域曲面  $S_{\bar{p}}$

在此局部座標系統，令一個最高次為  $n$  次多項式來表示此近似區域曲面  $S_{\bar{p}}$ ，如下式所示：

$$S_{\bar{p}} = g(u_i, v_i) = \sum_{a+b \leq n} c_{ab} u_i^a v_i^b = c_{n0} u_i^n + c_{n-1,0} u_i^{n-1} v_i + \dots + c_{00}, \quad (2.16)$$

由(2.15)式將計算出來的投影點集合  $Q_{iH} = \{\mathbf{q}_{iH}\}$  代入(2.16)式中，並計算出該區域曲面  $S_{\bar{p}}$  的多項式係數  $\mathbf{c} = [c_{n0} \ c_{n-1,0} \ \dots \ c_{01} \ c_{00}]^T$ ，其中係數的總個數為  $\frac{n(n+1)}{2}$ 。

因為  $g(u, v)$  要使此加權最小平方誤差為最小，要滿足下式：

$$\sum_i (g(u_i, v_i) - f_i)^2 \theta (\|\mathbf{p}_i - \mathbf{q}\|), \quad (2.17)$$

其中， $g(u_i, v_i)$  為投影點  $\mathbf{q}_{iH}$  跟區域曲面  $S_{\bar{p}}$  的距離。

步驟五：最後求得  $\mathbf{r}_{\bar{k}}$  投影至區域曲面  $S_p$  的投影點  $\mathbf{q}_{\bar{k}}$

定義  $\mathbf{r}_{\bar{k}}$  在區域曲面  $S_p$  的投影點  $\mathbf{q}_{\bar{k}}$ ，如下式所示：

$$\mathbf{q}_{\bar{k}} = \mathbf{q} + g(0,0) \mathbf{n}_{\bar{r}_k}, \quad (2.18)$$

其中， $g(0,0)$  是  $\mathbf{q}$  跟區域曲面  $S_{\bar{p}}$  的垂直距離，即區域曲面  $S_{\bar{p}}$  之多項式的常數係數  $c_{00}$ ，再將原本  $\mathbf{q}$  沿著法向量方向移動長度  $g(0,0)$ ，即可得到區域曲面  $S_p$  上的近似投影點  $\mathbf{q}_{\bar{k}}$ ，再將之前設  $\mathbf{q} = \mathbf{r}_{\bar{k}} + t \mathbf{n}_{\bar{r}_k}$  代入(2.18)式，最後可得  $\mathbf{q}_{\bar{k}}$ ，如下式所示：

$$\mathbf{q}_{\bar{k}} = \mathbf{r}_{\bar{k}} + (t + c_{00}) \mathbf{n}_{\mathbf{r}_{\bar{k}}}. \quad (2.19)$$

## 2.3 複雜場景分割

本節將介紹分割複雜場景的演算法及概念，並介紹如何分割深度影像，使之成為多個單一場景，並且從多個單一場景中分離來多個單一物體出來。2.3.1 節將介紹多重臨界值法，使得複雜場景變成許多單一場景；2.3.2 節將介紹深度影像中像素之間的連通性，以分離出每個物體。

### 2.3.1 多重臨界值法

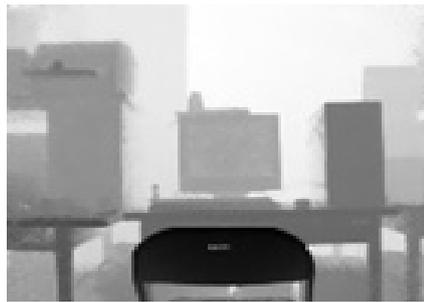
在影像處理中，常利用臨界值法(thresholding)[22]來分割影像，找出適當的臨界值將影像分割成目標物體與背景兩部分，而臨界值法中，最常使用直方圖來統計影像中的灰階值，可表示成(2.20)式，其中灰階值的區間範圍為 $[0, G]$ ，灰階值區間個數(bin size)為 $b$ ，區間大小 $\Delta_b = \frac{(G-0)}{b}$ ， $r_i$  是在整體區間 $[0, G]$  中的第  $i$  個區間的區間位準，其中  $r_i = \Delta_b i$ ，而  $h_i$  是該區間  $[r_{i-1}, r_i]$  內灰階值個數，

$$\mathbf{H} = [h_1 \ h_2 \ \dots \ h_b]_{b \times 1} \quad h_i = h(r_i), \quad (2.20)$$

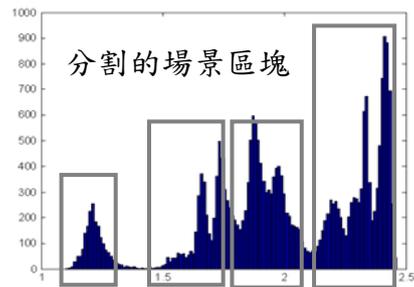
然後，可根據不同的分割問題來找出直方圖中的適當臨界值去分割影像，使影像成為兩個區域；然而，若尋找的臨界值數目不只一個，則稱之多重臨界值法(multilevel thresholding)。

在本論文中，將設計出一個多重臨界值演算法，是以雷射測距儀測量場景所獲得的深度影像為基礎，然後根據深度影像中像素的深度值變化，對複雜場景進行分割，最後可複雜場景分割成前景、中景、後景 等單一場景。分割的目的是為了提高三維物體辨識的速度。在進行物體辨識的時候，可以同步進行辨識且找出物體所在的正確位置。

一個深度影像，如圖 2-9(a)所示，利用直方圖對深度距離做統計，可得到圖 2-9(b)。



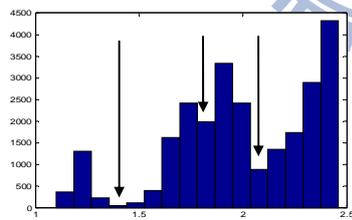
(a)



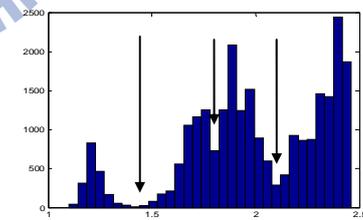
(b)

圖 2-9 深度影像與深度統計後的直方圖 (a)深度影像 (b)直方圖

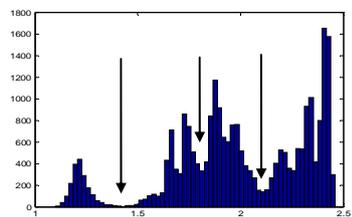
一般而言，根據經驗會將深度影像分割成前景、中景、與後景等，臨界值的個數不超過 5 為主，然後由圖 2-9(a)可以用看的出來，深度影像中物體的深度變化幾乎連續的，所以只需要找出深度變化不連續的部分，就可以有效的分割深度影像，如圖 2-9(b)所示，可將圖 2-9(a)的深度影像分割成四個場景，而每個場景的邊界範圍所在，就是直方圖中局部最小值(local minimum)。多重臨界值的數目以及大小，會隨著直方圖橫軸的深度值區間數目增加而增加；但是，可以發現初始局部最小值的位置大約不變，如箭頭所示，依然是後來的局部最小值，如圖 2-10 所示。



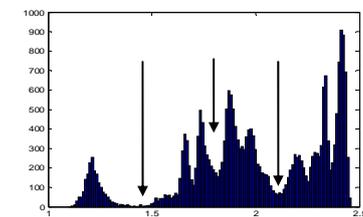
(a)



(b)



(c)



(d)

圖 2-10 不同區間個數的直方圖與臨界值位置

(a)  $b = 16$  (b)  $b = 32$  (c)  $b = 64$  (d)  $b = 128$

根據以上所描述深度影像的直方圖之局部最小值的特性，可依照下列的演算法來有效的分割深度影像：

步驟一：決定初始的樣本區間個數，並找出局部最小值

令直方圖的樣本區間個數為 2 的次方，即  $b = 2^k$ ,  $k \in \mathbb{N}$ ，然後決定初始的  $k$  值，即  $k_i$ ，然後找出直方圖中的局部最小值的位置及數值，若局部最小值的個數少於  $T$  個，則增加  $k$ ，直到局部最小值的個數超過  $T$  個，不再增加  $k$ ，則  $b_T = 2^{k_i-1}$ 。而局部最小值之集合  $D_{b_T}$ ，如下式所示：

$$D_{b_T} = \{d_1, d_2, \dots, d_T\} = \bigcup_{i=1}^b h_i, h_{i-1} \leq h_i \leq h_{i+1}. \quad (2.21)$$

步驟二：增加深度值取樣區間個數，以獲得更準確的臨界值

增加深度值區間個數  $b = 2^k$  其中  $b > b_T$ ，隨著  $k$  增加，所以原本獲得之局部最小值集合  $D_{b_T}$  中的元素  $d_j$  也會是後來局部最小值之集合  $D_b$  中的元素  $d_\ell$  的附近，可由圖 2-10 看的出來，隨著深度值區間個數是 2 的次方，隨著次方的增加，後來的部分局部最小值  $d_\ell$  會滿足下式：

$$\forall (d_j \in D_{b_T} \wedge d_\ell \in D_b) \rightarrow \exists d_\ell \in D_b, d_\ell = d_j \pm \frac{1}{2} \Delta_b, \quad (2.22)$$

之後再用後來的局部最小值更新去原來的局部最小值，即令  $d_\ell = d_j$ ，直到更新的局部最小值與經由步驟一計算所得的局部最小值，兩者的誤差小於某個值  $T_h$  時，也就是最後的局部臨界值不再變化，就可以停止增加深度值區間個數，其中  $T_h$  的大小是根據深度影像中所有相鄰兩點像素的深度值之差的平均值所決定的。

最後，利用多重臨界值法找出的局部臨界值，可用來分割深度影像圖 2-9，可計算出 3 個局部最小值將影像分割成四個單一場景，分割後的影像以二值影像表示，白色區域表示分割後的影像，如圖 2-11 所示。

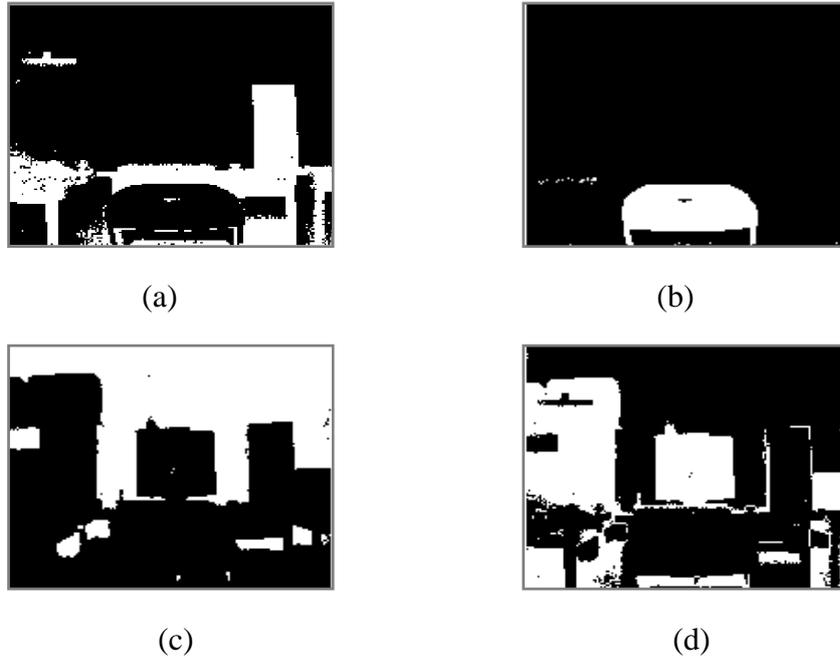


圖 2-11 深度影像之單一場景 (a)場景一(b)場景二(c)場景三(d)場景四

### 2.3.2 連通量分析

像素之間的連結性可以決定影像當中的區域或是邊界。首先，要確定兩個像素是否相鄰，必須要先定義出鄰接的方式，再依照這些鄰接的像素是否滿足某些特定的相似準則，來判斷這些像素是否為相鄰的像素。

首先， $N(P)$ 表示與  $P$  點相鄰的像素之集合。像素之間鄰接的方式[22][25]，常用的有 4-鄰接(4-neighbors)，以 $N_4(P)$ 表示，如圖 2-12(a)，8-鄰接(8-neighbors)，以 $N_8(P)$ 表示，如圖 2-12(b)。下圖中的灰色區塊表示  $P$  點相鄰之像素的位置，虛線表示其連通路徑，若對於任意像素  $P$  而言，其餘的像素  $S$  ( $S$  代表影像中像素的子集合)滿足上面的鄰接方式，並且存在著一條全部由  $S$  中的像素所組合的連通路徑，可以稱之  $P$  的連通成份 (connected component)，依據不同的連接方式，會有不同的連通方式，如圖 2-13 所示。

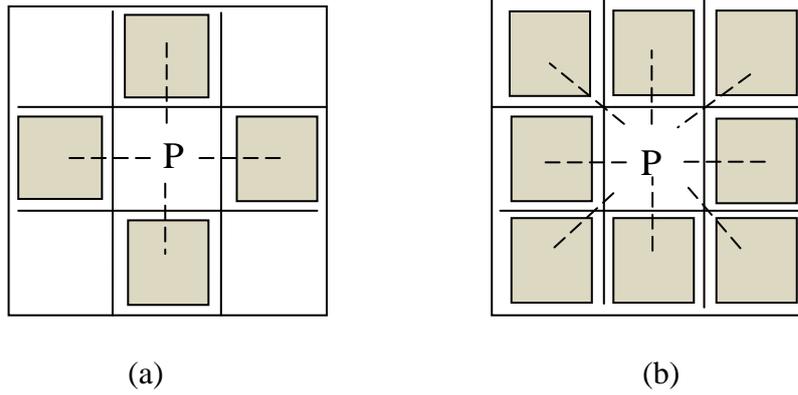


圖 2-12 不同的鄰接方式 (a) 4-鄰接  $N_4(P)$  (b) 8-鄰接  $N_8(P)$

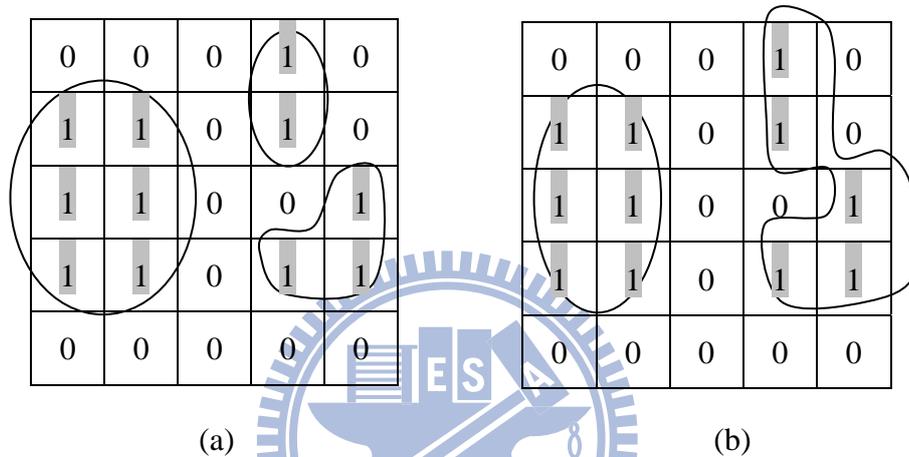


圖 2-13 不同的連通方式 (a) 三個 4 連通成份 (b) 兩個 8 連通成份

圖 2-13(a)是以 4-鄰接為鄰接方式來形成 4 連通(4-connected)，故二值影像可以被標記為三個 4 連通成份(4-connected component)；圖 2-13(b)是以 8-鄰接為鄰接方式來形成 8 連通(8-connected)，故二值影像可以被標記為兩個 8 連通成份 (8-connected component)；依照不同的鄰接定義，會有不同連通路徑，使得同樣的影像形成不同區域大小及個數。

## 2.4 邊緣資訊處理

本節將介紹如何找出物體的所有邊緣，並且使物體的邊緣成為封閉邊界，最後物體將根據此邊界而被分割成許多區塊。2.4.1 節介紹一個有效的邊緣偵測演算法：Canny 邊緣偵測器，用來尋找物體深度資訊中的邊緣，即步階邊緣；2.4.2 節將介紹如何利用法向量的變化，稱之法向量的梯度變化，進而找出物體曲面變化較大的邊緣，即屋脊邊緣；2.4.3 節將介紹影像處理中的形態學運算，目的是使物體的邊緣成為封閉邊界；2.4.4 節

將介紹區域成長法，用於取得物體每個區塊中的特徵。

## 2.4.1 邊緣偵測

在影像分割的方法中，主要是利用影像強度值的兩個基本特性來進行影像的分割：不連續性與相似性；其中邊緣偵測就是一種最常見的方法，是利用影像中強度值不連續的特性來分割影像的方法。

在影像當中的邊緣往往是介於兩個區域之間且有著不同灰階強度的邊界，兩個各別的区域可能是物體與背景或是物體與物體，使得邊緣處的灰階值變化較為劇烈，如圖 2-17 所示。

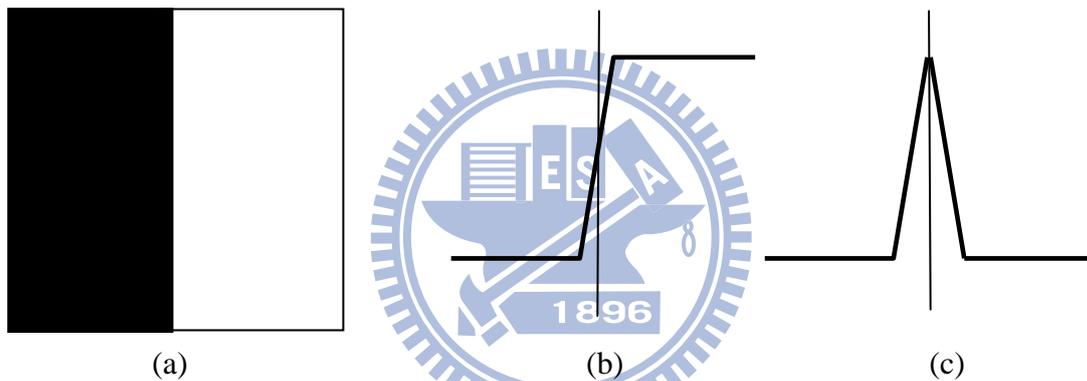


圖 2-14 影像邊緣之灰階與一階導數之變化

(a) 具明顯邊緣影像 (b) 水平灰階變化 (c) 一階導數變化

由圖 2-14(c) 可以觀察一階導數的極值所在，往往是影像中邊緣出現的區域，故找出區域導數的極值所在，即可找出區域邊緣。在數位影像中常利用梯度運算子去找出其區域導數極值的所在，進而尋找出邊緣點，除此之外，還有灰階矩量搜尋法、鏈狀輪廓追蹤法、空間矩量次像素搜尋法等，以下將介紹如何使用梯度運算子來搜尋邊緣的所在。

影像  $g$  上的某一點像素值  $g(x, y)$  在  $(x, y)$  處的梯度向量如下式所示：

$$\nabla \mathbf{g}(x, y) = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\partial g}{\partial x} \\ \frac{\partial g}{\partial y} \end{bmatrix}, \quad (2.23)$$

此向量的大小為  $\nabla g$ ，也是個在邊緣偵測上一個重要的量，如下式所示：

$$\nabla g(x, y) = \text{mag}(\nabla \mathbf{g}) = (G_x^2 + G_y^2)^{\frac{1}{2}}, \quad (2.24)$$

此梯度向量  $\nabla \mathbf{g}$  的方向指向  $g(x, y)$  在座標  $(x, y)$  改變最大率的方向上，而該點  $(x, y)$  梯度最大改變率之角度，如下式所示：

$$\alpha(x, y) = \tan^{-1} \left( \frac{G_x}{G_y} \right). \quad (2.25)$$

由於數位影像上的像素值為非連續的性質，所以在偵測數位影像的邊緣時，最常使用的方法是使用一個空間遮罩  $S_{xy}$ ，其大小是  $3 \times 3$ ，去通過數位影像  $g$  的每一點像素  $g(x, y)$ ，去計算遮罩  $S_{xy}$  內像素  $\tilde{g}(s, t)$  的線性響應  $R$ ，下式所示：

$$R = \sum_{s, t \in S_{xy}} w_{st} \tilde{g}(s, t) \quad (2.26)$$

以上列之方式來近似一階導數  $G_x$  及  $G_y$ ，其中又以 Sobel 邊緣偵測器[22][26]最常被使用，其遮罩中的係數如圖 2-15 所示。

-1	-2	-1	-1	0	1
0	0	0	-2	0	2
1	2	1	-1	0	1

(a)
(b)

圖 2-15 Sobel 運算子 (a)計算  $G_x$  (b)計算  $G_y$

經由上圖所列之係數，可獲得  $G_x$  及  $G_y$ ，如下式所示：

$$G_x = [\tilde{g}(3,1) + 2\tilde{g}(3,2) + \tilde{g}(3,3)] - [\tilde{g}(1,1) + 2\tilde{g}(1,2) + \tilde{g}(1,3)] \quad (2.27)$$

$$G_y = [\tilde{g}(1,3) + 2\tilde{g}(2,3) + \tilde{g}(3,3)] - [\tilde{g}(1,1) + 2\tilde{g}(2,1) + \tilde{g}(3,1)] \quad (2.28)$$

Sobel 邊緣偵測器計算簡單，所以常被使用在數位影像中，但是對於較弱的邊緣卻沒辦法偵測出來，所以接下來將介紹 Canny 邊緣偵測器 [22][27]，相對於 Sobel 或是其他線性或非線性的邊緣偵測器，是一個強而有力的邊緣偵測器，可同時找出主要的邊緣並降低不相干的細節。

Canny 邊緣偵測器的演算法說明如下：

步驟一：設定一個高斯遮罩

影像先經過一個特定標準差 $\sigma$ 的高斯濾波器來使影像平滑且降低雜訊。高斯濾波器為一個空間遮罩 $S_{xy}$ ，大小為 $S \times S$ ，其遮罩內的係數由中心開始往外分布為高斯函數分布，係數由下式中的 $\sigma$ 來決定：

$$w_{st} = \frac{1}{2\pi\sigma} e^{-\frac{[(s-\bar{s})^2 + (t-\bar{t})^2]}{2\sigma^2}}, \quad (2.29)$$

其中 $(\bar{s}, \bar{t})$ 為該遮罩 $S_{xy}$ 之中心位置，而高斯函數之分布的平均值為 0， $w_{st}$  取整數且 $\sum_{s,t \in S_{xy}} w_{st} = 1$ ，最後該遮罩的響應 $R$ 可由(2.26)式計算，其遮罩中的係數如下圖 2-16 所示，為一個 $5 \times 5$ 高斯濾波器的遮罩，其中 $\sigma = 1$ 。

	1	4	7	4	1
	4	16	26	16	4
$\frac{1}{273} \times$	7	26	41	26	7
	4	16	26	16	4
	1	4	7	4	1

圖 2-16 高斯濾波器遮罩

步驟二：計算每點梯度大小

由(2.23)(2.24)式之定義計算每點之局部梯度大小，即 $\nabla g(x, y)$ ，並可用 Sobel 運算子(2.27)式、(2.28)式算出  $G_x$  及  $G_y$  或是其他運算子算出 $G_x$  及  $G_y$ 並且由(2.25)式算出該點之邊緣方向  $\alpha(x, y)$ 。一個邊緣點 $(x, y)$ 被定義為該梯度方向有局部最大強度的點。

步驟三：形成山脊來劃分強弱邊界

將步驟二所求得的這些邊緣點，在梯度大小影像上造成山脊，接著沿著這些山脊上方追蹤，並將沒有真正在山脊頂點的所有點像素值為零，以便輸出中得到一條細長的線，

這個程序稱為非最大值的抑制(non-maximal suppression)。這些山脊像素將用兩個臨界值  $T_1$  及  $T_2$  來分界，其中  $T_1 < T_2$ 。若是其山脊像素的值大於  $T_2$ ，會被稱為強邊緣像素，若是其山脊像素的值介於  $T_1$  與  $T_2$  之間，會被稱為弱邊緣像素。

步驟四：利用連通性分析

最後利用 8 連通 (8-connected)路徑 將強邊緣像素與弱邊緣像素進行邊緣連接。

如圖 2-17 所示，分別使用 Sobel 邊緣偵測器與 Canny 邊緣偵測器於圖 2-17(a)，

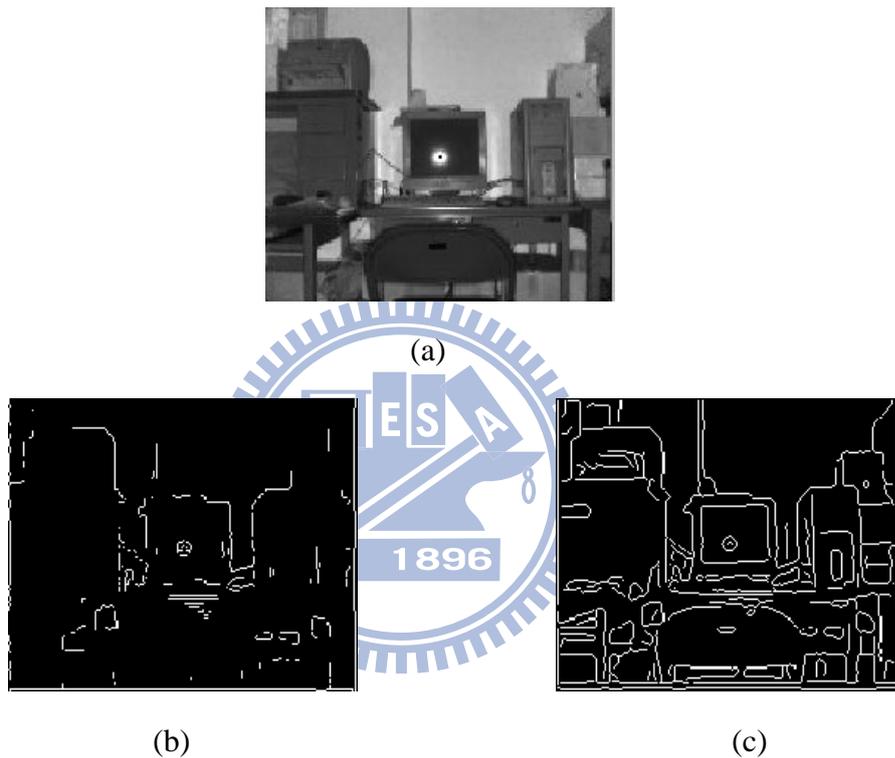


圖 2-17 數位影像使用不同邊緣偵測器之差異

(a) 數位影像 (b)使用 Sobel 邊緣偵測器 (c)使用 Canny 邊緣偵測器

其中，可以看的出圖 2-17(c)中的邊緣，不論是弱邊緣或是強邊緣，都可以有效的找出來。故本論文中，將使用 Canny 邊緣偵測器來偵測出深度影像中物體與物體或是物體與背景的邊緣。

## 2.4.2 法向量的梯度變化

由深度資訊中可以獲得物體表面上之點所包含的資訊，藉由這些資訊可以計算該點

之法向量、曲率或是其他資訊，來作為三維物體曲面之特徵，而這些特徵可以作為影像分割或物體辨識時的依據，其中在影像分割的領域中，一般的二維影像分割所需的資訊是影像像素的強度值變化作為影像分割時的依據，主要是依照像素間強度的不連續性或是相似性來進行影像的分割，如 2.4.1 節所探討的；但是，相對於二維影像中因像素強度值變化所造成的邊緣，在三維深度資訊中所定義的邊緣或是真實三維空間中人類肉眼看到的物體邊緣，是因為空間中點的位置變化，即空間中曲面的法向量變化過大所造成的，所以在三維的深度資訊中，一個物體與其他物體或是物體上本身的邊緣可分為兩種，步階邊緣(step edge)與屋脊邊緣[21][28][29]。

因為曲面上每點深度(即距離)的不連續所造成的邊緣，稱之步階邊緣，如圖 2-18(c)所示，還有因為曲面變化劇烈所造成的邊緣，即曲面上每點法向量的不連續所造成的邊緣，稱之屋脊邊緣，如圖 2-18(d)所示。

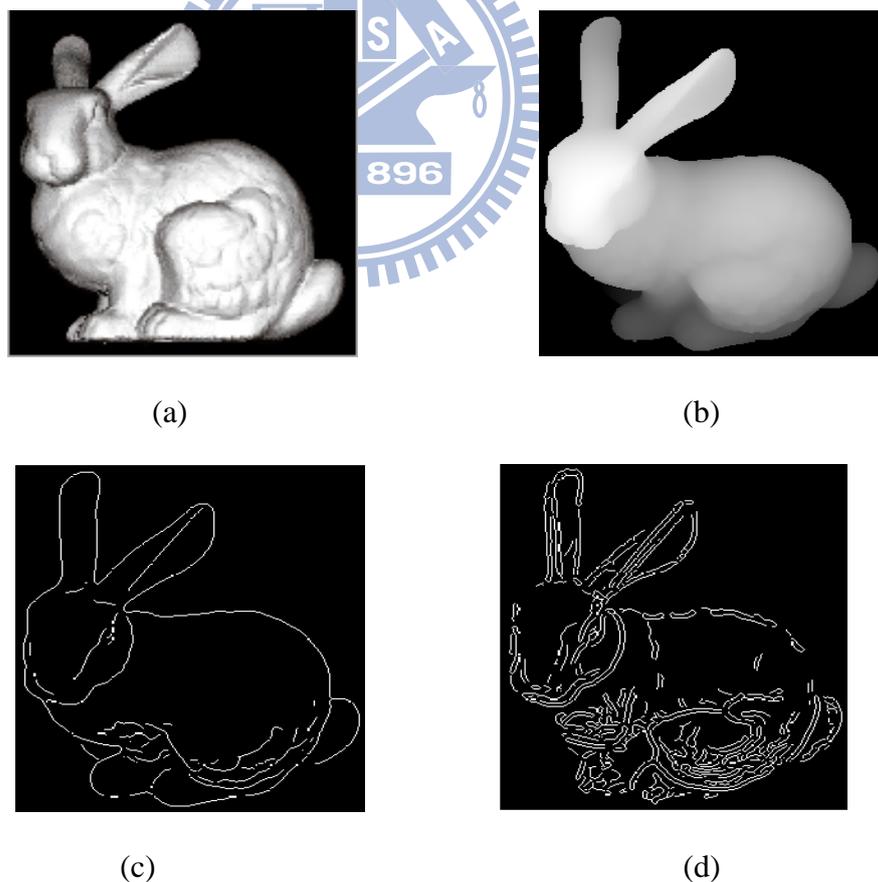


圖 2-18 藉由深度資訊所測量到物體之邊緣

(a) 真實物體 (b) 深度影像 (c) 步階邊緣 (d) 屋脊邊緣

然而，要獲得三維物體的步階邊緣，只要將深度影像中的深度值替換成原來二維影像中的強度值，再用 2.4.1 節中的 Canny 邊緣偵測來尋找深度影像中物體的邊緣；然而，本論文提出一個有效且抗雜訊的偵測屋脊邊緣之方法；首先，要偵測到三維物體的屋脊邊緣，如圖 2-18(d)所示，首先必須要取得深度資訊中每點的法向量，再利用每點與鄰近點法向量的變化，計算出每點所對應的梯度大小，以獲得梯度影像。

首先，由 2.1.1 節可獲得深度影像所對應的法向量陣列，如下式所示：

$$\mathbf{N} = \begin{bmatrix} \mathbf{n}_{11} & \cdots & \mathbf{n}_{1n} \\ \vdots & \ddots & \vdots \\ \mathbf{n}_{m1} & \cdots & \mathbf{n}_{mn} \end{bmatrix}_{m \times n} \quad (2.30)$$

其中  $\mathbf{n}_{ij}$  為深度資訊中每個點  $\mathbf{p}_{ij}$  對應的法向量，對法向量陣列使用一個空間遮罩  $S_{xy}$ ，其中遮罩下所對應的中心為  $\tilde{\mathbf{n}}_{st} = \mathbf{n}_{ij}$ ，其大小  $S \times S$ ，以  $3 \times 3$  的遮罩為例，如圖 2-19 所示。

1	1	1	$\tilde{\mathbf{n}}_{11}$	$\tilde{\mathbf{n}}_{12}$	$\tilde{\mathbf{n}}_{13}$
1	1	1	$\tilde{\mathbf{n}}_{21}$	$\tilde{\mathbf{n}}_{22}$	$\tilde{\mathbf{n}}_{23}$
1	1	1	$\tilde{\mathbf{n}}_{31}$	$\tilde{\mathbf{n}}_{32}$	$\tilde{\mathbf{n}}_{33}$

(a)
(b)

圖 2-19  $3 \times 3$  的遮罩內的係數與對應的法向量 (a)係數 (b)法向量

利用下式來計算每一點  $\mathbf{p}_{ij}$  的梯度值為  $G_{ij}$

$$G_{ij} = \frac{1}{S^2} \sum_{(s,t) \in S_{xy}} (1 - |\tilde{\mathbf{n}}_{st}^T \tilde{\mathbf{n}}_{st}|^\gamma), \quad \gamma \geq 0, \quad (2.31)$$

若  $G_{ij}$  越接近 1 表示  $S_{xy}$  所對應之區域，越有可能是兩平面的交接處，反之若  $G_{ij}$  越接近 0，則表示  $S_{xy}$  所對應之區域是一個平面。

將遮罩  $S_{xy}$  對法向量陣列中的每一點使用(2.31)式，就形成了梯度影像 (gradient image)  $\mathbf{G} = [G_{ij}]_{m \times n}$ ，將法向量之梯度影像顯示在三維座標，可由圖 2-20 可看出屋脊邊

緣所在，即梯度變化大的點所形成之邊緣。

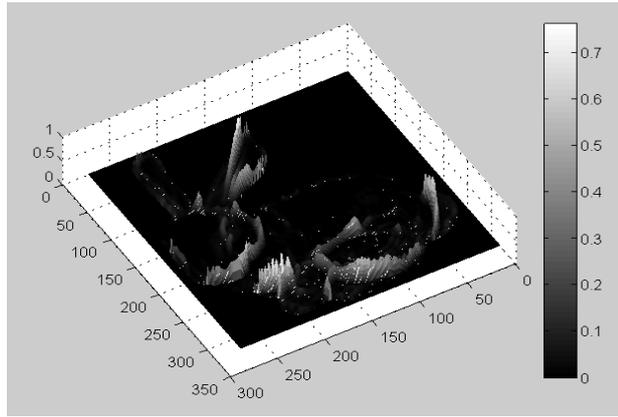


圖 2-20 法向量的梯度變化

(2.31)式主要計算遮罩中的每一點之法向量  $\tilde{\mathbf{n}}_{st}$  與中心的法向量  $\tilde{\mathbf{n}}_{\bar{s}\bar{t}}$ ，兩者的內積的絕對值再乘上  $\gamma$  次方，然後再計算此值與內積最大值 1 的差距，即  $1 - |\tilde{\mathbf{n}}_{st}^T \tilde{\mathbf{n}}_{\bar{s}\bar{t}}|^\gamma$ ，如果  $1 - |\tilde{\mathbf{n}}_{st}^T \tilde{\mathbf{n}}_{\bar{s}\bar{t}}|^\gamma$  很小，表示此兩點幾乎在同一平面上，如果很大，表示這兩點在不同平面上。對  $|\tilde{\mathbf{n}}_{st}^T \tilde{\mathbf{n}}_{\bar{s}\bar{t}}|^\gamma$  取絕對值，一方面是為了避免  $\tilde{\mathbf{n}}_{st}^T \tilde{\mathbf{n}}_{\bar{s}\bar{t}}$  小於 0 時，使得此數列  $\{1 - \tilde{\mathbf{n}}_{st}^T \tilde{\mathbf{n}}_{\bar{s}\bar{t}}\}_{(s,t) \in S_{xy}}$  分布在 0 的兩側時，之後再取平均值，其平均值可能很接近 0，使得遮罩中心可能與其他點在不同的曲面，但最後該點的法向量之梯度值卻很接近 0，導致做邊緣偵測時無法偵測到此點。此外，當  $\gamma$  很大時，梯度影像中的梯度值變化劇烈的點數目會變多，使得在偵測屋脊邊緣時，邊緣的點數目也會變多，使得物體的屋脊邊緣細節過多，由圖 2-21 所示，故一般而言，選擇  $\gamma \in [1, 2]$  即可。

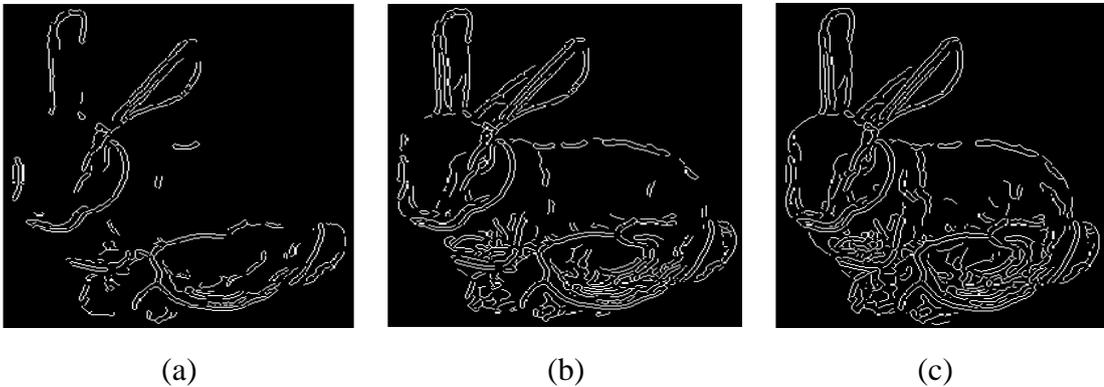


圖 2-21 不同的  $\gamma$  值使得屋脊邊緣細節改變 (a)  $\gamma = 0.5$  (b)  $\gamma = 2$  (c)  $\gamma = 10$

### 2.4.3 形態學運算

形態學運算(morphological operator) [22][25][30]在二值影像(binary image)中的主要應用在於抽取對於適合描述形狀有用的影像成份，例如：抽取出其邊界、連通成份、凸形封包以及區域骨架的形態，也應用於影像的前級與後級處理，例如：形態學濾波、細線化以及剪除，而形態學運算除了用來處理二值影像，也可處理灰階強度影像。本論文將會使用形態學處理其二值影像，故只介紹二值影像之形態學運算。

形態學是一種從二值影像  $A$  中的抽取出物體成分的工具，是由結構元素  $B$  (structuring element)來進行運算，如圖 2-22 所示，其中膨脹(dilation)與侵蝕(erosion)是形態學運算中的兩個最基本的運算，許多形態學演算法都是建立在這兩個原始的運算上，故首先介紹膨脹與侵蝕。

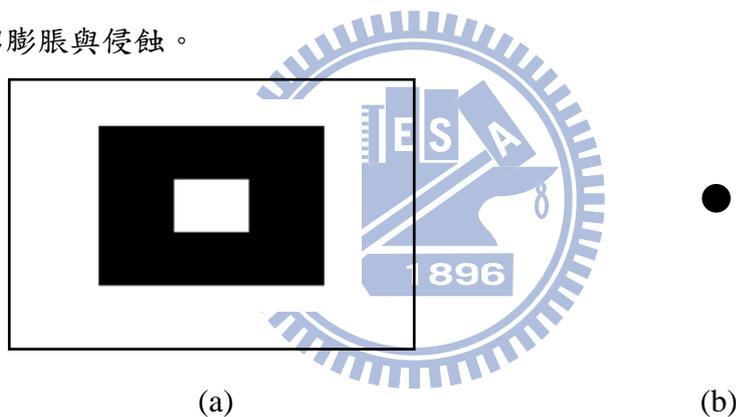


圖 2-22 二值影像與結構元素 (a)原本二值影像  $A$  (b)結構元素  $B$

膨脹是一種使二值影像  $A$  中的物體增大或是變厚，就是使物體向外擴充，在數學上是以集合運算的方式來定義， $A$  藉由  $B$  膨脹記為  $A \oplus B$ ，如下式：

$$C = A \oplus B = \left\{ z \mid (\hat{B})_z \cap A \neq \emptyset \right\}, \quad (2.32)$$

用圖 2-23 來說明其過程，結構元素  $B$  的中心在整個二值影像內平移，若結構元素內的點於物體上的部分點重疊，將會使物體上的點向物體外部膨脹，虛線是膨脹後的邊界。

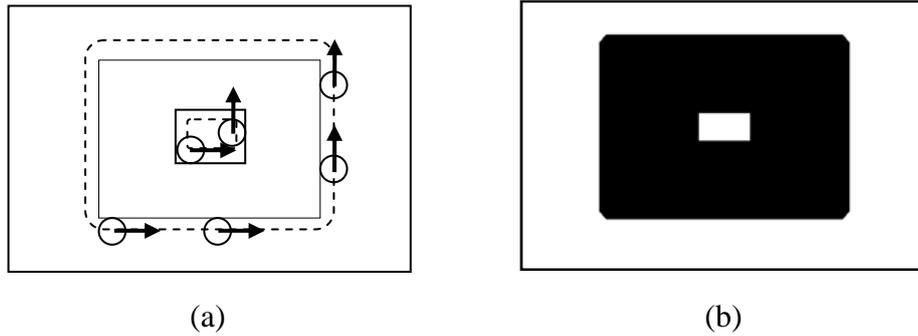


圖 2-23 二值影像的膨脹過程與結果 (a)二值影像進行膨脹 (b) 膨脹後的二值影像

侵蝕是一種使二值影像  $A$  中的物體收縮或是變薄，使物體向內收縮，收縮的方式也是由結構元素  $B$  所控制，在數學上也是以集合運算的方式來定義， $A$  藉由  $B$  侵蝕記為  $A \ominus B$ ，如下式所示：

$$C = A \ominus B = \{z | (\hat{B})_z \subset A\} \quad (2.33)$$

用圖 2-24 來說明，結構元素  $B$  的中心在整個二值影像內平移，若結構元素內的點於物體上的所有點重疊，將會使物體上的點向物體內部收縮，虛線是侵蝕後的邊界。

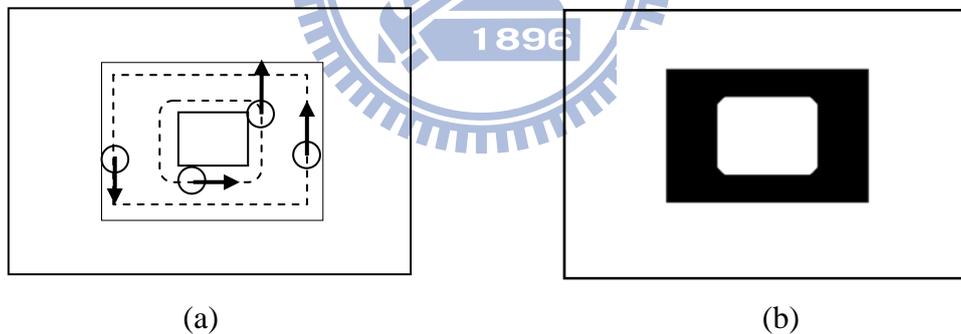


圖 2-24 二值影像的侵蝕過程與結果 (a)二值影像進行侵蝕 (b)侵蝕後的二值影像

在影像處理中，應用膨脹與侵蝕這兩個基本形態運算子的組合，形成其他的形態學演算法，本論文將會使用形態學上的閉合(morphological closing)以及形態學上的收縮(morphological shrinking)，來處理邊界資訊。

形態學上的閉合，是將二值影像  $A$  先以結構元素  $B$  進行膨脹再進行侵蝕，如下式所示

$$A \circ B = (A \ominus B) \oplus B \quad (2.34)$$

可以使物體的輪廓平滑，使窄的中段部分連接起來，填補細常缺口等功能，如圖 2-25 所示。

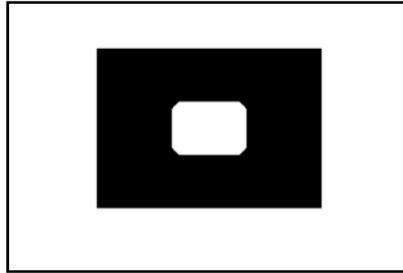


圖 2-25 閉合後的二值影像

形態學上的收縮[30]，主要的概念就是將二值影像中，沒有洞的物體收縮成點，有洞的物體收縮成環狀；此外，物體經過形態學的收縮處理後，依舊滿足尤拉數(Euler number)公式，例如：將圖 2-22 的二值影像，經過  $n$  次的收縮處理後，最後結果如圖 2-26 所示。

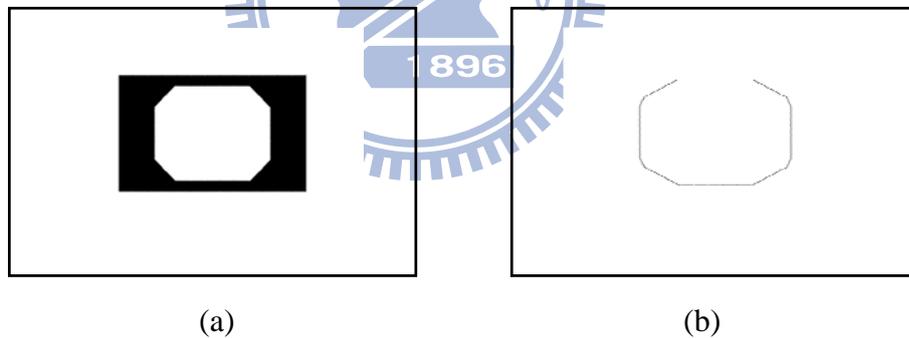


圖 2-26 收縮  $n$  次的二值影像 (a)  $n=20$  (b)  $n=50$

#### 2.4.4 區域成長法

區域成長(region growing)[22]是單一像素或是子區域根據預先定義的準則，然後成長成更大區域的過程。基本的方法是從一組種子點出發，把每個種子點具有相同性質的鄰近點像素添加進來一起進行區域成長。

假設一個影像區域  $\mathcal{R}$  可以根據某種準則分割成  $N$  個不同的子區域  $\mathcal{R}_i$ ， $\mathcal{R}_i$  為第

$i$  個子區域，使得  $\mathcal{R} = \{\mathcal{R}_1, \mathcal{R}_2, \dots, \mathcal{R}_i, \dots, \mathcal{R}_N\}$ ，會滿足下列式子：

- (a)  $\bigcup_{i=1}^N \mathcal{R}_i = \mathcal{R}$ 。
- (b)  $\mathcal{R}_i$  是連通的區域， $i = 1, 2, \dots, N$ 。
- (c)  $\mathcal{R}_i \cap \mathcal{R}_j = \emptyset, \forall i \neq j$ 。
- (d)  $P(\mathcal{R}_i) = \text{True}$ ,  $P$ 是屬性，表示子區域  $\mathcal{R}_i$  都具有相同的屬性。
- (e)  $P(\mathcal{R}_i \cap \mathcal{R}_j) = \text{False}, \forall i \neq j$ 。

本論文將會基於 2.4.3 節所提過的形態學運算以及 2.3.2 節連通性分析等概念，使用形態學重建(reconstruct)來進行區域成長的動作，並以像素之間的 8 連通作為準則，將二值影像中的區域  $\mathcal{R}$  分割成  $N$  個不同的子區域，並且取出這些子區域來計算其對應的三維特徵，以進行三維物體辨識。

形態學重建必須使用兩張影像以及一個結構元素  $B$ ，其中一張影像作為標記(marker)使用，以  $\mathbf{M}_i^k$  表示之，另外一張原始影像  $\mathbf{A}$  作為遮罩使用，以  $\mathbf{A}_k$  表示之，其重建步驟如下：

將初始種子點以隨機的方式撒在影像區域中，並標記初始種子點的位置，使其成為初始標記影像  $\mathbf{M}_i^k$ ，令初始遮罩為原始影像，即  $\mathbf{A}_1 = \mathbf{A}$ ；此外產生一個結構元素  $B$ 。

然後進行下列步驟，從  $k = 1$  開始。

步驟一：將影像種子點的位置變成標記影像  $\mathbf{M}_i^k$ ，從  $i = 1$  開始

隨機撒一個種子於遮罩影像  $\mathbf{A}_k$  中，對該種子點進行標記其遮罩影像中的位置，產生其標記影像  $\mathbf{M}_i^k$ 。

步驟二：形成下一張標記影像，以取代原本的標記影像

首先，利用結構元素對標記影像進行形態學膨脹，滿足下式：

$$\mathbf{M}_{i+1}^k = (\mathbf{M}_i^k \oplus B) \cap \mathbf{A}_k, \quad (2.35)$$

持續重覆步驟二，直到滿足  $\mathbf{M}_{i+1}^k = \mathbf{M}_i^k$ ，其中最後的標記影像滿足  $\mathbf{M}_i^k \subseteq \mathbf{A}_k$ ，即完成該子區域  $\mathcal{R}_i$  的區域成長。

步驟三：重新進行區域成長法於其他區域

重新定義其遮罩影像  $\mathbf{A}_{k+1} = \mathbf{A}_k - \mathbf{M}_i^k$ ，即後來的遮罩影像  $\mathbf{A}_{k+1}$  是原來的遮罩影

像  $A_k$  跟標記影像  $M_i^k$  的差集，然後再重覆步驟一，並且  $k = k + 1$ ，直到最後所有的標記影像的聯集為一開始的原始影像，當區域成長完畢時，會滿足下式：

$$\bigcup_{k=1}^N M_i^k = A, \quad (2.36)$$

其中，為了重建快速，本論文採用了 Vincent [31]提出的快速混合重建法為基礎，來進行上列演算法的步驟二，以加快重建區域的速度。

## 2.5 區域特徵抽取

本節將介紹如何抽取出三維物體表面的區域特徵，進而利用這些特徵來辨識三維物體。2.5.1 節介紹計算深度資訊中每個物體表面的深度變化，作為三維物體的區域特徵之一；2.5.2 節介紹計算深度資訊中物體表面上每一點的法向量分量之夾角，作為三維物體的區域特徵之一；2.5.3 介紹計算深度資訊中物體表面上每一點所在的曲面之曲率，然後利用形狀指標，轉換成三維物體辨識所用的區域特徵之一。

### 2.5.1 表面之深度變化

借由深度資訊可取得物體表面距離雷射測距儀的深度變化，物體表面的深度變化將可以作為比對物體的一個特徵，當物體從複雜場景中被切割出來後，此單一物體的表面上每一點的深度大小  $z_p$  值將在一個範圍  $[z_{\min} z_{\max}]$  之間，將此物體的表面深度變化作正歸化之動作，使得每個物體的表面深度範圍都介於  $[0 1]$  之間，最後物體表面上的每一點  $P$  經過正規化之深度值  $d_p$  為下式：

$$d_p = \frac{(z_p - z_{\min})}{(z_{\max} - z_{\min})}, \quad (2.37)$$

其中  $d_p \in [0 1]$ ，將對單一物體表面上每一點  $P$  的深度大小作運算，最後就得到  $D = [d_{ij}]_{m \times n}$ 。

## 2.5.2 法向量分量之夾角

在進行三維物體的辨識的時候，物體表面上的每一點所對應的法向量，是一個常用的全域特徵，其法向量的方向將由一開始的世界座標(global coordinate) 所決定，但對於利用深度資訊來計算此特徵，使用特徵在比對三維物體時，不是一個很好的特徵，主要是因為此特徵會隨著世界座標所設立的位置而改變。一般而言，都會將取像儀器所在的點將會是世界坐標中的原點  $[0,0,0]^T$ ，而深度資訊中的每一點座標值，都是相對取像儀器所在點的值，若取儀器所在的點相對於場景平移 $[a,b,c]^T$ ，將會使得後來計算物體表面上的點之法向量也會平移 $[a,b,c]^T$ ，如圖 2-27 所示。

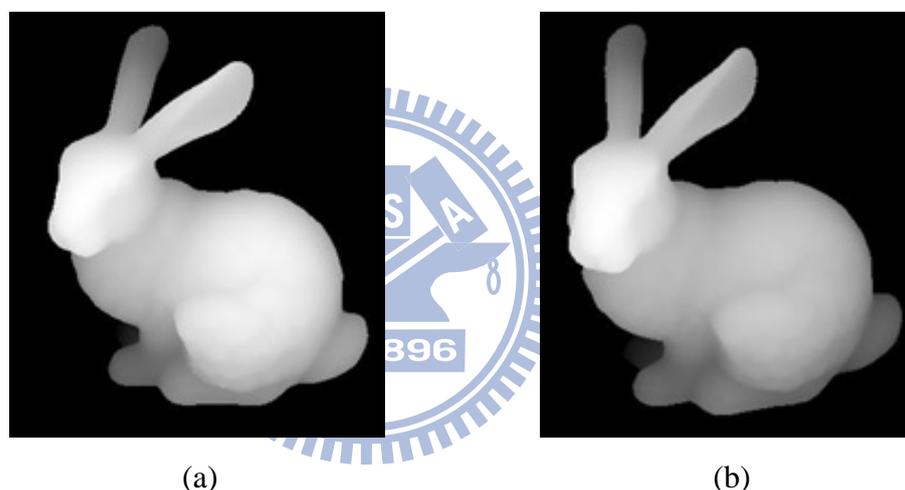


圖 2-27 不同視角下的深度影像 (a)原本的視角 (b)後來的視角

因為以上的原因，所以將法向量轉成區域特徵以作為比對三維物體所用是必要的。首先，計算深度資訊中的每一點  $\mathbf{P}$  之法向量  $\mathbf{n}_p = [n_x, n_y, n_z]^T$ ，計算該法向量於球座標 (sphere coordinate) 中，各分量的之間的夾角[32]，如圖 2-28 所示。

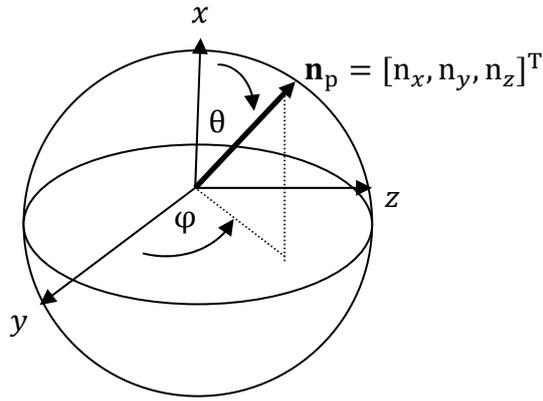


圖 2-28 法向量分量之夾角。

$$\phi_p = \tan^{-1}\left(\frac{n_z}{n_y}\right) + \frac{\pi}{2}, \quad \theta_p = \tan^{-1}\left(\frac{\sqrt{(n_y^2 + n_z^2)}}{n_x}\right) + \frac{\pi}{2}, \quad (2.38)$$

其中  $\phi_p \in [0, \pi]$ ,  $\theta_p \in [0, \pi]$ , 將對每一點  $\mathbf{P}$  的法向量  $\mathbf{n}_p$  作運算, 最後就得到  $\Phi = [\phi_{ij}]_{m \times n}$  與  $\Theta = [\theta_{ij}]_{m \times n}$  作為比對三維物體的特徵, 如圖 2-29 所示。

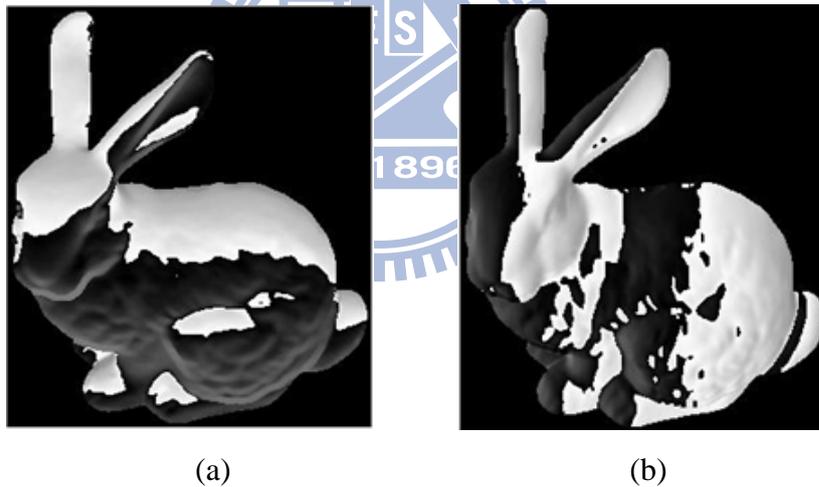


圖 2-29 物體的區域特徵 (a)  $\Phi = [\phi_{ij}]_{m \times n}$  (b)  $\Theta = [\theta_{ij}]_{m \times n}$

### 2.5.3 曲率之形狀指標

在計算曲率的形狀指標(shape index)前, 必須要先計算曲率。曲率(curvature)是一個描述幾何物體不平坦程度的量, 平坦對於不同幾何體不同的意義, 例如: 對於曲線而言, 平坦就是直線; 對於曲面而言, 平坦就是平面, 在二維空間  $\mathcal{R}^2$  中, 曲線  $\mathcal{C}: f(x, y)=0$  上的點  $\mathbf{P}$ , 可用一圓去近似該點  $\mathbf{P}$  附近所形成的曲線, 如圖 2-31 所示, 並定義曲率

大小  $K = \frac{1}{R}$ ， $R$  為此圓之半徑，也可以寫成  $K = \nabla \cdot \left( \frac{\nabla f}{\|\nabla f\|} \right)$ 。

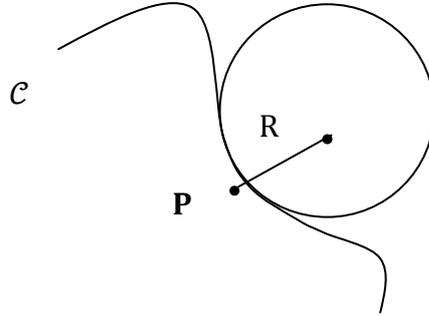


圖 2-30 二維曲率的定義之示意圖

然而，在三維空間  $\mathcal{R}^3$  中，曲面的曲率也可以用來描述該曲面在三維空間中的彎曲變化程度。曲面曲率在微分幾何學(Differential Geometry)[33][34]中，常用的主曲率(principal curvature)，高斯曲率(Gaussian curvature)以及平均曲率(mean curvature)。

主曲率是考慮區域曲面  $\mathcal{S}$  上該點  $\mathbf{p}$  的法向量  $\mathbf{n}_p$  與該點  $\mathbf{p}$  之切平面  $\mathcal{T}$  上任一向量所形成的平面  $\mathcal{H}$  與該曲面  $\mathcal{S}$  之交集，這個交集將是一個平面曲線，計算此平面曲線  $\mathcal{C}$  上的點  $\mathbf{p}$  的曲率。當任意的切平面向量  $\mathbf{T}_n$ ，會對應到其曲率值  $k_n(\mathbf{p})$ ，其中點  $\mathbf{p}$  的曲率值會有兩個極值  $k_{\max}(\mathbf{p})$  與  $k_{\min}(\mathbf{p})$ ，稱為主曲率(principal curvatures)；此外，極值的方向為  $\mathbf{T}_u, \mathbf{T}_v \in \mathcal{H}$  且  $\mathbf{T}_u \perp \mathbf{T}_v$ ，即對應的切平面向量之方向稱為主方向(principal direction)，如圖 2-31 所示。其中  $k_n(\mathbf{p}), k_{\max}(\mathbf{p})$  與  $k_{\min}(\mathbf{p})$  會有下列關係式如下式所示，稱之 Euler 公式[36]：

$$k_n(\mathbf{p}) = k_{\max}(\mathbf{p})\cos^2\theta + k_{\min}(\mathbf{p})\sin^2\theta, \quad (2.39)$$

其中  $\theta$  為  $\mathbf{T}_u$  與  $\mathbf{T}_n$  之夾角。

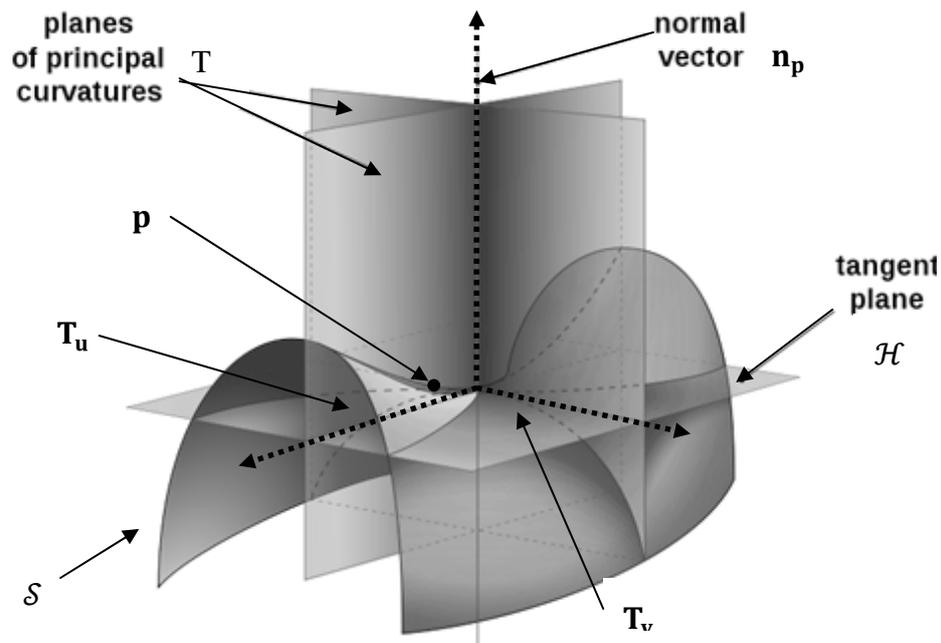


圖 2-31 馬鞍面的主曲率計算之示意圖

資料來源:Wikipedia, author: Eric Gaba.

高斯曲率  $K$  定義為兩個主曲率的乘積， $K = k_{\max}(\mathbf{p})k_{\min}(\mathbf{p})$ ，而平均曲率  $H$  定義為兩個主曲率的平均， $H = \frac{1}{2}(k_{\max}(\mathbf{p}) + k_{\min}(\mathbf{p}))$ 。

高斯曲率可以決定該局部曲面是凸的(convex)，若高斯曲率的值大於 0；該局部曲面是局部鞍點(locally saddle)，則高斯曲率的值小於 0，例如：對於球面、橢球面、單葉雙曲面、橢圓拋物面時，高斯曲率為正，對於偽球面、雙葉雙曲面的一葉、雙曲拋物面為負，對於平面與圓柱體為 0；反之，對於平均曲率而言，平面的平均曲率為 0，但圓柱體非 0。

主曲率之計算[34]，可利用微分幾何學中的曲面， $S$ 可表示下式：

$$S: z = f(x, y), x = x(u, v), y = y(u, v), z = g(u, v), \quad (2.40)$$

$u, v$ 是  $\mathbf{T}_u, \mathbf{T}_v$ 兩向量所形成平面之局部座標值  $(u, v)$ ， $x, y, z$  可以標示成  $\mathbf{T}_u, \mathbf{T}_v$  為基底的參數式。

曲面第一基本形式(first fundamental form)，其係數可寫對稱矩陣形式，如下式所示：

$$\mathbf{I} = \begin{bmatrix} E & F \\ F & G \end{bmatrix} \triangleq \begin{bmatrix} \mathbf{T}_u^T \mathbf{T}_u & \mathbf{T}_u^T \mathbf{T}_v \\ \mathbf{T}_v^T \mathbf{T}_u & \mathbf{T}_v^T \mathbf{T}_v \end{bmatrix}, \quad (2.41)$$

也可寫成下列形式，如下式所示：

$$\mathbf{I} = \begin{bmatrix} E & F \\ F & G \end{bmatrix} \triangleq \begin{bmatrix} \mathbf{G}_u^T \mathbf{G}_u & \mathbf{G}_u^T \mathbf{G}_v \\ \mathbf{G}_v^T \mathbf{G}_u & \mathbf{G}_v^T \mathbf{G}_v \end{bmatrix}, \quad (2.42)$$

其中  $\mathbf{G}_u = [\frac{\partial x}{\partial u}, \frac{\partial y}{\partial u}, \frac{\partial z}{\partial u}]^T$  與  $\mathbf{G}_v = [\frac{\partial x}{\partial v}, \frac{\partial y}{\partial v}, \frac{\partial z}{\partial v}]^T$  為曲面上點  $\mathbf{p}$  之主方向的一階偏微分向量，即沿著  $\mathbf{T}_u, \mathbf{T}_v$  兩向量方向作偏微分。

曲面第二基本形式(second fundamental form)，其係數也可寫對稱矩陣形式，下式所示：

$$\mathbf{II} = \begin{bmatrix} L & M \\ M & N \end{bmatrix} = \begin{bmatrix} \mathbf{G}_{uu}^T \mathbf{n}_p & \mathbf{G}_{uv}^T \mathbf{n}_p \\ \mathbf{G}_{vu}^T \mathbf{n}_p & \mathbf{G}_{vv}^T \mathbf{n}_p \end{bmatrix}, \quad (2.43)$$

其中  $\mathbf{n}_p$  為曲面  $S = F(x, y, z)$  上點  $\mathbf{p}$  的法向量， $\mathbf{G}_{uu} = [\frac{\partial^2 x}{\partial u^2}, \frac{\partial^2 y}{\partial u^2}, \frac{\partial^2 z}{\partial u^2}]^T$ ， $\mathbf{G}_{vv} = [\frac{\partial^2 x}{\partial v^2}, \frac{\partial^2 y}{\partial v^2}, \frac{\partial^2 z}{\partial v^2}]^T$ ， $\mathbf{G}_{uv} = [\frac{\partial^2 x}{\partial u \partial v}, \frac{\partial^2 y}{\partial u \partial v}, \frac{\partial^2 z}{\partial u \partial v}]^T$ ， $\mathbf{G}_{uu}$  表示沿著  $\mathbf{T}_u$  的方向作二次偏微分， $\mathbf{G}_{vv}$  表示沿著  $\mathbf{T}_v$  的方向作二次偏微分； $\mathbf{G}_{uv}$  表示沿著  $\mathbf{T}_u$  的方向作一次偏微分，再沿著  $\mathbf{T}_v$  的方向再做一次偏微分會等於  $\mathbf{G}_{vu} = \mathbf{G}_{uv}$ 。

算出曲面的第一基本形式的係數  $E, F, G$  與第二基本形式的係數  $L, M, N$ ，可以計算下式方程式的兩根：

$$(EG - F^2)k^2 - (EN - 2FM + GL)k + (LN - M^2) = 0, \quad (2.44)$$

可得  $k_1 = k_{\max}(\mathbf{p})$  與  $k_2 = k_{\min}(\mathbf{p})$ 。

也可算出高斯曲率  $K = \frac{LN - M^2}{EG - F^2}$  及平均曲率  $H = \frac{EN - 2FM + GL}{2(EG - F^2)}$ ，若是沿著任意垂直的切平面向量  $\mathbf{T}_n$  及  $\mathbf{T}_{n\perp}$  計算(2.42)式、(2.43)式，也可得到(2.44)式，然後得到的曲率值  $k_n(\mathbf{p})$  及  $k_{n\perp}(\mathbf{p})$  滿足(2.39)式。

最後可由 Koenderink 與 A. van Doorn [35] 定義曲率形狀指標(shape index)， $SI_p$  如下式所示：

$$SI_p = \frac{1}{2} - \frac{1}{\pi} \tan^{-1} \left( \frac{k_{\max}(\mathbf{p}) + k_{\min}(\mathbf{p})}{k_{\max}(\mathbf{p}) - k_{\min}(\mathbf{p})} \right), SI_p \in [0, 1], \quad (2.45)$$

最後算出該點  $\mathbf{p}$  的區域特徵  $SI_{\mathbf{p}}$ 。

在深度資訊中要計算出該陣列中每一點的主曲率，可參考 A. Gray[36]或 P. Krsek, G. Lukacs 與 R. R. Martin[37]所提出的方法去計算某點之曲率；然而，本論文提出了一個計算三維曲率的演算法。首先，跟計算 2.1.1 節的深度資訊  $\mathbf{P} = [\mathbf{p}_{ij}]_{m \times n}$  中每點  $\mathbf{p}_{ij}$  的法向量作法一樣，首先，定義出一個遮罩(mask)  $S_{xy}$ ，遮罩大小為  $S \times S$ ，其遮罩內的係數為  $w_{st} = 1$ ，計算遮罩下所對應的點集合  $P = \{\mathbf{p}_{11}, \mathbf{p}_{12}, \dots, \mathbf{p}_{st}, \dots, \mathbf{p}_{SS}\}$  所形成的曲面  $\mathcal{S}$ ，其遮罩中心對應的點為  $\mathbf{p}_{st}$  為曲面的代表點，以  $3 \times 3$  為例，其中  $\mathbf{p}_{st} = \mathbf{p}_{22}$  如下圖 2-32 所示。

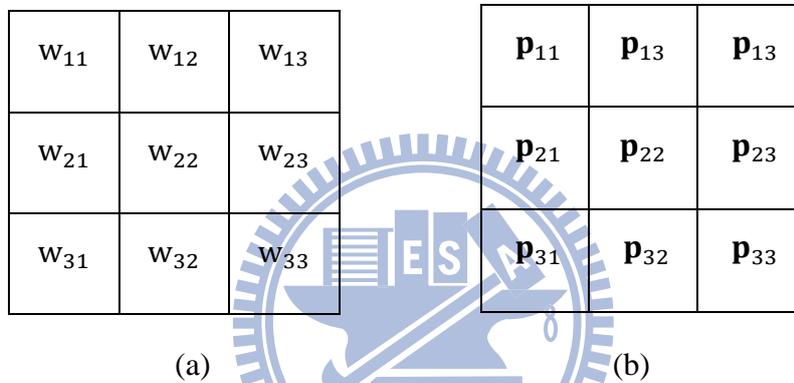


圖 2-32  $3 \times 3$  的遮罩內的係數與對應的點集合 (a)係數 (b)點集合

在本論文中，為了快速計算主曲率，首先先計算  $\mathbf{p} = \mathbf{p}_{st}$  法向量  $\mathbf{n}_{\mathbf{p}}$ ，如 2.1.1 節所敘述的，然後任選  $\mathcal{S}_{\mathbf{p}}$  直切平面向量(tangent direction)  $\mathbf{T}_n$  與  $\mathbf{T}_{n\perp}$  作為近似曲率主方向  $\mathbf{T}_u, \mathbf{T}_v$  的切平面向量，而  $\mathbf{T}_n, \mathbf{T}_{n\perp}$  投影在深度影像所在之切平面  $E$  上的向量為  $\mathbf{T}_i$  及  $\mathbf{T}_j$ ，即深度資訊陣列中的鉛直方向與水平方向，如圖 2-33 所示。

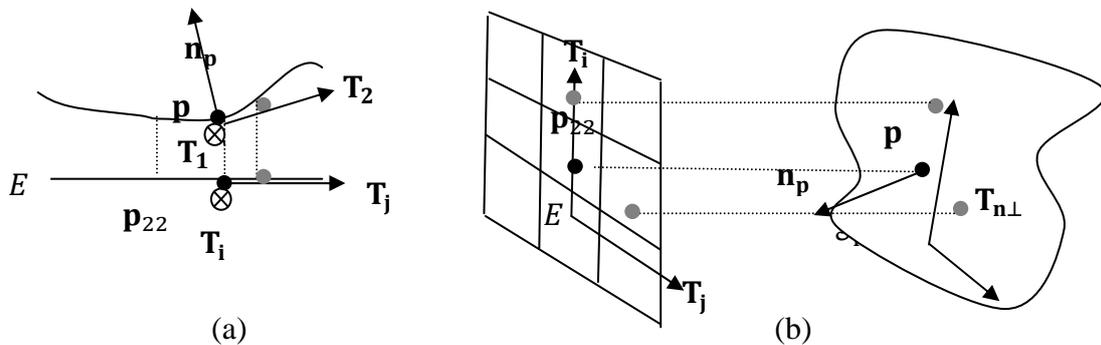


圖 2-33 計算曲率方式之示意圖 (a)側面投影圖 (b)三維空間投影圖

若要沿著該點  $\mathbf{p}$  與附近的點形成的曲面  $S_p$  之切平面向量  $\mathbf{T}_u, \mathbf{T}_v$ ，先利用內插法計算出遮罩下所對應的內插點，以獲得曲面上的點以便計算該點  $\mathbf{p} = \mathbf{p}_{st}$ ，再沿著深度陣列中鉛直方向  $\mathbf{T}_i$  的計算與鄰近點之梯度向量  $\mathbf{G}_i$ ，如，即沿著深度陣列中水平方向  $\mathbf{T}_j$  的鄰近點之梯度向量  $\mathbf{G}_j$ 。

由圖 2-4 可以看得出來，因為深度資訊是物體上的點取樣後投影到深度影像所在的平面陣列上，故鄰近點的相對位置是不變的，所以  $\mathbf{G}_i = \mathbf{G}_u$  且  $\mathbf{G}_j = \mathbf{G}_v$ 。之後再對  $\mathbf{G}_i$  及  $\mathbf{G}_j$  個別沿  $\mathbf{T}_i$  及  $\mathbf{T}_j$  作偏微分，即再計算一次  $\mathbf{G}_i$  及  $\mathbf{G}_j$  的梯度向量，得  $\mathbf{G}_{ii}$ 、 $\mathbf{G}_{jj}$ 、 $\mathbf{G}_{ij}$ 、 $\mathbf{G}_{ji}$ ，其中  $\mathbf{G}_{ii} = \mathbf{G}_{uu}$  依此類推，代入(2.42)式、(2.43)式，再由(2.44)式算得兩個曲率值  $k_i(\mathbf{p}) = k_n(\mathbf{p})$ ， $k_j(\mathbf{p}) = k_{n\perp}(\mathbf{p})$ ，並且滿足(2.39)式，再代入(2.46)式算得近似的  $\tilde{SI}(\mathbf{p})$  曲率之形狀指標[18][32]，如下式所示：

$$\tilde{SI}_p = \frac{1}{2} - \frac{1}{\pi} \tan^{-1} \left( \frac{k_i(\mathbf{p}) + k_j(\mathbf{p})}{k_i(\mathbf{p}) - k_j(\mathbf{p})} \right), \quad (2.46)$$

其中  $k_i \geq k_j$ ，最後就得到  $\mathbf{SI} = [\tilde{SI}_{ij}]_{m \times n}$ ，如圖 2-34 所示。



圖 2-34 曲面形狀指標

## 2.6 區域特徵比對

本節將介紹如何將 2.5 節介紹的區域特徵，統計成直方圖，然後進行比對。2.6.1 節說明如何使多個區域特徵利用直方圖做統計，形成多維度直方圖[32]，其中包括單一物體的整體直方圖以及部分直方圖；2.6.2 說明利用多維度直方圖作比對時，整體直方圖的比對方式與部份直方圖的比對方式。

## 2.6.1 多維度直方圖

直方圖(histogram)在統計上是一個統計樣本之整體分布變化的一個圖，橫軸是樣本區間寬度(bin width)，縱軸顯示該區間內所對應到樣本個數或是機率，也可表示成正規化(normalized)直方圖，以  $\mathbf{H}$  表示之。圖 2-35(a)為圖 2-1(a)強度影像的直方圖，樣本類別是強度值，圖 2-35(b)為圖 2-1(b)深度影像的直方圖，樣本類別是深度值，其類別個數均為 1，均可表示成維度為 1 的陣列，如下式所示：

$$\mathbf{H} = [h_1 \ h_2 \ \dots \ h_b]_{b \times 1} \quad h_i = h(r_i) = \frac{n_i}{n_t}, h_i \in [0, 1], \quad (2.47)$$

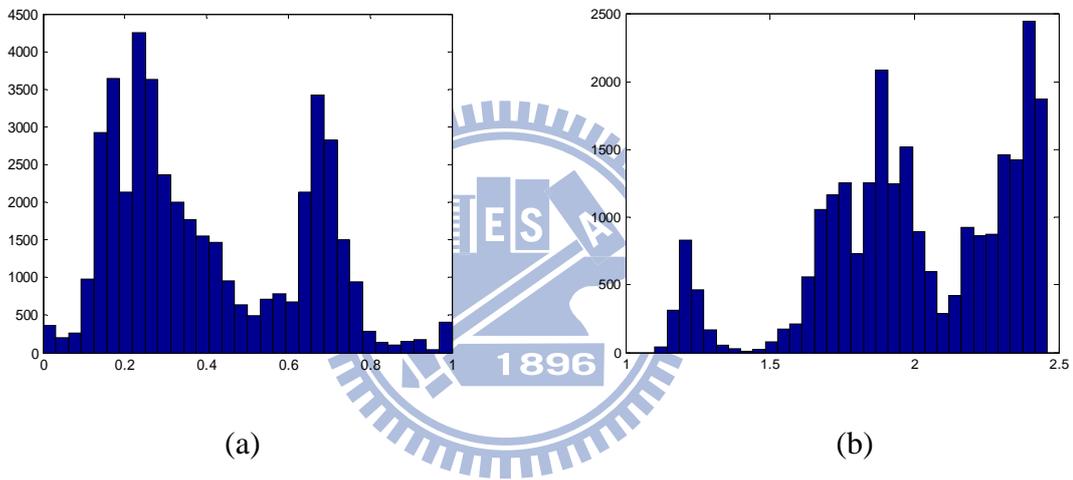


圖 2-35 不同影像的直方圖 (a)強度影像的直方圖 (b)深度影像的直方圖

在(2.47)式中，樣本總個數為  $n_t = \sum_{i=1}^b n_i$ ，整體樣本區間範圍為  $[0, G]$ ，樣本區間個數(bin size)為  $b$ ， $r_i$  是在整體區間  $[0, G]$  中的第  $i$  個區間的區間位準  $r_i = \frac{(G-0)}{b} i$ ， $h_i$  是該區間  $[r_{i-1}, r_i]$  內樣本個數佔整體樣本個數的比例，即機率。

當樣本類別多於一種時，就變成了多維度直方圖(multidimensional histogram)，可推廣到  $d$  維度，即樣本類別數為  $d$ 。以  $d = 3$  為例，如下式所示：

$$\mathbf{H} = [h_{ijk}]_{b_1 \times b_2 \times b_3}, h_{ijk} = h(r_{ijk}) = \frac{n_{ijk}}{n_t}, h_{ijk} \in [0, 1], \quad (2.48)$$

其中  $\mathbf{H}$  為正規化的直方圖，並且滿足下式：

$$\sum_{k=1}^{b_3} \sum_{j=1}^{b_2} \sum_{i=1}^{b_1} h_{ijk} = 1, \quad (2.49)$$

$b_1$ 、 $b_2$ 、 $b_3$ 分別為樣本 1、樣本 2、樣本 3 的區間個數，樣本總個數為 $n_t$ ，會滿足

$$n_t = \sum_{k=1}^{b_3} \sum_{j=1}^{b_2} \sum_{i=1}^{b_1} n_{ijk},$$

個別的整體樣本區間範圍分別為 $[0, L_1]$ 、 $[0, L_2]$ 、 $[0, L_3]$ ， $r_{ijk}$ 是在樣本 1 的整體區間 $[0, G_1]$  中的第  $i$  個區間的區間位準  $\frac{(G_1-0)}{b_1} i$  且在樣本 2 的整體

區間 $[0, G_2]$  中的第  $j$  個區間的區間位準  $\frac{(G_2-0)}{b_2} j$  且在樣本 3 的整體區間 $[0, G_3]$  中的第

$k$  個區間的區間位準  $\frac{(G_3-0)}{b_3} k$ ，而  $h_{ijk}$  是該區間 $[r_{i-1 j-1 k-1}, r_{ijk}]$ 內樣本個數佔整體樣

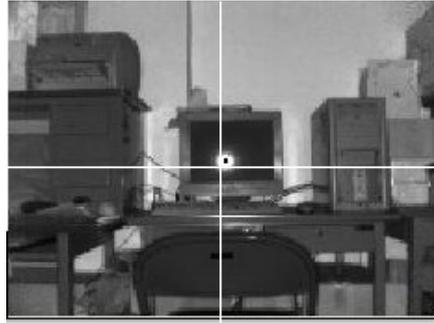
本個數的比例。

直方圖也常被使用在特徵比對上，統計各種特徵的整體變化，使用多維度直方圖作為比對特徵的工具，其好處是在於：(1) 比對時的計算量低。(2)相似物體所取得之特徵，會有相似的直方圖。(3) 即使物體取得特徵時，有雜訊干擾，直方圖的比對受雜訊影響較其他比對方法低。

假設單一物體  $O$  表面的特徵類別數為  $d$ ，以  $d = 3$  為例，此物體的整體特徵被統計成的直方圖，稱之整體直方圖(unity histogram)  $\mathbf{H}^0 = [h_{ijk}]_{b_1 \times b_2 \times b_3}$ 。若物體表面被分成  $N$  個區域，物體的第  $\ell$  個區域  $O_\ell$  之特徵經過統計後，會對應其直方圖，稱之部分直方圖(partial histogram)，會滿足下列(2.50)式。

$$\mathbf{H}_\ell = [h_{ijk}]_{b_1 \times b_2 \times b_3}, \quad \mathbf{H}^0 = \sum_{\ell=1}^N \mathbf{H}_\ell, \quad (2.50)$$

例如：將圖 2-36(a)的強度影像切割成四個區域，各部分直方圖的總和會等於整張強度影像的整體直方圖，如圖 2-38(b)。



(a)

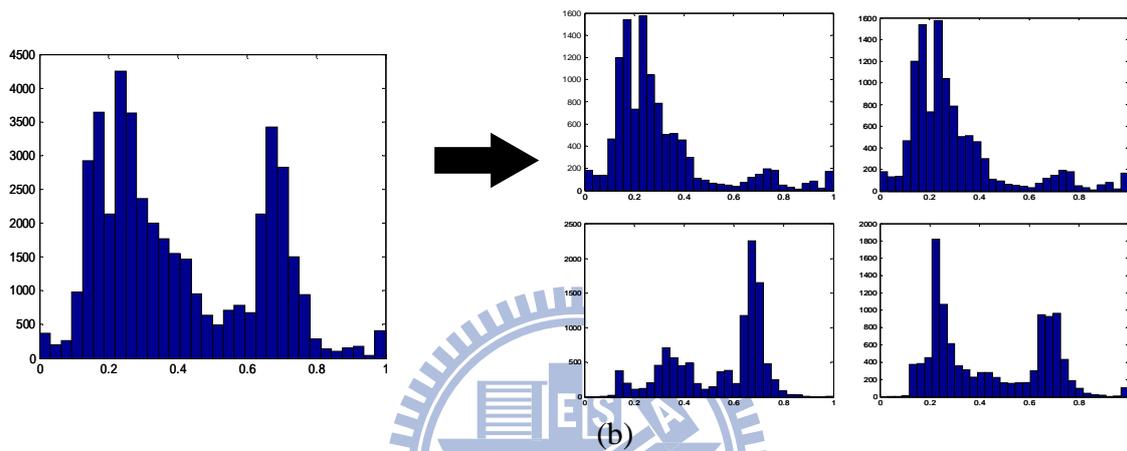


圖 2-36 強度影像的整體直方圖與部分直方圖關係  
 (a) 強度影像 (b) 整體直方圖為部分直方圖之總和

## 2.6.2 直方圖比對

要比對兩個直方圖  $H^Q$  與  $H^V$  是否相似，常用的方法是有三個。第一個方法是由 Swain and Ballard[38]提出的交集測量(intersection measurement)，原本用來比對兩個彩色直方圖的交集程度，如下式所示：

$$\bigcap(H^Q, H^V) = \sum_{i=1}^b \min(h_i^Q, h_i^V), \quad (2.51)$$

第二個方法是由 Kullback 與 Leibler 提出的 Kullback-Leibler divergence[39]，將直方圖視為機率分布函數(probability function)，來比對樣本直方圖  $H^Q$  相對於原本直方圖  $H^V$  的相異程度，如下式所示：

$$D_{\text{KL}}(\mathbf{H}^Q || \mathbf{H}^V) = \sum_{i=1}^b (h_i^Q - h_i^V) \ell n \left( \frac{h_i^Q}{h_i^V} \right), \quad (2.52)$$

最後一個是統計學上常用的  $\chi$ -divergence[40]，也是將兩個直方圖視為機率分布函數，比對任意兩直方圖的相異程度，如下式所示：

$$\chi(\mathbf{H}^Q, \mathbf{H}^V) = \sum_{i=1}^b \frac{(h_i^Q - h_i^V)^2}{(h_i^Q + h_i^V)}, \quad (2.53)$$

而本論文文將應用  $\chi$ -divergence 作為計算不同物體的整體直方圖之相異度與區域直方圖之相異。

在本論文中須個別比對兩個物體  $O_1$  與  $O_2$  的某個姿態下，即  $O_1^{p_1}$  與  $O_2^{p_2}$ ，其各自表面特徵所對應的整體直方圖  $\mathbf{H}^{O_1^{p_1}}$  與  $\mathbf{H}^{O_2^{p_2}}$  以及須個別比對兩物體某個姿態的區域直方圖  $\mathbf{H}_{\ell_1}^{O_1^{p_1}}$  與  $\mathbf{H}_{\ell_2}^{O_2^{p_2}}$ ，其中  $O_1^{p_1}$  表示物體  $O_1$  的第  $p_1$  個姿態，而  $O_1^{p_1}$  可以被分割成  $N_1$  個區域，每個區域  $R_{\ell_1}$  會滿足  $O_1^{p_1} = \bigcup_{\ell_1=1}^{N_1} R_{\ell_1}$ ，而  $\ell_1$  表示物體  $O_1^{p_1}$  第  $\ell_1$  區域；同理， $O_2^{p_2}$  表示物體  $O_2$  的第  $p_2$  個姿態，而物體  $O_2^{p_2}$  可以被分割成  $N_2$  個區域，每個區域  $R_{\ell_2}$  會滿足  $O_2^{p_2} = \bigcup_{\ell_2=1}^{N_2} R_{\ell_2}$ ，而  $\ell_2$  表示物體  $O_2^{p_2}$  第  $\ell_2$  區域。

要比對兩物體某個姿態的整體直方圖  $\mathbf{H}^{O_1^{p_1}}$  與  $\mathbf{H}^{O_2^{p_2}}$  代入(2-53)式，可得兩物體個別姿態的相異度，如下式所示：

$$\chi(O_1^{p_1}, O_2^{p_2}) = \chi(\mathbf{H}^{O_1^{p_1}}, \mathbf{H}^{O_2^{p_2}}) = \sum_{k=1}^{b_3} \sum_{j=1}^{b_2} \sum_{i=1}^{b_1} \frac{(h_{ijk}^{O_1^{p_1}} - h_{ijk}^{O_2^{p_2}})^2}{(h_{ijk}^{O_1^{p_1}} + h_{ijk}^{O_2^{p_2}})}, \quad (2.54)$$

稱之整體直方圖比對，其中直方圖的維度  $d = 3$ ，因為本論文中之特徵類別數為 3。

要比對兩物體某姿態下的所有區域直方圖，以計算兩物體的整體相異度，本論文結合了 Swain and Ballard 提出的 intersection measurement (2.51)式、Kullback 及 Leibler 提出的 Kullback-Leibler divergence(2.52)式與  $\chi$ -divergence (2.53)式，形成了下式：

$$D(O_2^{p_2} || O_1^{p_1}) = \sum_{\ell_2=1}^{N_2} \min_{(R_{\ell_1} \subset O_1^{p_1})} \{ \chi(\mathbf{H}_{\ell_1}^{O_1^{p_1}}, \mathbf{H}_{\ell_2}^{O_2^{p_2}}) \}, \quad (2.55)$$

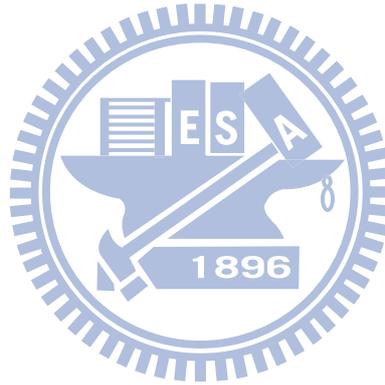
(2.54)式稱之部分直方圖比對，其中  $\chi \left( \mathbf{H}_{\ell_1}^{O_1^{P_1}}, \mathbf{H}_{\ell_2}^{O_2^{P_2}} \right)$  利用(2.53)式之定義，獲得下式：

$$\chi \left( \mathbf{H}_{\ell_1}^{O_1^{P_1}}, \mathbf{H}_{\ell_2}^{O_2^{P_2}} \right) = \sum_{k=1}^{b_3} \sum_{j=1}^{b_2} \sum_{i=1}^{b_1} \frac{\left( h_{ijk}^{R_{\ell_1}} - h_{ijk}^{R_{\ell_2}} \right)^2}{\left( h_{ijk}^{R_{\ell_1}} + h_{ijk}^{R_{\ell_2}} \right)}, \quad (2.56)$$

(2.55)式可計算出物體某姿態下  $O_2^{P_2}$  的某區域  $R_{\ell_2}$  相對於物體某姿態下  $O_1^{P_1}$  的某區域  $R_{\ell_1}$  之相異程度。首先，先須計算  $O_2^{P_2}$  的第  $\ell_2$  區域  $R_{\ell_2}$  的直方圖  $\mathbf{H}_{\ell_2}^{O_2^{P_2}}$  與  $O_1^{P_1}$  每個區域的直方圖之相異程度，即區塊相異度。

$$\min_{(R_{\ell_1} \subset O_1^{P_1})} \left\{ \chi \left( \mathbf{H}_{\ell_1}^{O_1^{P_1}}, \mathbf{H}_{\ell_2}^{O_2^{P_2}} \right) \right\} = \min \left\{ \chi \left( \mathbf{H}_1^{O_1^{P_1}}, \mathbf{H}_{\ell_2}^{O_2^{P_2}} \right), \dots, \chi \left( \mathbf{H}_{N_1}^{O_1^{P_1}}, \mathbf{H}_{\ell_2}^{O_2^{P_2}} \right) \right\}, \quad (2.57)$$

然後取出最小值，重覆此步驟於  $O_2^{P_2}$  的每個區域，再將每次所得之最小值總合起來，會得到兩物體的區塊相異度總和，如(2.56)式所定義的。



### 第三章 系統流程說明

本章分成六節，先說明整個三維物體分割與辨識系統的流程，再詳細說明流程中每個演算法的參數設定方式。3.1 節介紹整個三維物體分割與辨識的系統架構，使用在實際拍攝之複雜場景上。3.2 節說明使用適應性中值濾波以及移動式最小平方法來作為處理三維雜訊之目的；3.3 節說明如何分割深度影像，當深度影像是一個複雜場景時，將場景分割成許多單一場景並分離每個單一場景中的不同物體；3.4 節介紹如何偵測物體的邊緣，並且使邊緣形成封閉邊界，稱為物體的邊緣圖像(edge map)；3.5 節介紹對單一物體之深度影像使用區域成長法，獲得該物體每個區塊的區域特徵，並且用多維度直方圖統計；3.6 節介紹兩階段式的辨識系統來進行三維物體辨識。

#### 3.1 辨識系統架構

本論文的系統架構是應用在實際取得多物體在同一個場景中之深度資訊，此資訊是利用雷射測距儀去拍攝真實世界中的物體，再利用本論文提出的系統去處理雜訊、分割物體、抽取特徵，最後對每個物體進行辨識，以下的章節將介紹系統的每個方塊之細節以及參數設定。：

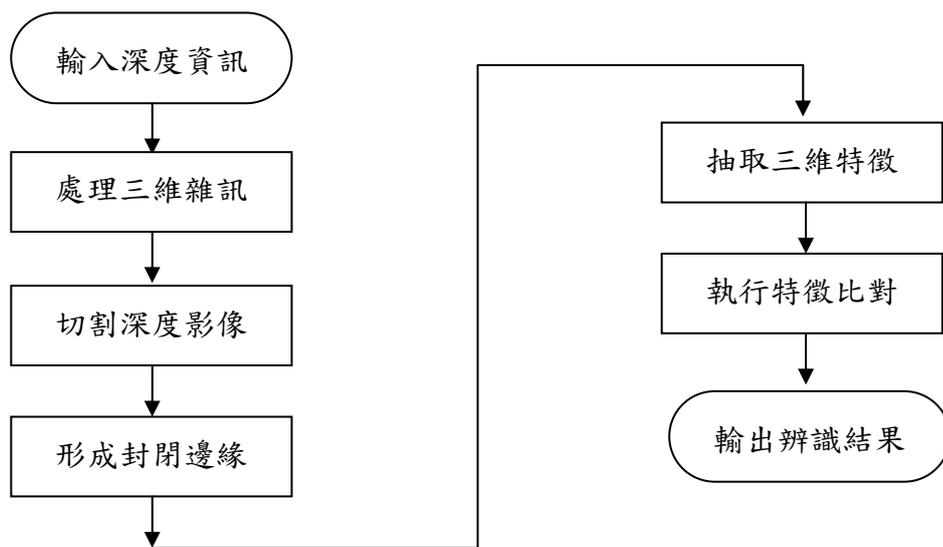


圖 3-1 應用於實際拍攝的深度資訊之整體系統架構

此辨識系統架構中，輸入是經由雷射測距儀實際拍攝所取得的，考慮場景中物體不只一個，所以必須要將每個單一物體各自分離出來，所以有切割深度影像這個步驟，之後再針對每個單一物體抽取出三維特徵，然後進行辨識，其中關於資料庫的組成，是先取得單一物體的深度資訊，進行同樣的步驟，最後抽取此單一物體的三維特徵，作為進行比對特徵時使用。

## 3.2 三維雜訊處理

在本論文中，為了要有效的處理三維雜訊，先使用 2.2.1 節的適應性中值濾波器，再使用 2.2.2 節的移動式最小平方法來處理雜訊，其中適應性中值濾波器所允許的最大遮罩尺寸  $S_{\max} = 5$ ，目的是為了在使用移動式最小平方法來處理雜訊時，必須要計算法向量，為了提高法向量的準確性，所作的前處理，之後再使用移動式最小平方法來更新深度資訊，圖 3-2 為處理深度影像中的三維雜訊之系統架構。

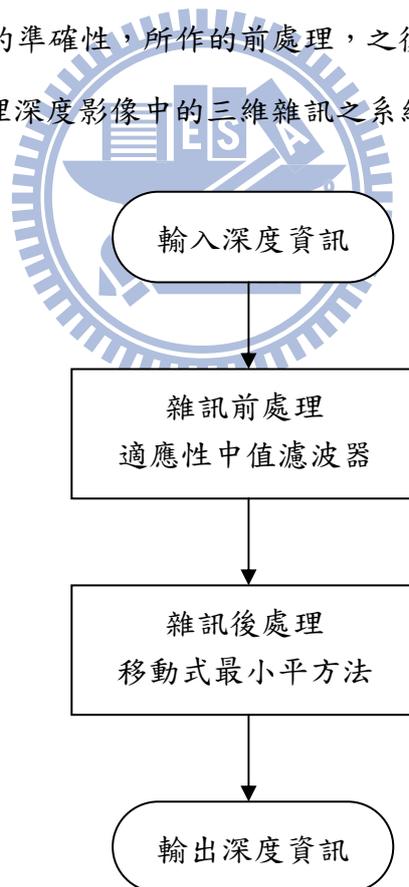


圖 3-2 深度資訊之雜訊處理架構

### 3.2.1 雜訊前處理

在三維雜訊的處理，本論文先用 2.2.1 節的適應性中值濾波器來處理深度資訊中的深度值，即  $z$  方向的深度值，之後再根據下式：

$$\tilde{x} = x \cdot \left(\frac{\tilde{z}}{z}\right), \tilde{y} = y \cdot \left(\frac{\tilde{z}}{z}\right), \quad (3.1)$$

來更新  $x$  方向及  $y$  方向的座標值，(3.1)式是根據雷射測距儀的取像原理所得的式子，由於雷射測距儀原理是將雷射測距儀放在球坐標系的原點，測得物體上的點與雷射測距儀之兩點距離  $r$ ，再根據下式：

$$x = r \sin \theta \cos \varphi, y = r \sin \theta \sin \varphi, z = r \cos \theta, \quad (3.2)$$

將  $r$  轉換成卡氏座標  $(x, y, z)$ ，並假設用雷射測距儀測量與物體上的點之兩點距離  $r$  有誤差，但  $\theta$  與  $\varphi$  為測距儀內部設定參數，所以沒有誤差，故對  $z$  方向的深度值做處理後，可得(3.1)式來更新  $x, y$  值。

### 3.2.2 雜訊後處理

本論文利用一個遮罩  $S_{xy} = \{\mathbf{p}_{11}, \mathbf{p}_{st}, \dots, \mathbf{p}_{ss}\}$ ，計算 2.2.2 節的移動式最小平方法，其大小為  $S = 7$ ，遮罩下所對應的中心點為  $\mathbf{p}_{st}$ ，遮罩移動於深度資訊陣列  $\mathbf{P} = [\mathbf{p}_{ij}]_{m \times n}$  中的每一個點  $\mathbf{p}_{ij} = \mathbf{p}_{st}$ ，而  $\mathbf{p}_{st} = \mathbf{p}_{44}$ 。對於點集合  $\mathbf{P}$  不進行重新取樣的動作，即  $\mathbf{P} = S_{xy} = \{\mathbf{p}_{11}, \mathbf{p}_{12}, \dots, \mathbf{p}_{st}, \dots, \mathbf{p}_{77}\} = \mathbf{R}$ 。

在 2.2.3 節的步驟一，計算每點  $\mathbf{p}_{ij}$  的法向量  $\mathbf{n}_{ij}$  時，利用的是 2.1.2 節所提出演算法，並且不採取任何加權，即  $w_{st} = 1$ 。在 2.2.3 節中的步驟四，根據(2.16)式，本論文令一個最高次  $n = 3$  的區域曲面來近似物體的表面，如下式所示：

$$S_{\bar{p}} = g(u_{st}, v_{st}) = \sum_{a=0, b=0}^{a+b \leq 3} c_{ab} u_{st}^a v_{st}^b = c_{30} u_{st}^3 + \dots + c_{01} v_{st} + c_{00} \quad (3.3)$$

總共 10 個係數， $\mathbf{C} = [c_{30} \ c_{21} \ c_{12} \ \dots \ c_{20} \ \dots \ c_{01} \ c_{00}]^T$ ，再 2.2.2 節之概念來計算出  $\mathbf{C}$ ，

可得(3.4)式、(3.5)式，其中高斯函數 $\theta(\| \mathbf{p}_{st} - \mathbf{q} \|) = e^{\frac{-\| \mathbf{p}_{st} - \mathbf{q} \|^2}{h^2}}$ 。在本論文中用電腦模擬其深度資訊辨識時，令 $h = 1$ ，實際取得深度資訊時，令 $h = 0.1$ ， $h$ 的大小是根據最後希望曲面的平滑程度所設定；一般而言， $h$ 的決定是根據深度資訊中所有相鄰兩點的距離值之平均值以及雜訊嚴重的程度來決定，故在本論文中，電腦模擬的雜訊處理架構中的 $h = 1$ ，實際拍攝的雜訊處理架構中的 $h = 0.1$ 。

欲計算  $\mathbf{C}$  使得可滿足(2.10)式，可得(3.4)式  $\mathbf{F} = \boldsymbol{\theta} \mathbf{G} \mathbf{C}$

$$\begin{bmatrix} f_{11} \\ \vdots \\ f_{77} \end{bmatrix}_{7^2 \times 1} = \begin{bmatrix} \theta_{11} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \theta_{77} \end{bmatrix}_{7^2 \times 7^2} \begin{bmatrix} u_{11}^3 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ u_{77}^3 & \cdots & 1 \end{bmatrix}_{7^2 \times 10} \begin{bmatrix} c_{30} \\ \vdots \\ c_{00} \end{bmatrix}_{10 \times 1}, \quad (3.4)$$

$\boldsymbol{\theta}$  是一個對角矩陣，其矩陣之元素  $\theta_{st}$  是每群點集合內的點  $\mathbf{p}_{st}$  所對應的加權係數， $\mathbf{G}$  是多項式值所對應的矩陣，最後可推導出下式：

$$\mathbf{C} = (\mathbf{G}^T \boldsymbol{\theta} \mathbf{G})^{-1} (\mathbf{G}^T \boldsymbol{\theta} \mathbf{F}), \quad (3.5)$$

最後根據(2.18)式及(2.19)式，可得  $\mathbf{p}_{ij}$  投影到區域曲面上的點  $\mathbf{q}_{ij}$ 。

$$\mathbf{q}_{ij} = \mathbf{p}_{ij} + (t + c_{00}) \mathbf{n}_{ij}, \quad (3.6)$$

將遮罩移動於深度資訊陣列中的每個點，進而更新該點位置，可解決因雷射測距儀之測量誤差所形成的雜訊，以兩個不同物體的深度資訊為例，實際拍攝的兩個物體的深度影像，如圖 3-3 所示，有雜訊影響的法向量  $z$  方向分量，如圖 3-4 所示，經過雜訊處理後的法向量  $z$  方向分量，圖 3-5 所示。



圖 3-3 深度影像 (a)椅子 (b)電腦



圖 3-4 去除雜訊前的法向量之z方向分量 (a)椅子 (b)電腦

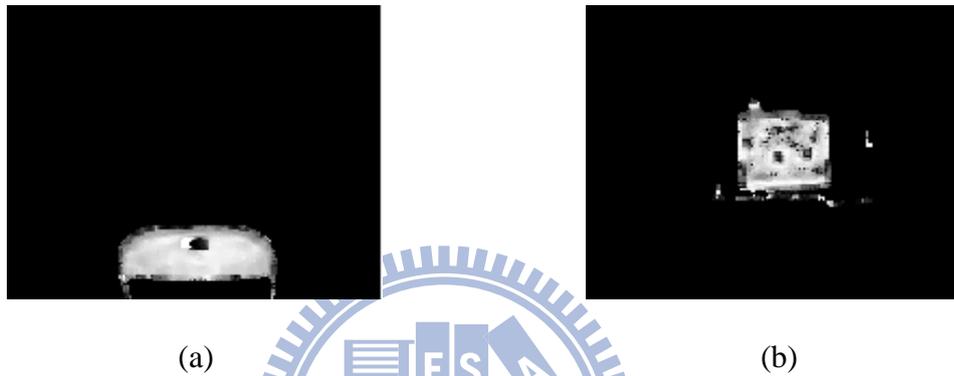


圖 3-5 去除雜訊後的法向量之z方向分量 (a)椅子 (b)電腦

可以看出在沒有經過雜訊處理前，兩物體的法向量之z方向分量之影像有黑點，表示z方向的法向量是不連續的，物體的表面是連續的，所以法向量應該也是連續分布的，若法向量不連續，則表示物體表面呈現非連續分布，即三維雜訊，後來經過雜訊處理後，由圖 3-5 可以看得出來，表面法向量的變化變得連續許多，表示此三維雜訊處理架構是必須的。

### 3.3 深度影像切割

為了要辨識一個複雜場景的多個物體，有必要將場景內的多個物體各自獨立開來，形成單一物體，再來對個別的單一物體進行三維辨識，所以先用 2.3.1 的多重臨界值法，依照物體表面深度的不連續來分割場景，形成單一場景，例如：前景、中景、後景等，之後再用 2.3.2 連通量分析，在每一個單一場景中利用深度影像中的像素間連通性來分離物體。假設直接只有利用深度影像中的像素間連通性來分離物體，可能發生不同物體

彼此遮蔽住，但卻以為是同一個物體的情形，圖 3-6 為切割深度影像的系統架構。

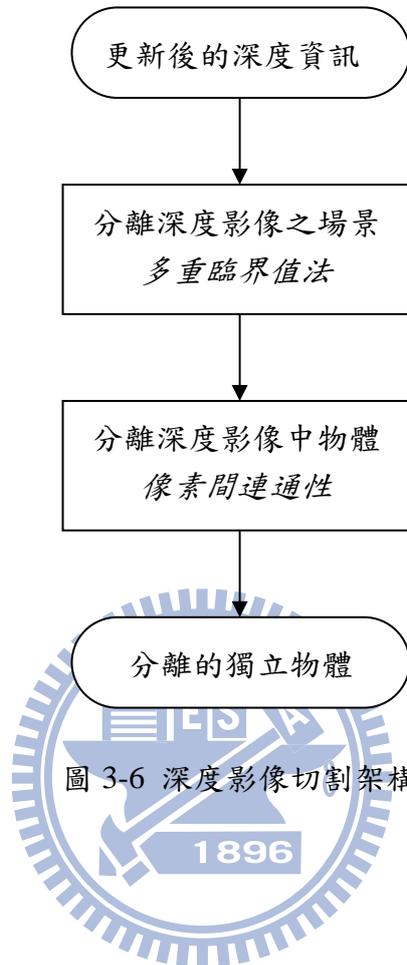
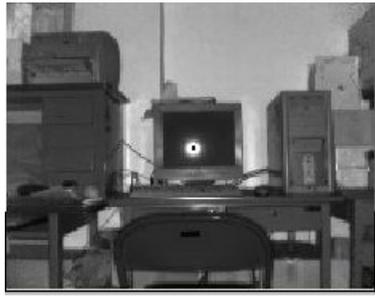


圖 3-6 深度影像切割架構

### 3.3.1 場景分離

在本論文中，將使用多重臨界值法來使深度影像分離成不同的場景，由 2.3.1 節的 (2.2) 式，本論文設定直方圖的初始樣本區間個數  $b = 2^k, k = 2$ ，然後增加  $k$ ，直到所找到的局部最小值個數大於  $T = 4$ ，將單一場景的個數限制住 5 個以下，如 2.3.1 節的圖 2-3 所示，複雜場景很大，可將  $T$  設定大一點，但若  $T$  過大時，會使得原本的物體可能被切割成在不同的單一場景下，如圖 3-7 所示，以  $T = 6$  為例，共有七個單一場景。



(a)



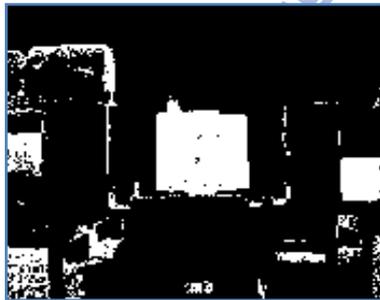
(b)



(c)



(d)



(e)



(f)



(g)



(h)

圖 3-7 T=6 時，由真實場景中所分離的物體  
(a) 真實場景 (b)物體 1 (c)物體 2 (d)物體 3 (e)物體 4 (f)物體 5 (g)物體 6 (h)物體 7

可看出設定 $T = 6$ 來分割圖 3-7(a)，所得的單一場景過多，所以  $T$  過大會使單一

場景過多，同一個物體都被分割成很多不同的單一場景中，例如：圖 3-7(c)與圖 3-7(d)中的印表機被分割成在兩個不同場景中，圖 3-7(g)與圖 3-7(h)中的牆面被分割成兩個不同的場景中。

### 3.3.2 物體分離

在本論文中，利用深度影像  $\mathcal{R}$  中的像素間連通性來分離物體，利用 8 連通成份，並將同樣的 8 通成份作標記，使有相同標記的像素點變成一區域  $\mathcal{R}_i$ ，此區域就可被分離成單一物體；除此之外，在本論文中，為了避免因為雜訊的影響，導致區域數過多，使得辨識速度的變慢，故設置了一個濾波器，使得某區域中的點數佔整個場景中的點數比例過低時，刪除掉此區域，如(3.1)式。

$$\text{If } \frac{\text{num}(\mathcal{R}_i)}{\text{num}(\mathcal{R})} \leq \text{Th}, \text{ then } \mathcal{R}_i = \{\emptyset\} \quad (3.7)$$

當  $\text{Th} = 0$  時，即沒有過濾時，經過兩步驟的分割，可將圖 3-7(a)分成 55 個物體， $\text{Th} = 0.01$ ，經過這二階段的分割，最後可分成 10 個物體，如圖 3-8 所示，若  $\text{Th} = 0.03$ ，可分成 6 個物體，在本論文中，令  $\text{Th} = 0.05$ ，可分成 5 個物體，如圖 3-9 所示，以此類推。  
Th 的大小可由雷射測距儀的解析度以及測距範圍來決定。

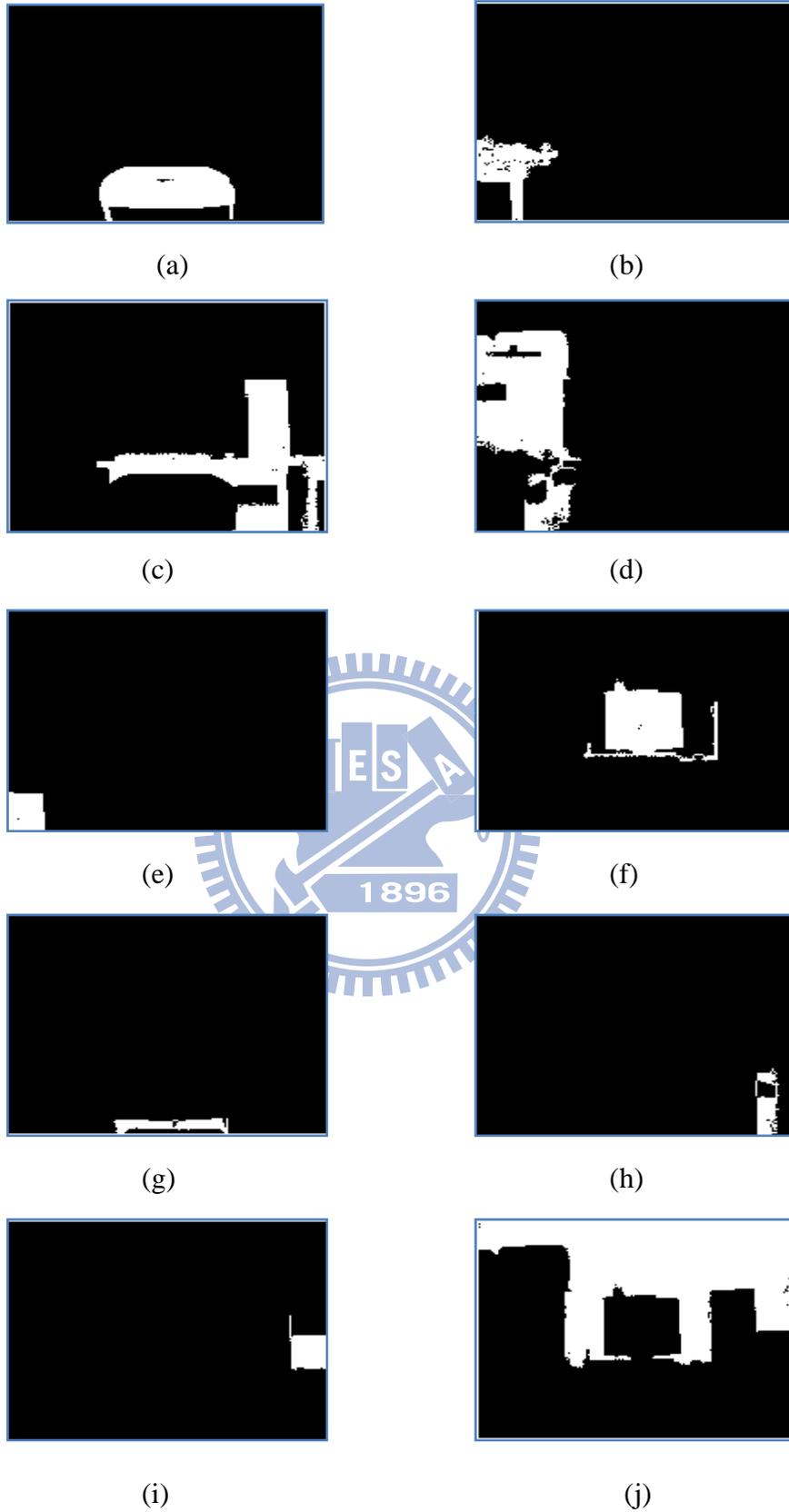


圖 3-8  $Th=0.01$  時，由真實場景中所分離的物體。(a)物體 1 (b)物體 2  
(c)物體 3 (d)物體 4 (e)物體 5 (f)物體 6 (g)物體 7 (h)物體 8 (i)物體 9 (j)物體 10

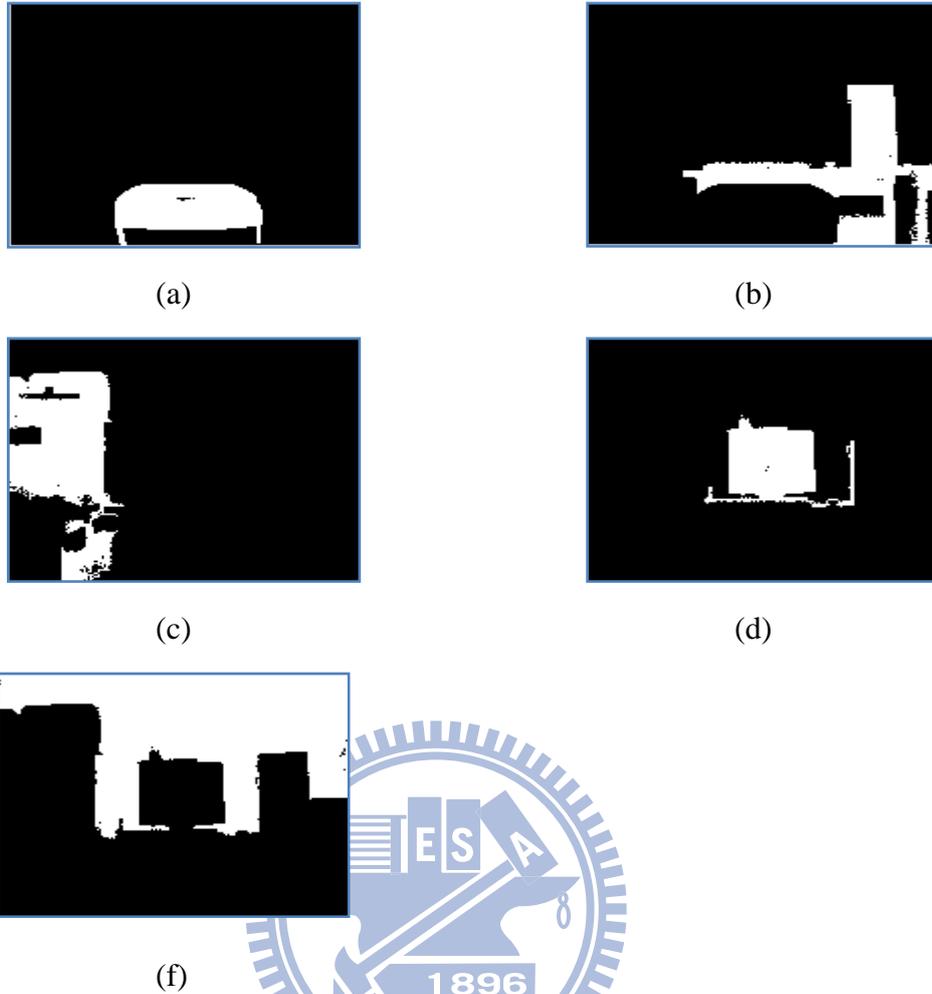


圖 3-9  $Th=0.05$  時，由真實場景中所分離的物體  
 (a) 物體 1 (b) 物體 2 (c) 物體 3 (d) 物體 4 (e) 物體 5

### 3.4 封閉邊緣處理

一般而言，在深度影像中分離完場景後，然後再進一步的分離出個別物體，就可以抽取出單一物體的特徵來進行三維物體的比對，但是考慮單一物體的比對時，往往沒法處理當物體彼此有遮蔽情形的時候或是深度資訊受到大量的雜訊影響的時候，故本論文提出了一個根據單一物體之表面變化特性，包括表面上的點之深度變化與表面上的點之法向量變化，依據某規則有效地切割此單一物體的表面，形成一個封閉的邊緣圖像 [28][32]，每一個封閉邊緣形成一個區塊，其中每個區塊會形成一種基礎曲面，然後對每個區塊抽取其特徵，最後去比對此物體的每個區塊特徵，如此一來，即使物體有遮蔽情形發生或是某部分有嚴重雜訊，仍可確保其他區塊的比對是正確的。圖 3-15 為形成

邊緣影像並將邊緣影像的邊緣使之封閉，形成邊緣圖像的系統架構。

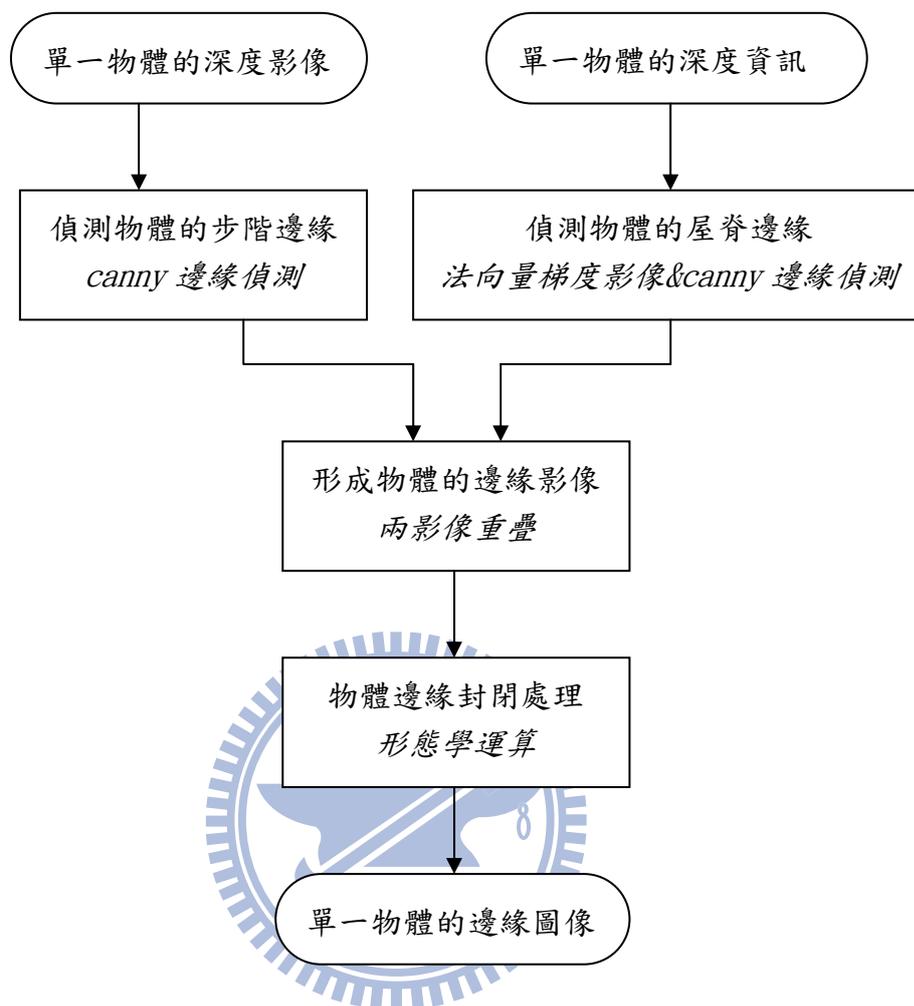


圖 3-10 邊緣處理架構

### 3.4.1 邊緣影像

在本論文中，使用 2.4.1 節所介紹的 Canny 邊緣偵測器來偵測出深度影像的邊緣，邊緣偵測器原本偵測強度影像中像素的強度值變化，改成偵測深度影像中像素的深度變化，並設定 Canny 邊緣偵測器中的 $\sigma = 2$ ，見(2.29)式，最後形成步階邊緣影像。圖 3-11 為實際拍攝中兩個不同物體的步階邊緣。



圖 3-11 物體的步階邊緣影像 (a)椅子 (b)電腦

除了找尋因為深度變化所造成的邊緣，也找尋表面法向量變化所造成的邊緣，即屋脊邊緣，本論文改良 1989 N. Yokoya 以及 M. Levine [29]所提出的屋脊邊緣偵測概念，因為 N. Yokoya 計算屋脊邊緣的演算法會受到三維雜訊的影像，導致效果過差。首先，利用 2.1.2 節來計算深度資訊中每點的法向量，其中選擇  $5 \times 5$  的遮罩且遮罩內係數均為 1，見(2.1)式，然後使用 2.4.2 節來計算出法向量變化所造成的梯度影像，之後計算法向量之梯度影像時，再使用  $3 \times 3$  遮罩且遮罩內係數均為 1 於梯度影像上，使用  $3 \times 3$  的大小是避免遮罩尺寸過大，使得梯度變化模糊；此外，在(2.31)式中，選擇  $\gamma = 1$ ，避免後來屋脊細節過多。形成法向量的梯度影像後，再使用 Canny 邊緣偵測器來偵測法向量之梯度影像，但在實際拍攝的架構，Canny 邊緣偵測器中的  $\sigma = 2$ ，因為實際拍攝所得的深度資訊，解析度較低且受到雜訊影響程度較高，故  $\sigma = 2$ 。

圖 3-12 為實際拍攝中兩個不同物體的屋脊邊緣。



圖 3-12 物體的屋脊邊緣影像 (a)椅子 (b)電腦

最後將這步階影像與屋脊影像重疊，就形成了邊緣影像，如圖 3-13 所示。



圖 3-13 物體的邊緣影像(a)狗 (b)機器

### 3.4.2 邊緣圖像

形成了單一物體的邊緣影像後，必須使邊緣影像中的開放邊緣成為封閉邊緣，之後才能使用區域成長法，取出每一個封閉區塊內的特徵，做為區塊辨識比對用。本論文使用一連串的形態學運算將邊緣影像經過形態學閉合、形態學收縮，最後再移除掉每個區塊內部的孤立像素，即完成了邊緣圖像，邊緣圖像中的每個區域是封閉區域，使邊緣封閉的目的是為了之後對每個封閉區域使用區域成長法，以取得每個區域對應的特徵。

首先要對邊緣影像進行形態學閉合，使用 2.4.3 節所介紹過的概念，使用  $3 \times 3$  的結構元素，利用(2.32)式，最後經過形態學閉合的影像，如圖 3-14 所示。



圖 3-14 形態學閉合後的邊緣影像 (a)椅子 (b)電腦

然後，再對經過形態學閉合後的影像，利用 2.4.3 節介紹的形態學收縮，使用  $3 \times 3$  大

小的結構元素收縮邊緣使之細化，將有洞的邊緣將收縮成環，沒洞的邊緣將收縮成點，如圖 3-15 所示。



圖 3-15 形態學收縮後的邊緣影像 (a)椅子 (b)電腦

最後，利用 2.3.2 節連通量分析的概念，移除內部孤立像素，該像素的 4 鄰域像素皆是背景像素，則移除該孤立像素，最後所得到的就是邊緣圖像，如圖 3-16 所示。



圖 3-16 物體的邊緣圖像 (a)椅子 (b)電腦

### 3.5 三維特徵抽取

經由雷射測距儀所取得深度影像，可以明確地描述物體的外形，故將用物體表面之幾何描述來做為三維物體辨識用之特徵。但是利用雷射測距儀來測量三維物體時，會因為測量的視角不同，造成會得到三維物體不同的姿態，為了要避免視角差異下，造成物體姿態的不同，所以本論文將用區域特徵(local feature)，來做為比對三維物體時的特徵，因為區域特徵跟測量視角的相關性較低，為了加快辨識三維物體的速度，故將區域特徵

利用統計的方式形成整體直方圖[32]，然後比對單一物體所對應的整體直方圖；但是，若物體被遮蔽的情形發生時，為了提高辨識率，本論文提出了依照物體表面幾何特性的邊緣偵測系統，將物體的表面切割成許多區塊，而每個區塊所對應的區域特徵可利用統計的方式形成部分直方圖，以解決在遮蔽下造成辨識率下降的問題，同時也可解決測量視角略有差異之影響。圖 3-17 是將單一物體的深度資訊，抽取出區域特徵並且統計成整體直方圖與部分直方圖的系統架構。

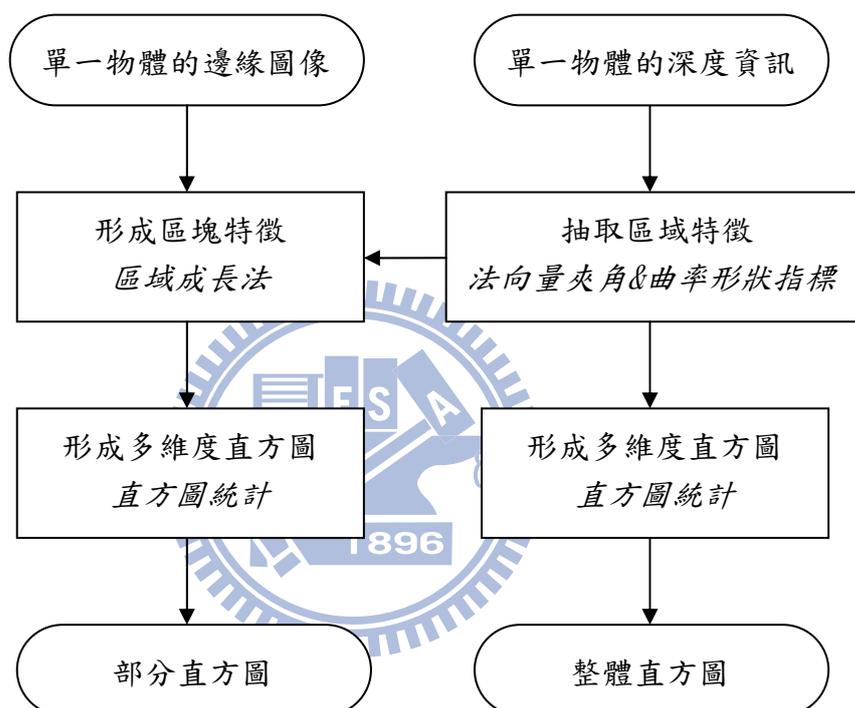


圖 3-17 特徵抽取架構

### 3.5.1 整體特徵統計

由每個物體的深度資訊，計算其區域特徵，主要使用的區域特徵是根據三維物體表面的幾何特性，例如：表面的法向量以及表面的曲率變化。在本論文中，用來比對的特徵有四個，如 2.5 節的所介紹的，利用可獲得每點的深度變化  $d_p$ ，如圖 3-18 所示，並利用(2.38)式可得曲面上的每點法向量分量之夾角  $\phi_p$ ，如圖 3-19 所示，以及  $\theta_p$ ，如圖 3-30 所示。最後再使用 2.5.3.節所介紹的曲率之形狀指標  $SI_p$ ，如圖 3-31 所示。在本論文使用  $5 \times 5$ 的遮罩去計算遮罩下所對應的點，這些點所形成的曲面之曲率，最後再

計算曲率的形狀指標，利用(2.44)式，其中遮罩避免選太小，因為是利用內插法去計算中心點與附近點的梯度變化，若選取的點太少，會造成計算曲率有誤差，若太大會造成計算速度變慢。圖 3-18 至圖 3-21 是實際拍攝時兩個不同物體的區域特徵。



(a)

(b)

圖 3-18 物體的整體區域特徵影像  $D$  (a)椅子 (b)電腦



(a)

(b)

圖 3-19 物體的整體區域特徵影像  $\Phi$  (a)椅子 (b)電腦



(a)

(b)

圖 3-20 物體的整體區域特徵影像  $\theta$  (a)椅子 (b)電腦



圖 3-21 物體的整體區域特徵影像  $S_I$  (a)椅子 (b)電腦

### 3.5.2 區塊特徵統計

抽取出單一物體的表面特徵後，繼續處理由 3.4.2 節獲得該單一物體的邊緣圖像，利用 2.4.3 節所介紹的區域成長法演算法來獲得該邊緣圖像中每個封閉區域的資訊，取出每一個封閉區域內的點所對應的區域特徵，形成區塊特徵，如圖 3-22 所示，邊緣圖像中的區塊都經過了區域成長法，然後顯示不同的灰階值。



圖 3-22 物體的區塊圖像 (a)椅子 (b)電腦

### 3.5.3 直方圖統計

獲得單一物體的整體特徵後，利用 2.6.1 節多維度直方圖所介紹的概念，將整體特徵使用直方圖來進行統計，形成一個三維直方圖，稱之整體直方圖，如(2.49)式，其中整體直方圖為一個正規化直方圖，如(2.50)式；此外，單一物體的區塊特徵，也用 2.6.1 節的概念，使用直方圖統計，也會形成一個三維直方圖，稱之部分直方圖，一個物體可

能被分為數百個區塊，每個區塊都會有個部分直方圖，如(2.51)式。

在本論文中，設定區域特徵依序為 $[d_p \ \phi_p \ \theta_p \ SI_p]^T$ ，所對應的特徵之區間範圍分別為 $[0, L_d] = [0 \ 1]$ 、 $[0, L_\phi] = [0 \ \pi]$ 、 $[0, L_\theta] = [0 \ \pi]$ 、 $[0, L_{SI}] = [0 \ 1]$ ，根據 Gunter Hetzel [32]選定許多不同的區間個數來將三維物體表面的區域特徵作直方圖統計，然後比對直方圖，發現選定區間個數為  $b_d = 4$ 、 $b_\phi = 4$ 、 $b_\theta = 8$ 、 $b_{SI} = 8$ ，會有比較高的辨識率，故本論文區間個數就選定為  $[b_d \ b_\phi \ b_\theta \ b_{SI}]^T = [4 \ 4 \ 8 \ 8]^T$ 。

### 3.6 辨識系統介紹

為了要有效地辨識出深度資訊中的三維物體，有以下幾點問題需要考量：第一，考慮在使用雷射測距儀測量時，所測量的物體姿態跟資料庫的物體姿態大略相同，但是兩者所拍攝的視角有些微差異，也就是如 2.5.1 節所介紹的；第二，假設拍攝的場景中，物體有被遮蔽的情形時，必須考慮遮蔽情形；第三，為了同時提高辨識率且不可過度地犧牲掉辨識速度。基於以上這些因素，本論文採用兩階段式的辨識系統，先比對該未知物體的某姿態之整體直方圖與資料庫中已知物體的某姿態之整體直方圖，然後列出可能是該未知物體的候選清單，再進一步的比對未知物體的某姿態之部分直方圖與候選清單上之物體的部分直方圖，才候選清單中找出相似度最高的物體，即該未知物體。圖 3-1 是未知物體如何與資料庫中的已知物體比對之辨識系統架構。

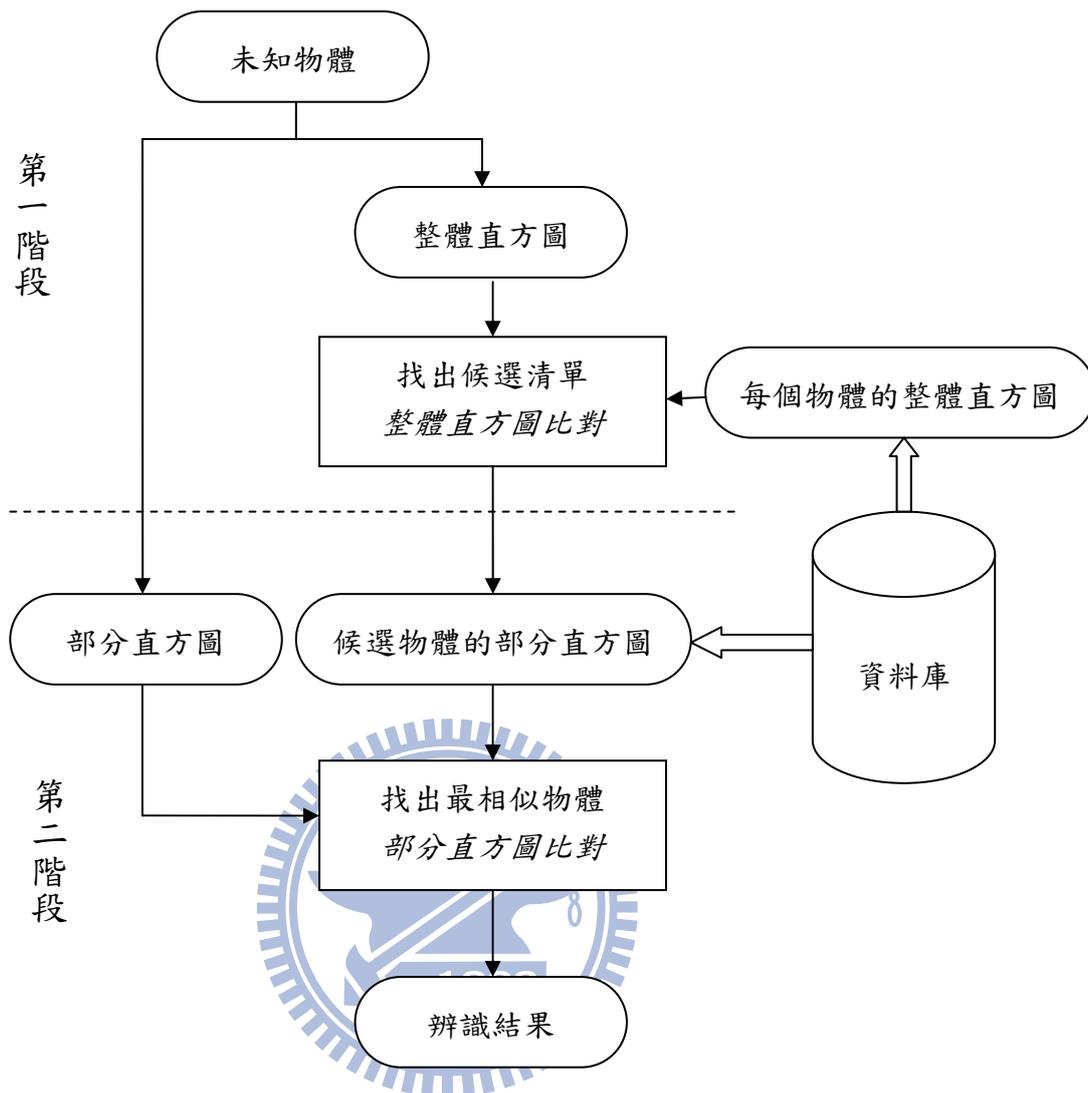


圖 3-23 辨識系統架構

### 3.6.1 整體直方圖比對

首先，利用 2.6.2 直方圖比對的概念中的(2.54)式來比對單一未知物體的整體直方圖與資料庫中每個已知物體的整體直方圖，然後找出其有可能是未知物體  $K$  個。即計算未知物體與資料庫中的每個物體的整體相異度並且進行排序，將兩物體的整體相異度由小排到大，然後將整體相異度較小的前  $K$  個物體將列在候選清單上，作為第二階段辨識使用，其中  $K$  的大小將由資料庫中物體的種類、個數以及視辨識的速度來決定；此外，若是一個複雜場景內有多個未知物體需要辨識時，可利用之前提出的分割架構，每個單一物體從場景中分離出來，然後利用整體直方圖，得知每個未知物體與資料庫中的

每個物體的整體相異度，若某未知物體 $O_{\text{query}}$ 的第 $j$ 姿態 $O_{\text{query}}^{\text{pj}}$ 與某已知物體 $O_k$ 的第 $i$ 姿態 $O_k^{\text{pi}}$ 的整體相異度 $\chi(O_{\text{query}}^{\text{pj}}, O_k^{\text{pi}})$ 小於某個臨界值 $\text{Th}$ ，則該已知物體可以列入後選清單。

在本論文中，在實際拍攝的辨識系統，均是場景中多個單一未知物體和資料庫物體作比對，故選擇兩者整體相異度較小的 $K$ 個物體作為比對，其中資料庫中總共有8種不同的物體，每個物體有13種姿態，故選擇 $K=5$ 。

### 3.6.2 部分直方圖比對

得到辨識清單後，執行更進一步的比對，利用部分直方圖比對，來比對未知物體的部分直方圖與在候選清單上每個已知物體的部分直方圖。首先，計算某姿態下的未知物體之某區塊與某候選物體的某區塊相異度 $\chi\left(\mathbf{H}_{\ell_1}^{O_{\text{query}}^{\text{pj}}}, \mathbf{H}_{\ell_2}^{O_k^{\text{pi}}}\right)$ ，即(2.54)式，然後找出某候選物體中哪個區塊與未知物體的某固定區塊的區塊相異度為最小的，為即(2.57)式，最後利用(2.56)式計算未知物體的每個區塊與某候選物體的最相似區塊之相異程度，即區塊相異度，最後把區塊相異度相加，即為區塊相異度總和 $D(O_{\text{query}} \parallel O_k)$ ，即(2.55)式。最後，找出候選清單上哪個物體與未知物的區塊相異度總和為最低者，為最終辨識結果。

## 第四章 實驗結果與討論

本章分成三節，在 4.1 節將說明實驗設備，包括如何取得深度資訊並利用這些訓練資料來建立資料庫，並介紹作測試資料用的複雜場景之深度資訊，再介紹用來評估三維辨識系統的準則；4.2 節將顯示實驗結果並統計所有實驗數據；4.3 節將對實驗結果作分析與討論。

### 4.1 實驗設備與效能評估

本節分為兩小節，在 4.1.1 節將介紹用於實際拍攝的雷射測距儀之規格，並利用雷射測距儀對多個單一物體取得其深度資訊，作為訓練資料(training data)用，同時建立資料庫，再使用雷射測距儀另外取得多張複雜場景之深度資訊，作為測試資料(testing data)用，用測試資料來驗證本論文提出的三維物體分割與辨識系統；在 4.1.2 節將介紹評估多個不同辨識系統之辨識率的準則，用來比較本論文所提出的兩階段三維物體辨識系統與其他論文提出的辨識系統之效能。

#### 4.1.1 實驗設備

本論文將使用雷射測距儀進行實際拍攝以取得複雜場景之深度資訊，並且同時辨識場景中多個物體，而本論文取得真實場景中深度資訊的雷射測距儀是 MESA-Imaging 公司所出產的 SR3000 如圖 4-1 所示，適合測量室內複雜場景的深度資訊，有效測量範圍 0 到 7.5 公尺，測量誤差小於為 0.3 公分，掃描真實場景後，所獲得的深度資訊為  $176 \times 144 \times 3$  的陣列，每個維度之陣列各儲存該點的  $x, y, z$  值，此雷射測距儀是利用紅外線光源作為測距光源，所以很容易受到物體表面材質之影響，導致測量到的物體表面會有嚴重的雜訊波動，例如：在黑色材質上，雷射測距儀將測不到數據或是很大的誤差，以及在可透光的材質上也有同樣的現象出現，所以此雷射測距儀更適合測試本論文提出的三維物體辨識系統對抗雜訊的強健性。



圖 4-1 雷射測距儀

圖片來源：MESA-Imaging 公司

在本論文中選擇八個日常生活中常見且不同幾何形狀的物體作為資料庫的物體，如表 4-1 所示：

表 4-1 資料庫中的物體

物體 1	物體 2	物體 3	物體 4
CRT 電腦螢幕	玩具人偶	折疊椅	桌子(無抽屜)
物體 5	物體 6	物體 7	物體 8
檯燈	LCD 電腦螢幕	旋轉椅	桌子(有抽屜)

並實際拍攝單一場景下的單一物體，將物體與背景分離後，使用 3.2 節的流程對此物體進行去除雜訊的動作以及再利用 3.3 節 3.4 節 3.5 節之流程來取得該物體之幾何特徵以作為資料庫中的比對特徵；其中，每個物體有 13 個不同測量俯視視角、水平視角、仰視視角，此測量視角即為訓練角度，如下表 4-2 所式，令物體的朝向雷射測距儀之正面為 0 度，將物體由上往下以順時針方向朝著雷射測距儀旋轉以改變其視角，每旋轉 15 度以取得此物體的某一種姿態，每個視角可以取得 13 種姿態，共有三種視角，總計資料庫中每一種物體有 39 種姿態，其中資料庫中的物體表面都具有左右對稱的特性，使得對稱的姿態取得該姿態下的區域特徵後，其區域特徵將會是一模一樣，基於這種特性，所以在旋轉時，最多只需要沿著水平方向旋轉 180 度即可，因為對稱的物體左邊旋轉跟右邊旋轉相同角度後，該姿態下物體有著相同的表面變化，故計算出來的區域特徵是相

同的。

表 4-2 已知物體旋轉角度與姿態對照表

訓練角度 (俯角 45 度)	0	15	30	45	60	75	90	105	120	135	150	165	180
物體姿態	1	2	3	4	5	6	7	8	9	10	11	12	13
訓練角度 (仰角 45 度)	0	15	30	45	60	75	90	105	120	135	150	165	180
物體姿態	14	15	16	17	18	19	20	21	22	23	24	25	26
訓練角度 (z 軸)	0	15	30	45	60	75	90	105	120	135	150	165	180
物體姿態	27	28	29	30	31	32	33	34	35	36	37	38	39
訓練角度 (x 軸)	0	15	30	45	60	75	90	105	120	135	150	165	180
物體姿態	40	41	42	43	44	45	46	47	48	49	50	51	52
訓練角度 (y 軸)	0	15	30	45	60	75	90	105	120	135	150	165	180
物體姿態	53	54	55	56	57	58	59	60	61	62	63	64	65

在本論文中，除了要判定未知物體是資料庫中的是哪個已知物體，也要判定未知物體的姿態是否跟資料庫中的該已知物體某姿態相符，實驗時會拍攝處於某個視角下不同姿態的物體作為輸入，輸入的角度跟一開始訓練資料的角度不同，所以本論文訂定表 4-3 之規則來判定未知物體的姿態，此外再考慮拍攝時，操控變因為物體是否有被遮蔽，所以輸入的測試資料的視角為水平，以變控制物體遮蔽的情形。

當未知物體的水平視角落在該角度範圍下，即為該姿態，此角度範圍為訓練角度的前後 7.5 度，若拍攝物體的角度介於該角度之間，即屬於該姿態，因為物體左右對稱，所以 0

度的範圍只有 0 度~7.5 度，180 度也是同理。例如：當拍攝未知物體的測試角度的範圍介於 22.5 度與 37.5 度之間時，即認定該未知物體的姿態屬於姿態 16，而經過辨識系統後也辨識出未知物體的姿態為姿態 16，則該物體姿態之辨識為正確，依此類推。

最後，資料庫有八種物體，如圖 4-2 所示，而每種物體有 65 種姿態，分別如圖 4-3 至圖 4-10 所示，依序為 CRT 電腦螢幕、玩具人偶、檯燈、摺疊椅、桌子(無抽屜)、LCD 電腦螢幕、旋轉椅、桌子(有抽屜)，其中 CRT 電腦螢幕如圖 4-2 所示，左上第一張圖是該物體的強度影像，然後由左至右，由上至下分別是沿著俯角 45 度拍攝，為圖 4-3(a)、仰角 45 度拍攝為圖 4-3(b)，水平方向且固定 z 軸拍攝為圖 4-3(c)、水平方向且固定 x 軸拍攝為圖 4-3(d)、水平方向且固定 y 軸拍攝為圖 4-3(e)，然後每個角度下將物體依順時針方向依序旋轉 15 度後，所獲得的 39 張深度影像，而以下所有的 520 張深度資訊將成為訓練資料用，並建立特徵資料庫。

表 4-3 未知物體的角度範圍與姿態對照表

角度範圍(z 軸)	0~7.5	7.5~22.5	22.5~37.5	37.5~52.5	52.5~67.5
物體姿態	14	15	16	17	18
角度範圍(z 軸)	67.5~82.5	82.5~97.5	97.5~112.5	112.5~127.5	127.5~142.5
物體姿態	19	20	21	22	23
角度範圍(z 軸)	142.5~157.5	157.5~172.5	172.5~180		
物體姿態	24	25	26		

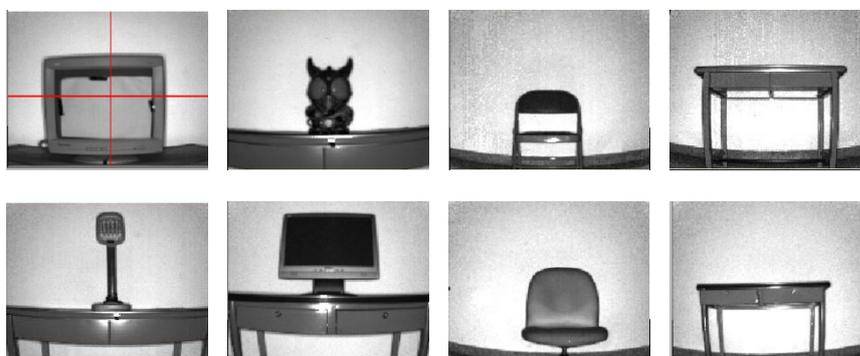
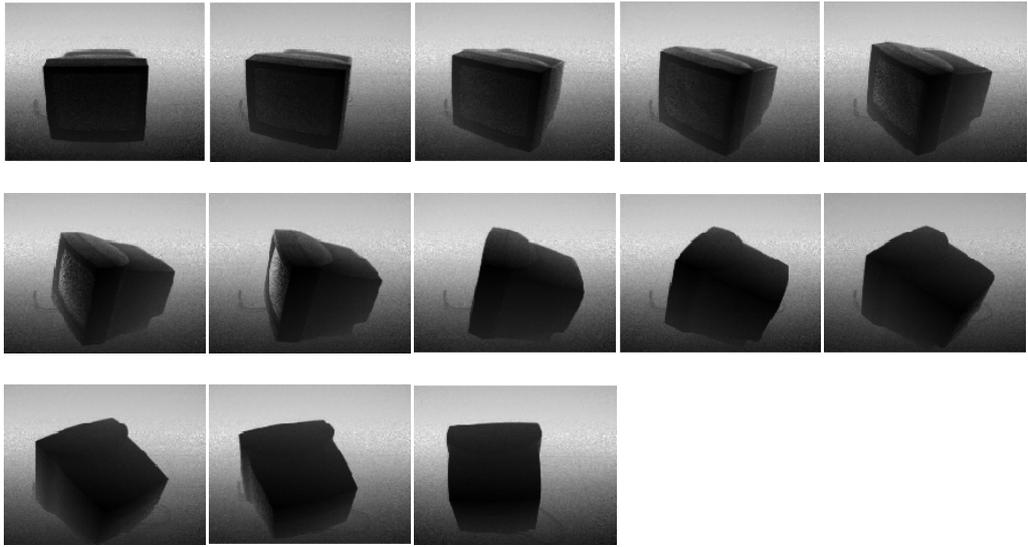
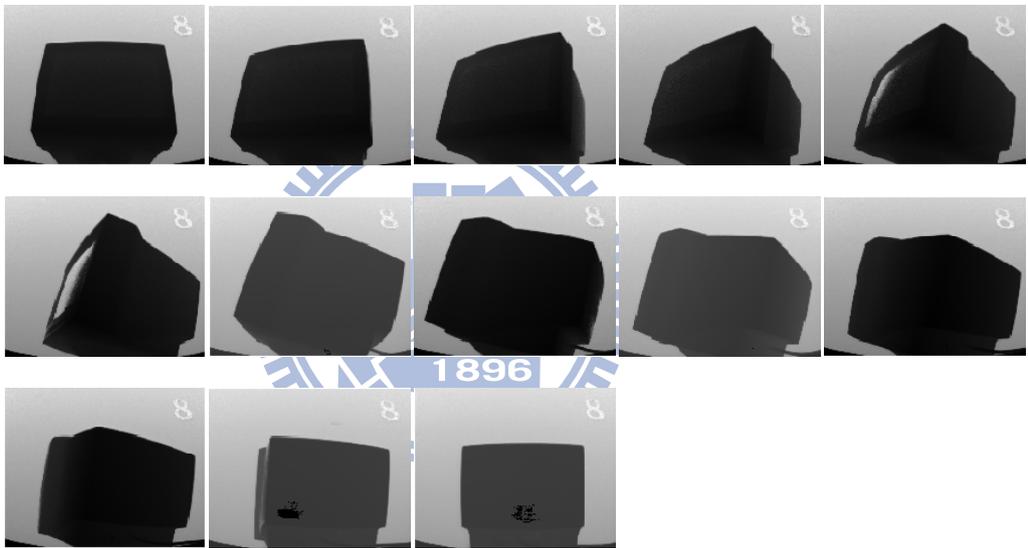


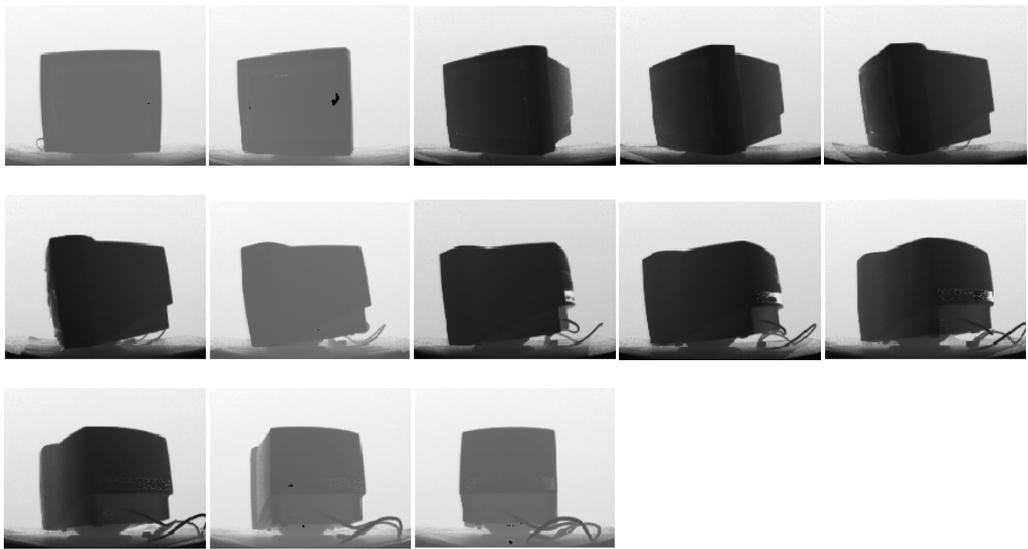
圖 4-2 資料庫中的八個物體



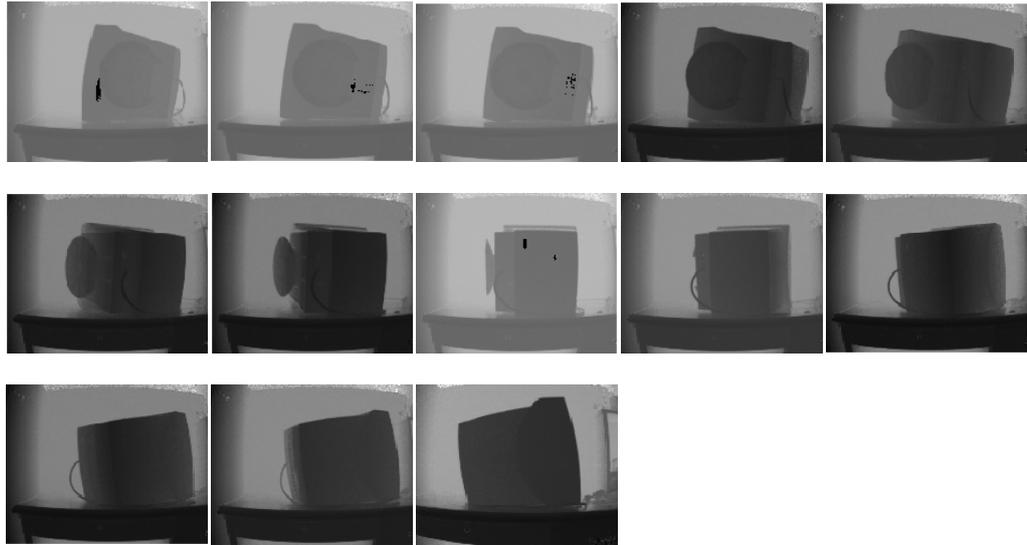
(a)



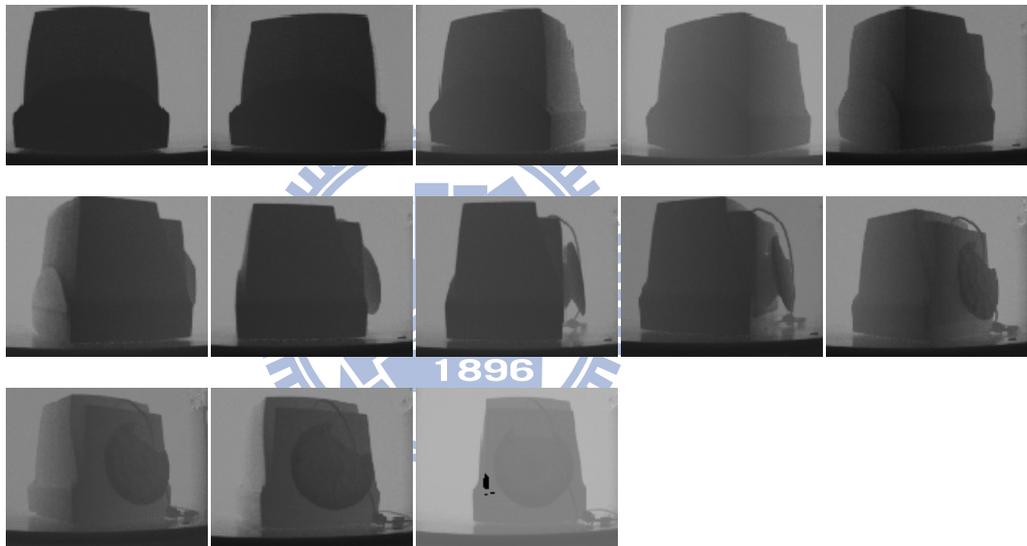
(b)



(c)



(d)

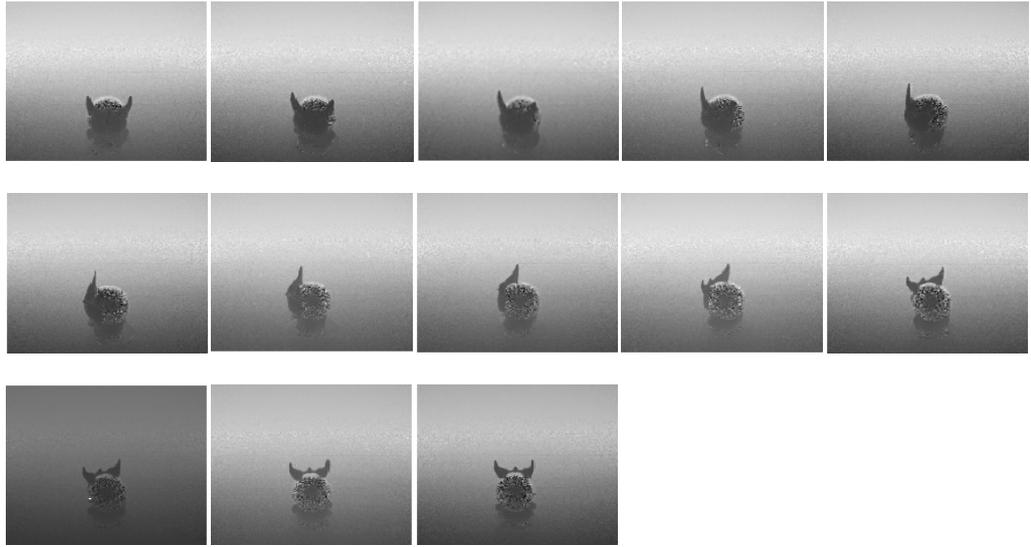


(e)

圖 4-3 CRT 電腦螢幕的強度影像與 39 張姿態的深度影像

(a)俯角 45 度，z 軸固定 (b)仰角 45 度，z 軸固定 (c)水平 0 度，z 軸固定

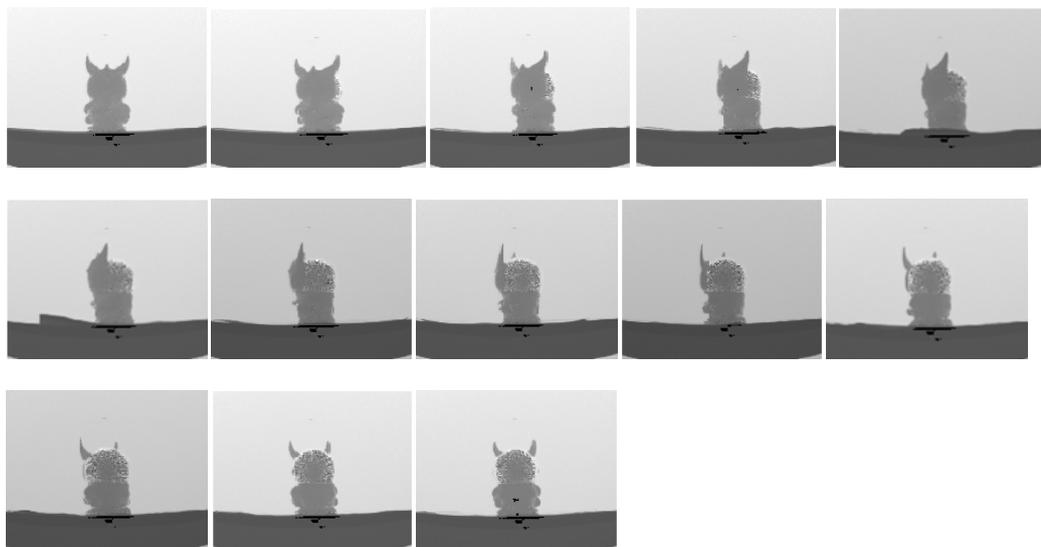
(d)水平 0 度，x 軸固定 (e)水平 0 度，y 軸固定



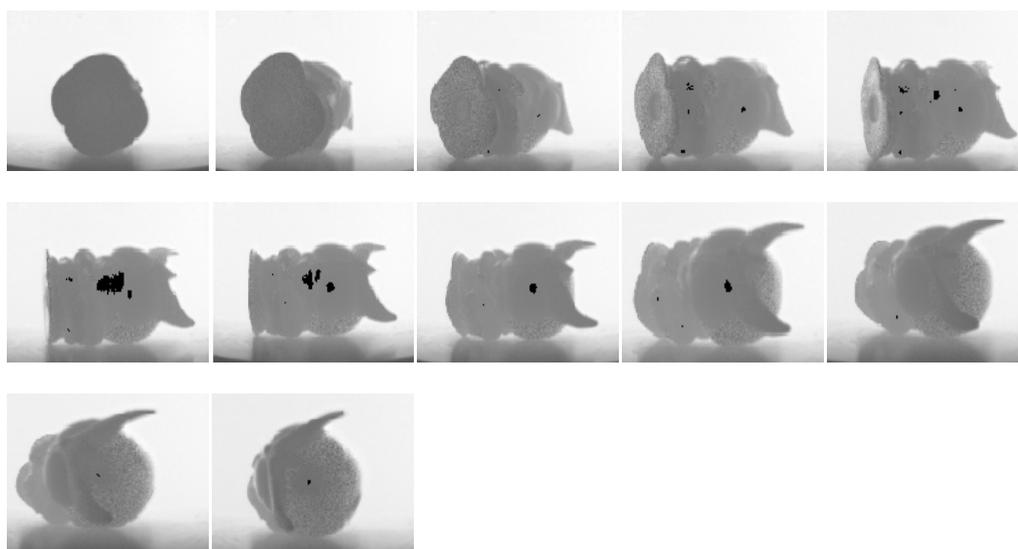
(a)



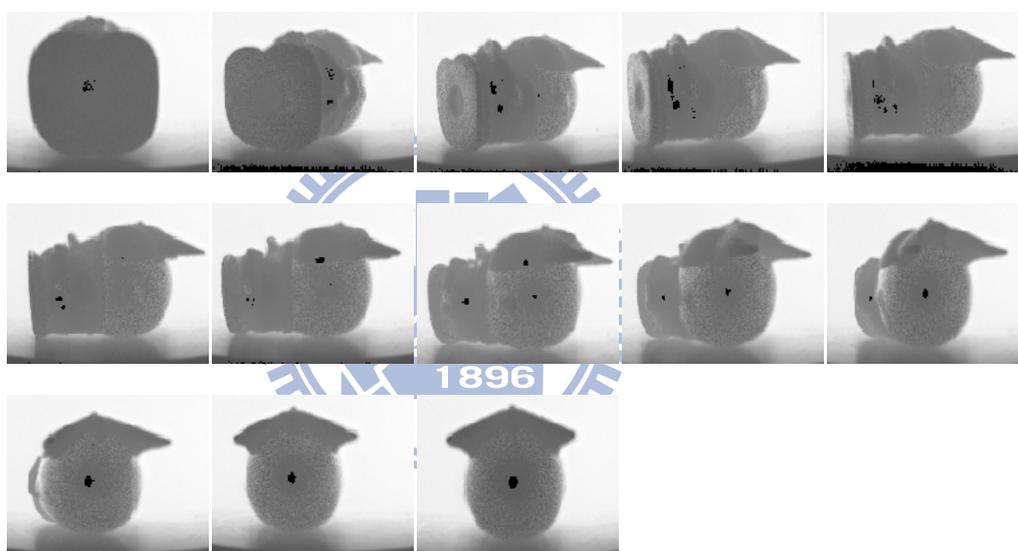
(b)



(c)



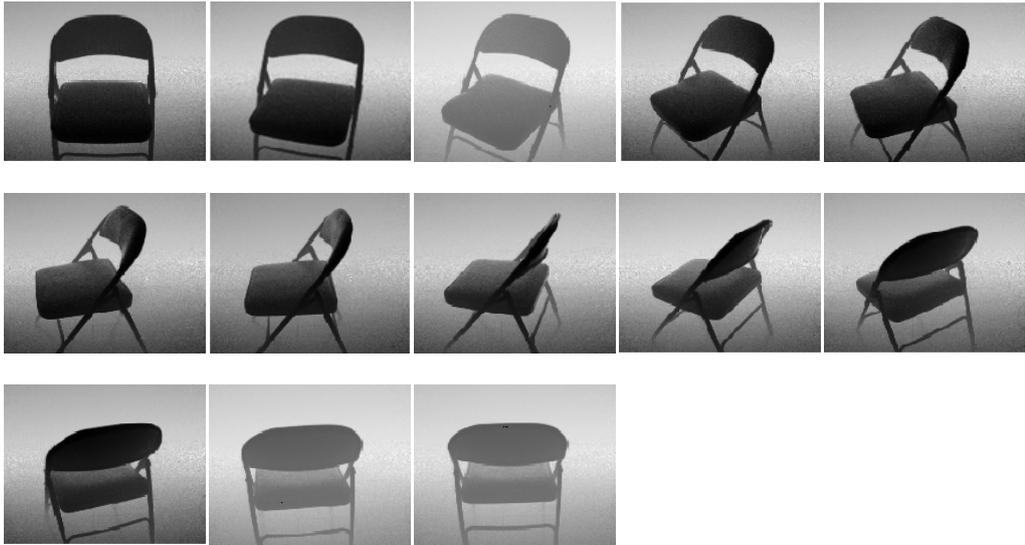
(d)



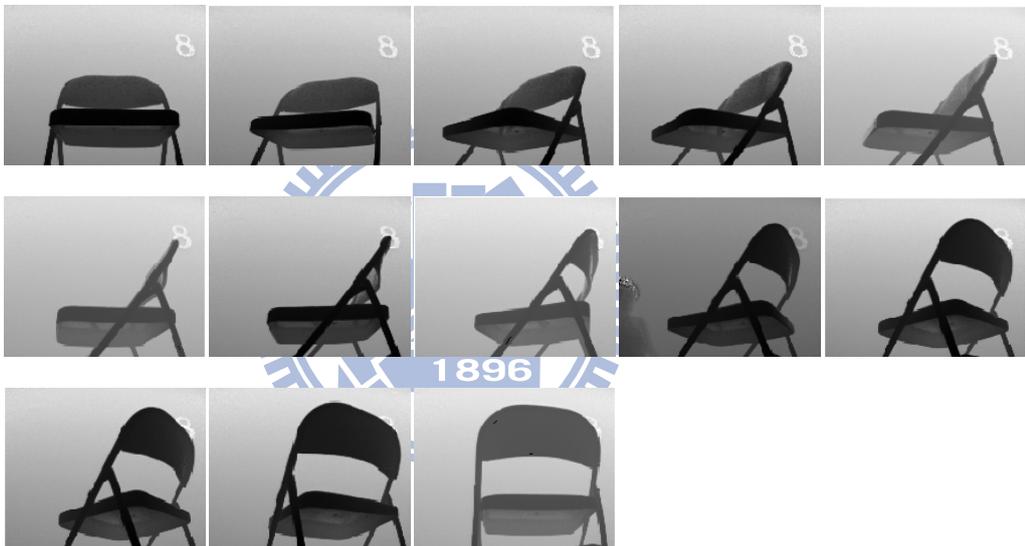
(e)

圖 4-4 玩具人偶的強度影像與 39 張姿態的深度影像

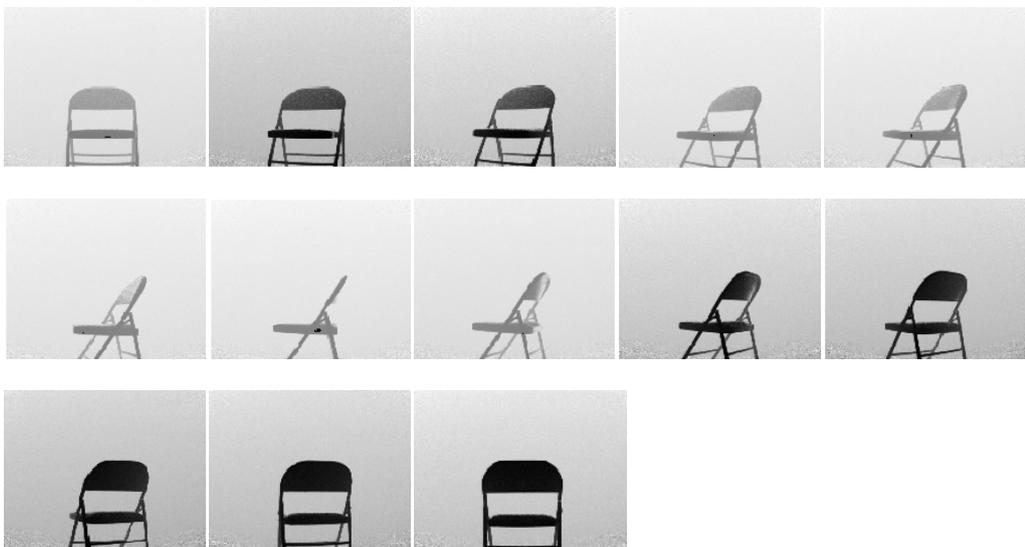
- (a)俯角 45 度，z 軸固定 (b)仰角 45 度，z 軸固定 (c)水平 0 度，z 軸固定  
 (d)水平 0 度，x 軸固定 (e)水平 0 度，y 軸固定



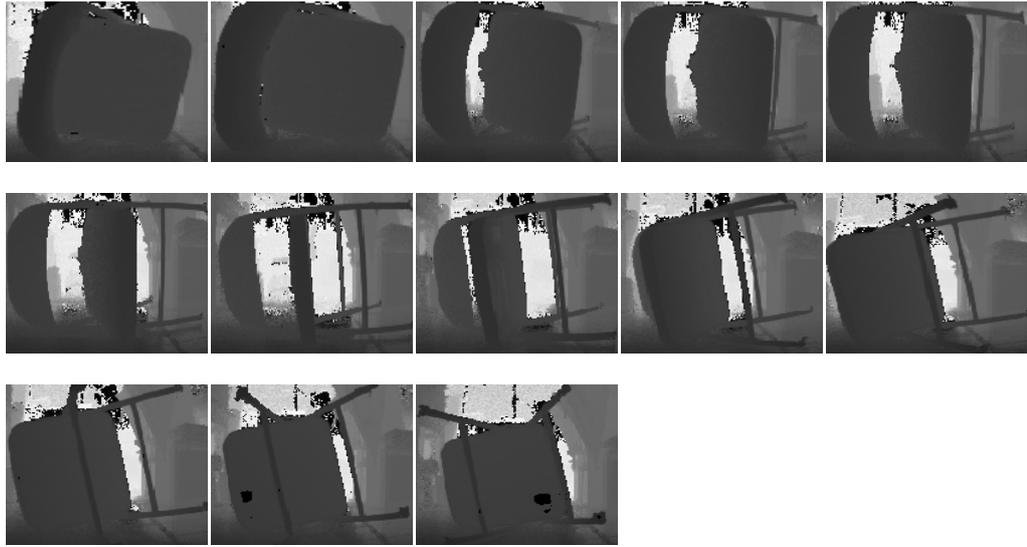
(a)



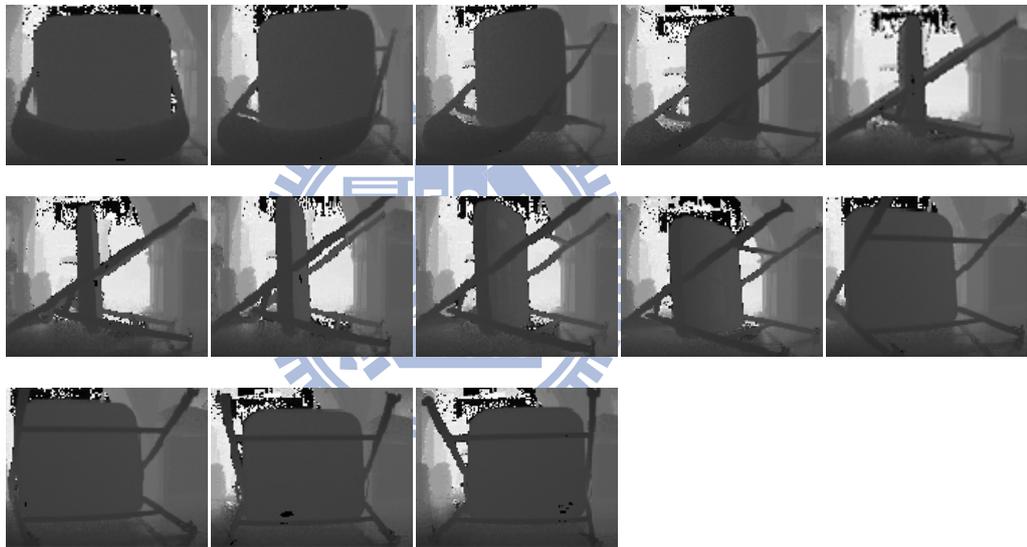
(b)



(c)



(d)

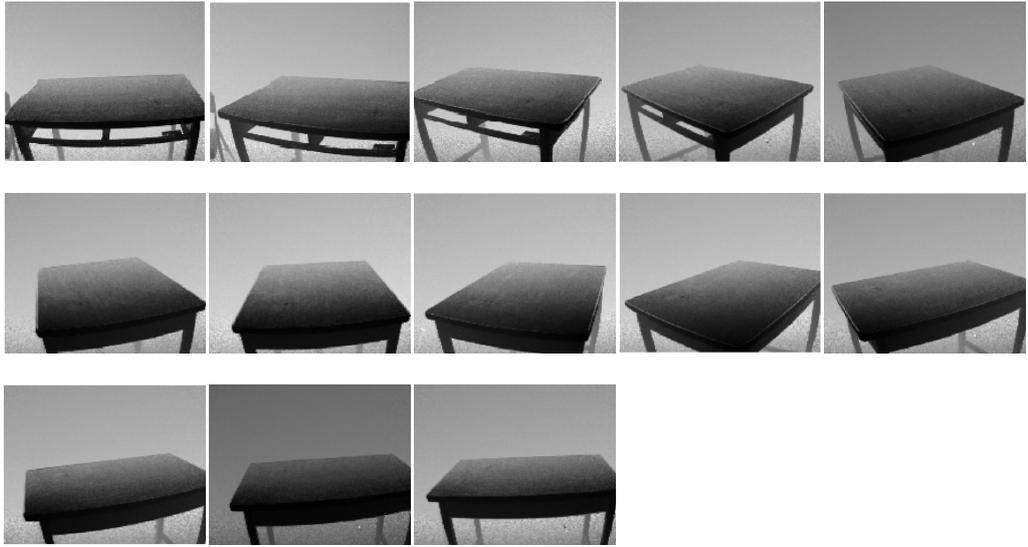


(e)

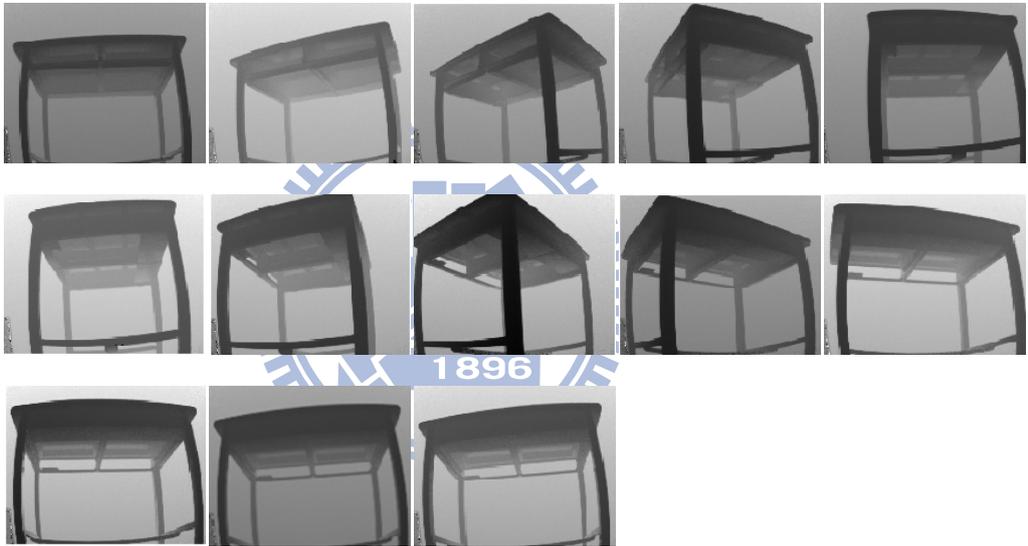
圖 4-5 摺疊椅的強度影像與 39 張姿態的深度影像

(a)俯角 45 度，z 軸固定 (b)仰角 45 度，z 軸固定 (c)水平 0 度，z 軸固定

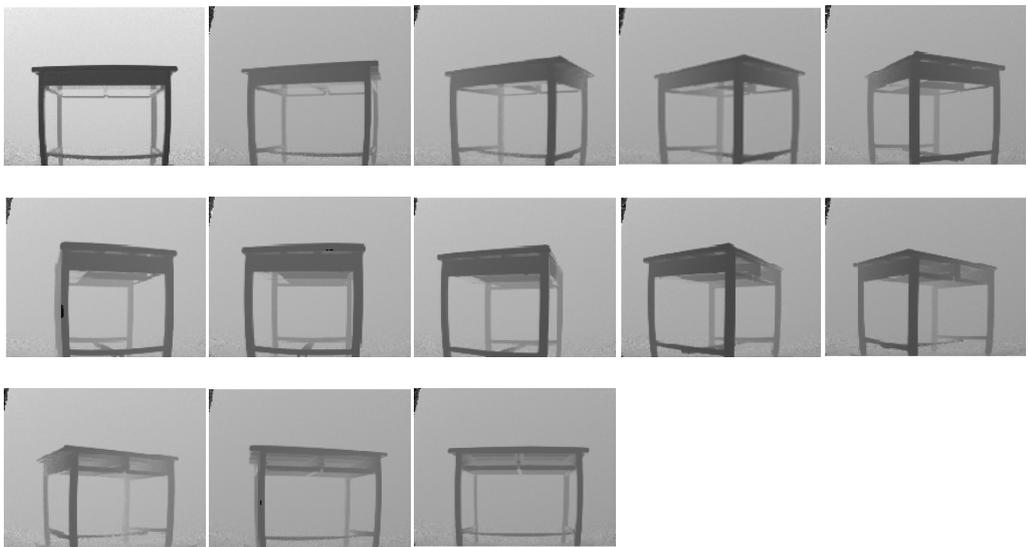
(d)水平 0 度，x 軸固定 (e)水平 0 度，y 軸固定



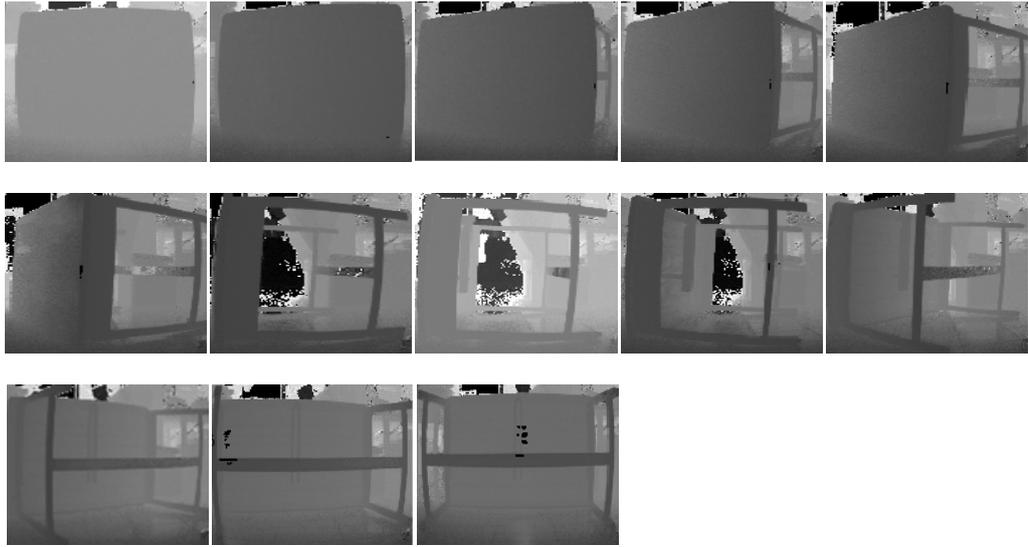
(a)



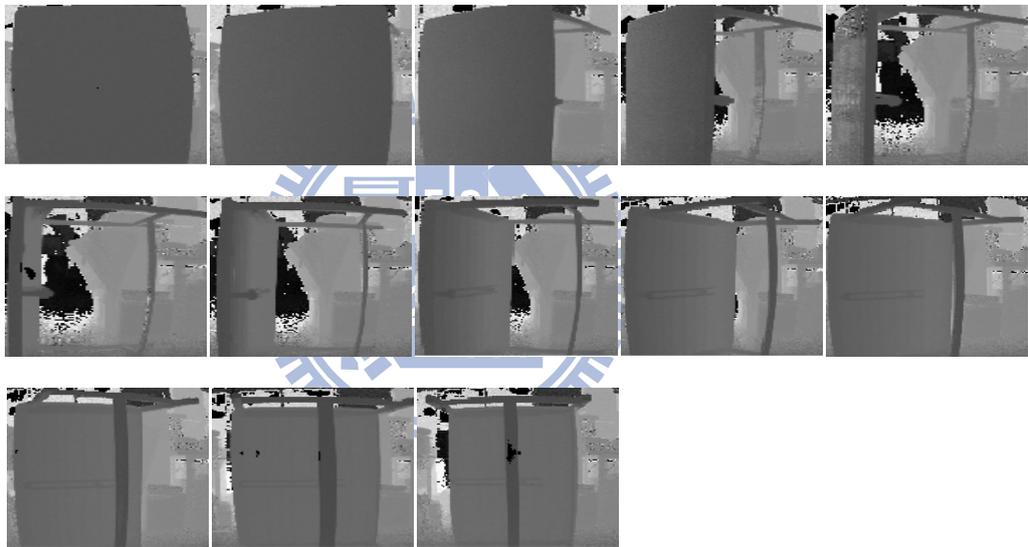
(b)



(c)



(d)

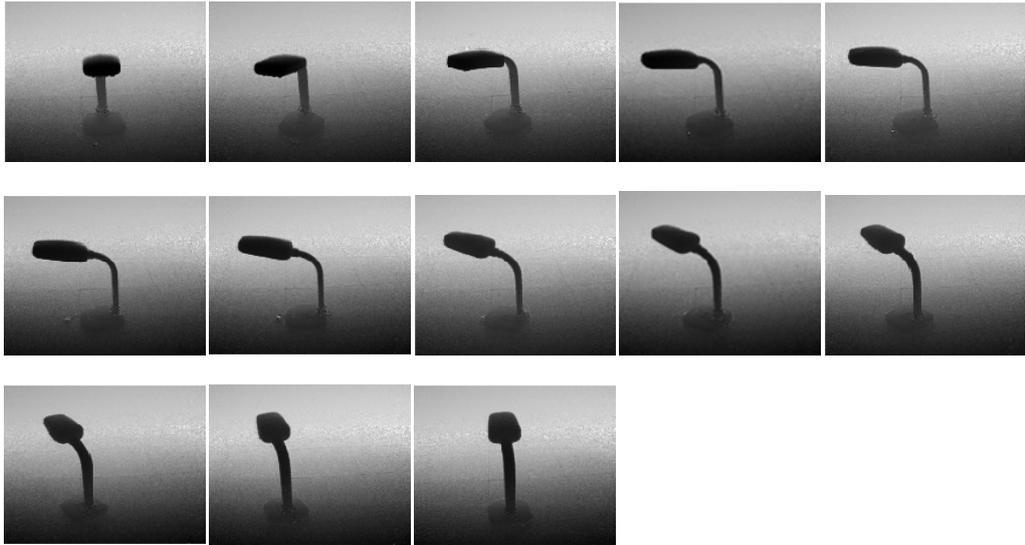


(e)

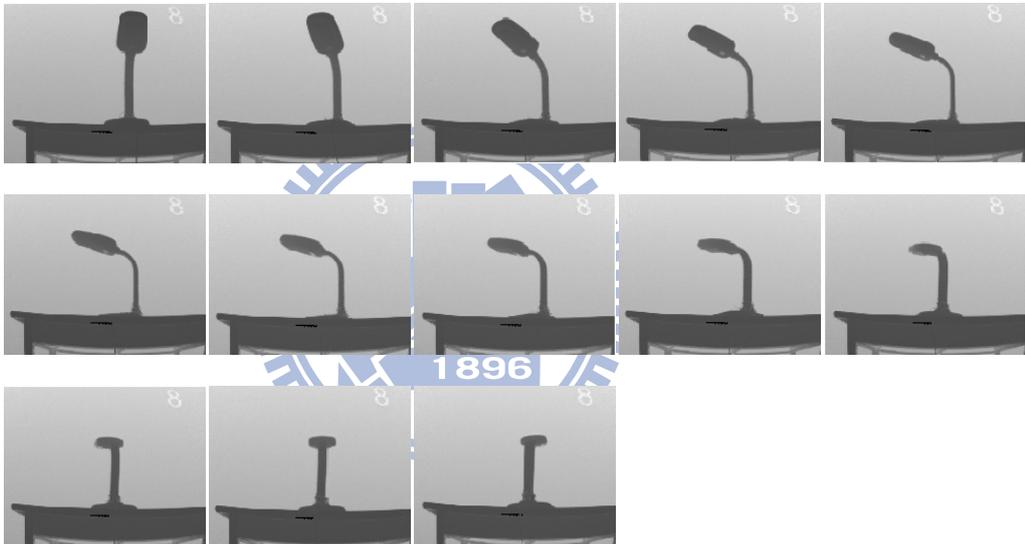
圖 4-6 桌子(無抽屜)的強度影像與 39 張姿態的深度影像

(a)俯角 45 度，z 軸固定 (b)仰角 45 度，z 軸固定 (c)水平 0 度，z 軸固定

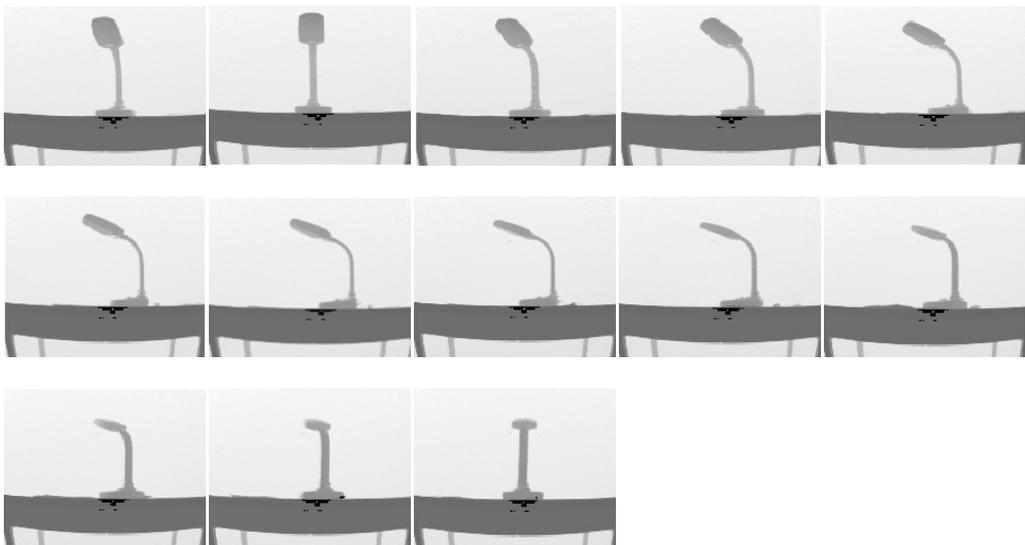
(d)水平 0 度，x 軸固定 (e)水平 0 度，y 軸固定



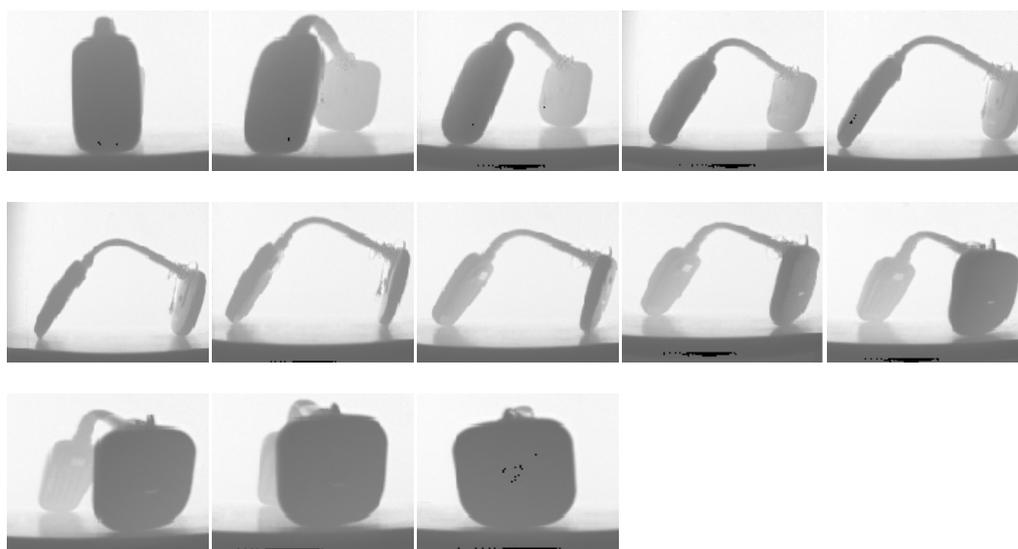
(a)



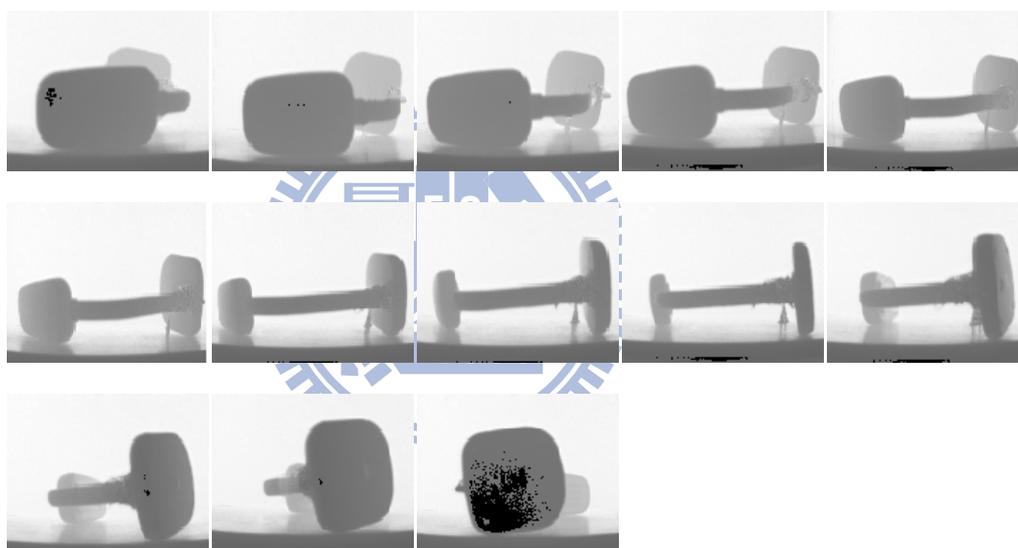
(b)



(c)



(d)

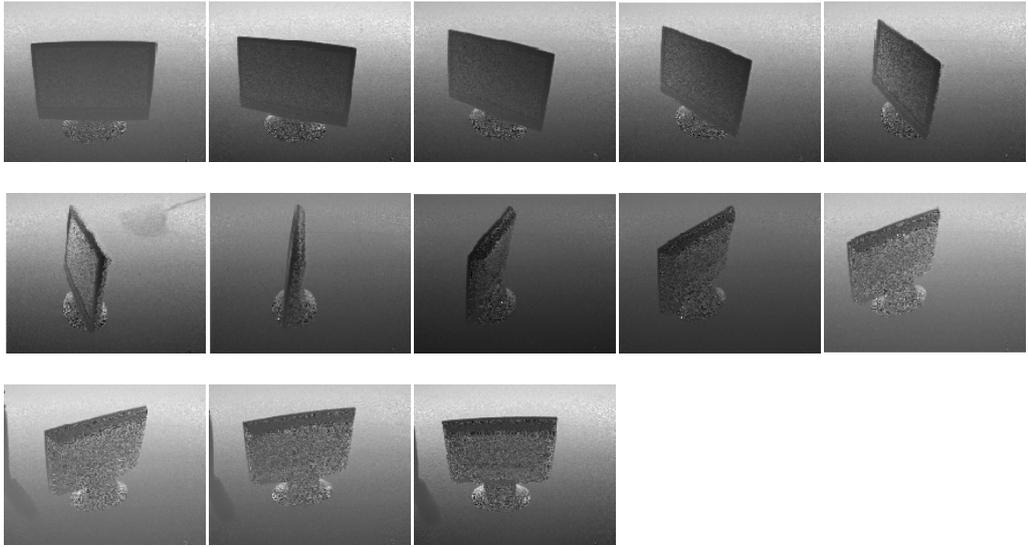


(e)

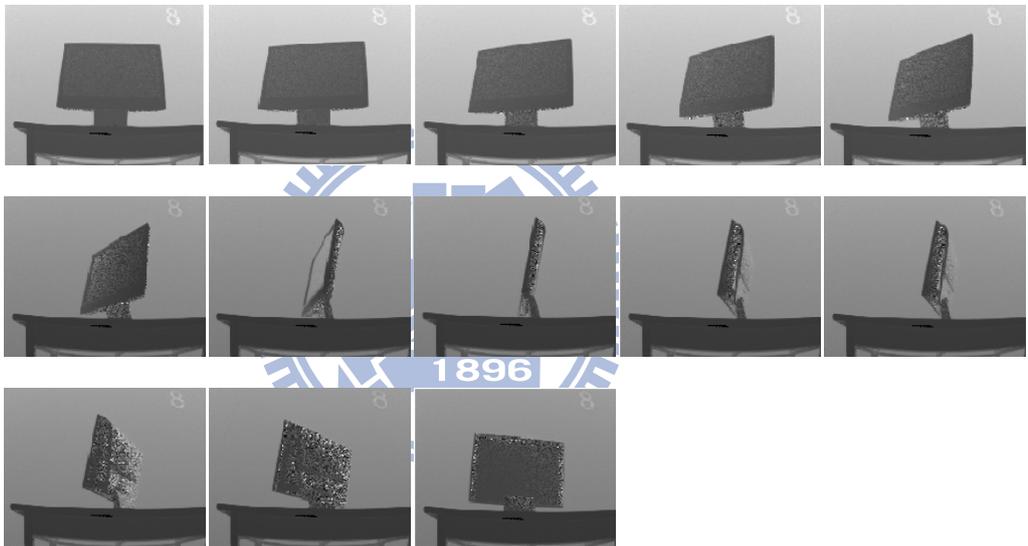
圖 4-7 檯燈的強度影像與 39 張姿態的深度影像

(a)俯角 45 度，z 軸固定 (b)仰角 45 度，z 軸固定 (c)水平 0 度，z 軸固定

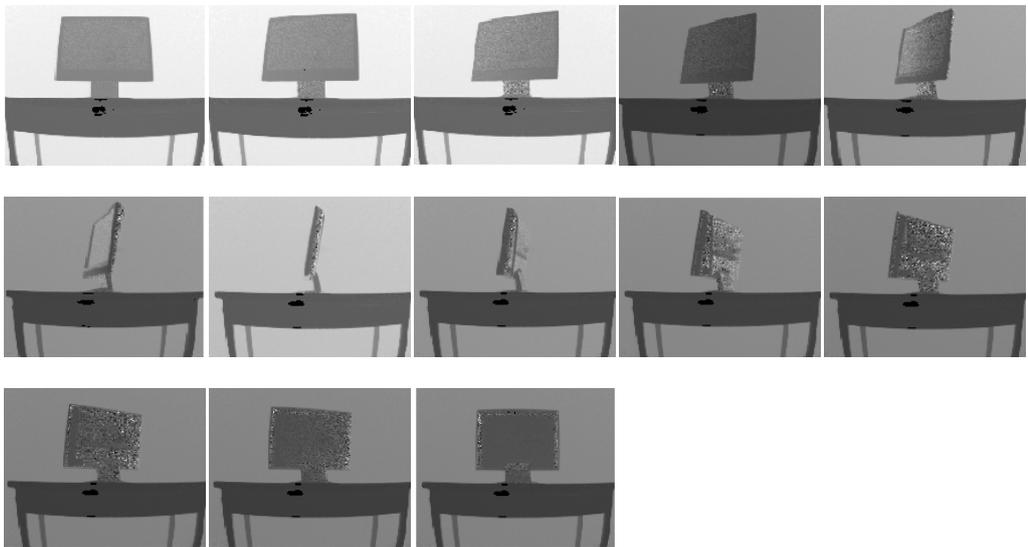
(d)水平 0 度，x 軸固定 (e)水平 0 度，y 軸固定



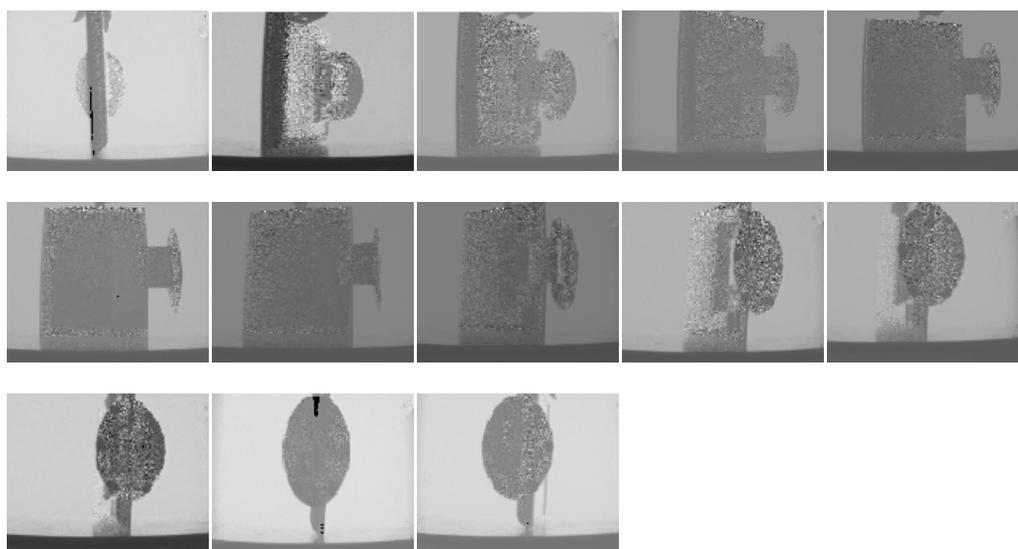
(a)



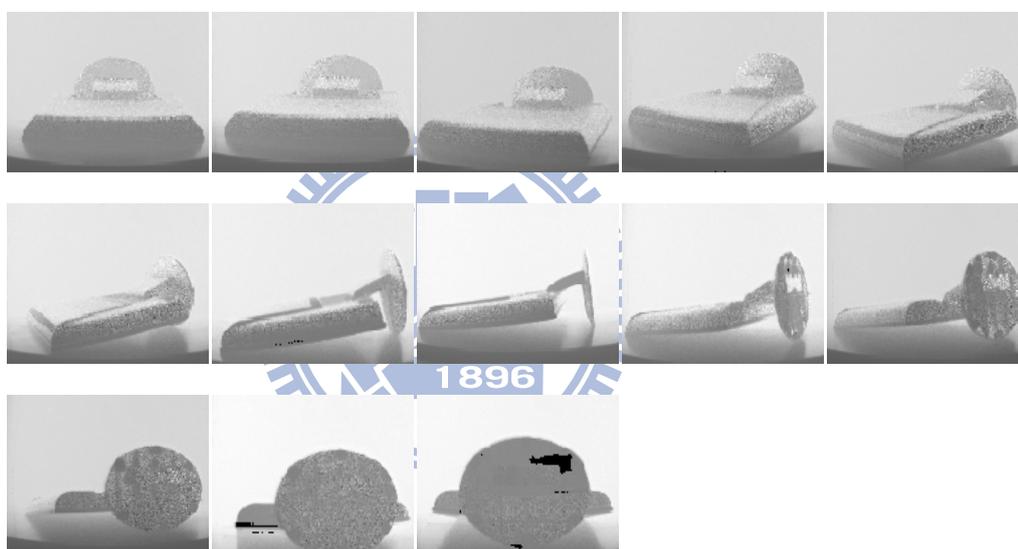
(b)



(c)



(d)

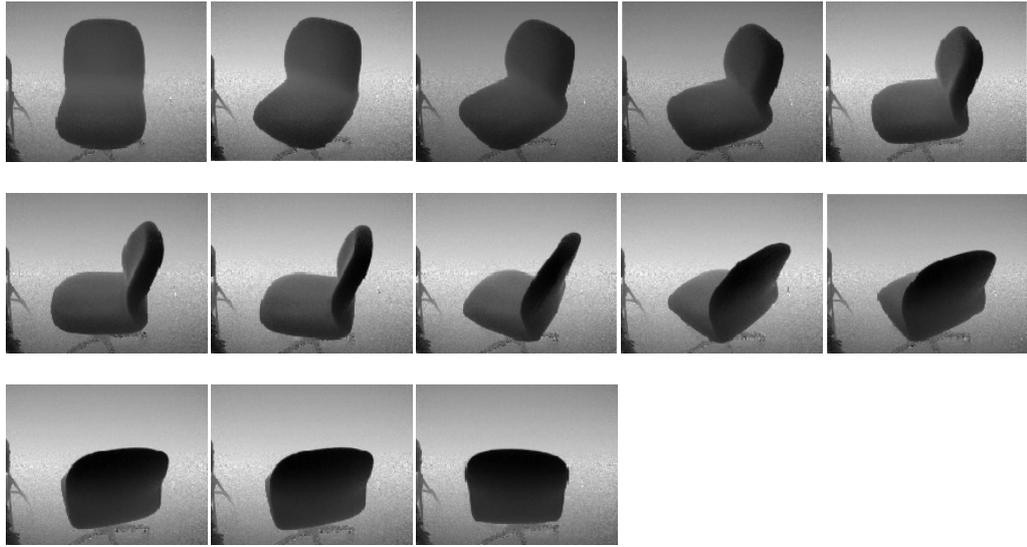


(e)

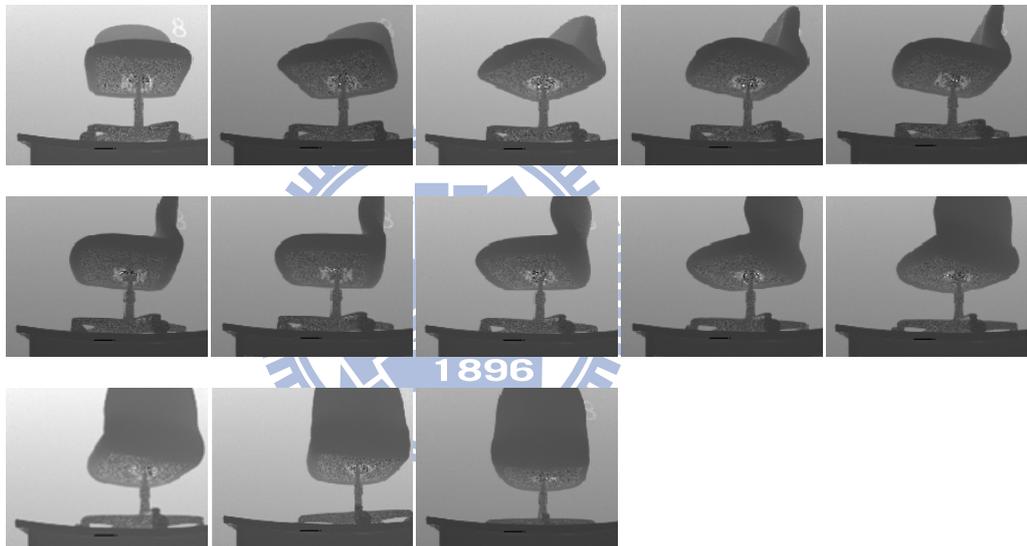
圖 4-8 LCD 電腦螢幕的強度影像與 39 張姿態的深度影像。

(a)俯角 45 度，z 軸固定 (b)仰角 45 度，z 軸固定 (c)水平 0 度，z 軸固定

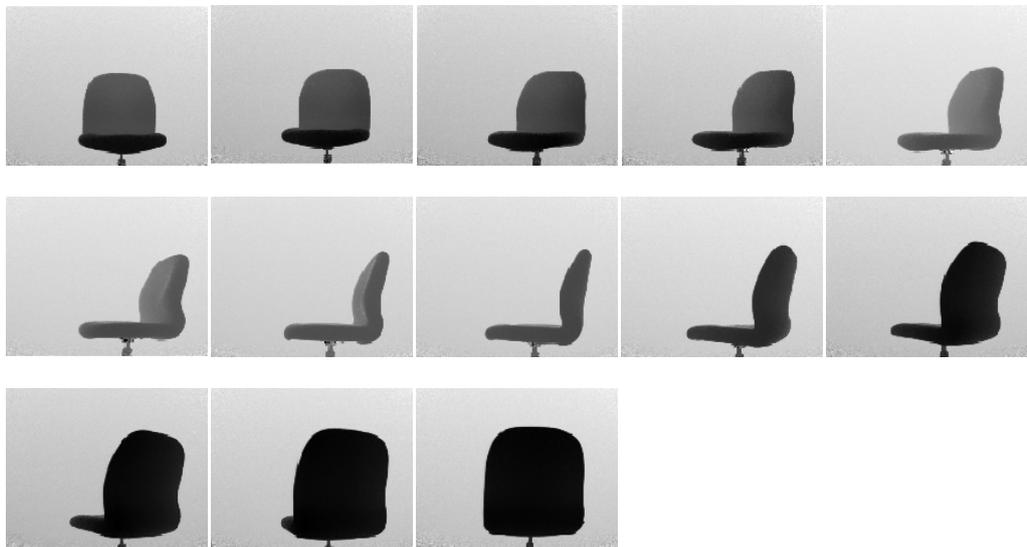
(d)水平 0 度，x 軸固定 (e)水平 0 度，y 軸固定



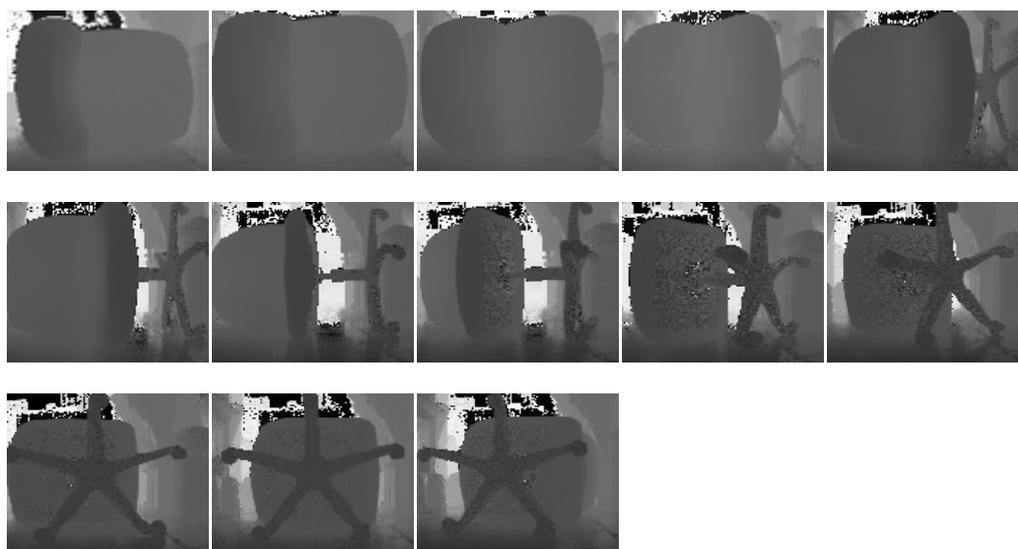
(a)



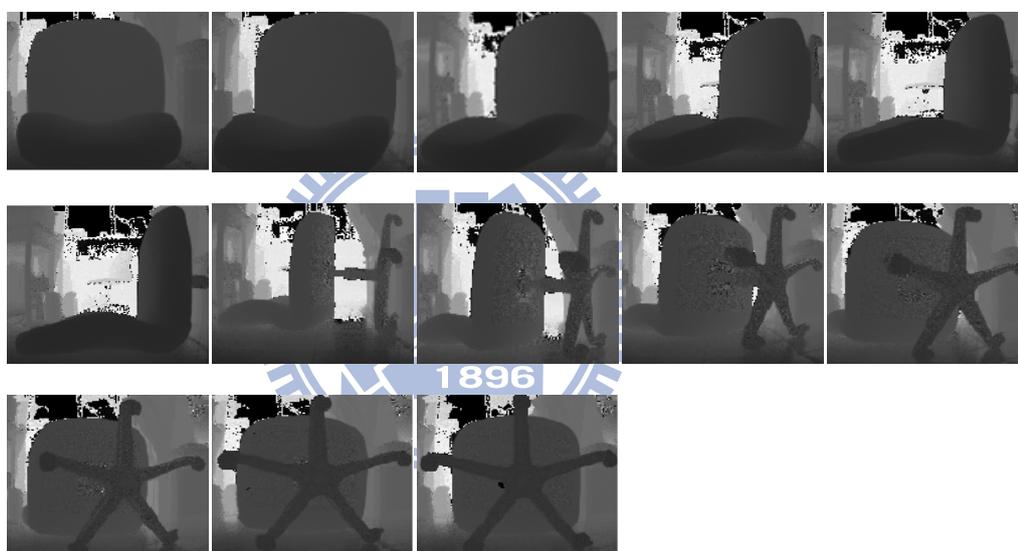
(b)



(c)



(d)

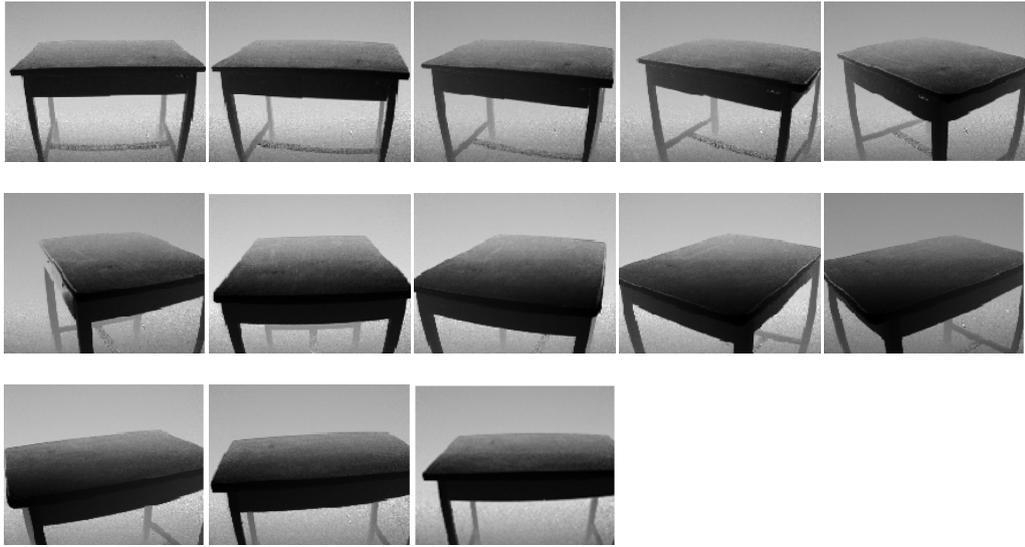


(e)

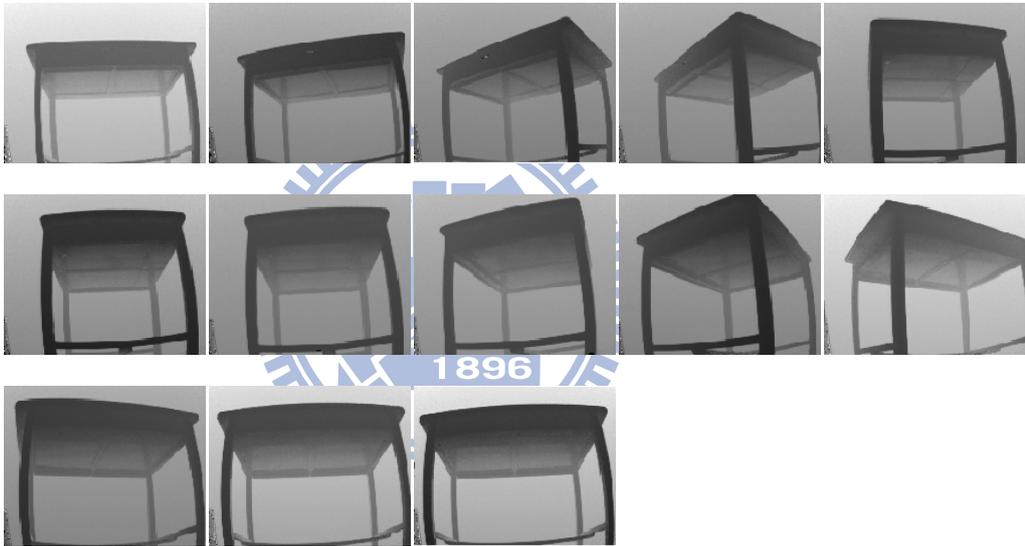
圖 4-9 旋轉椅的強度影像與 39 張姿態的深度影像

(a)俯角 45 度，z 軸固定 (b)仰角 45 度，z 軸固定 (c)水平 0 度，z 軸固定

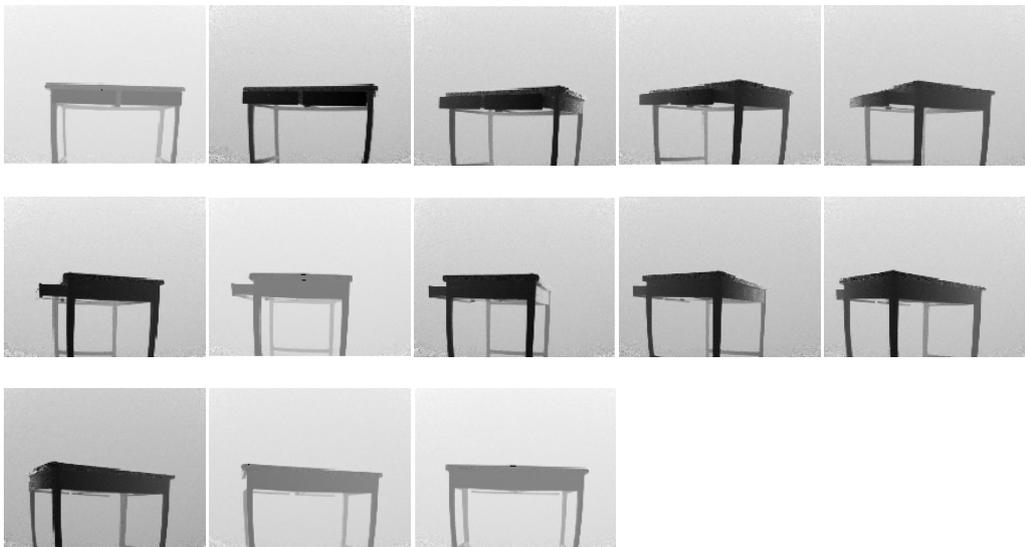
(d)水平 0 度，x 軸固定 (e)水平 0 度，y 軸固定



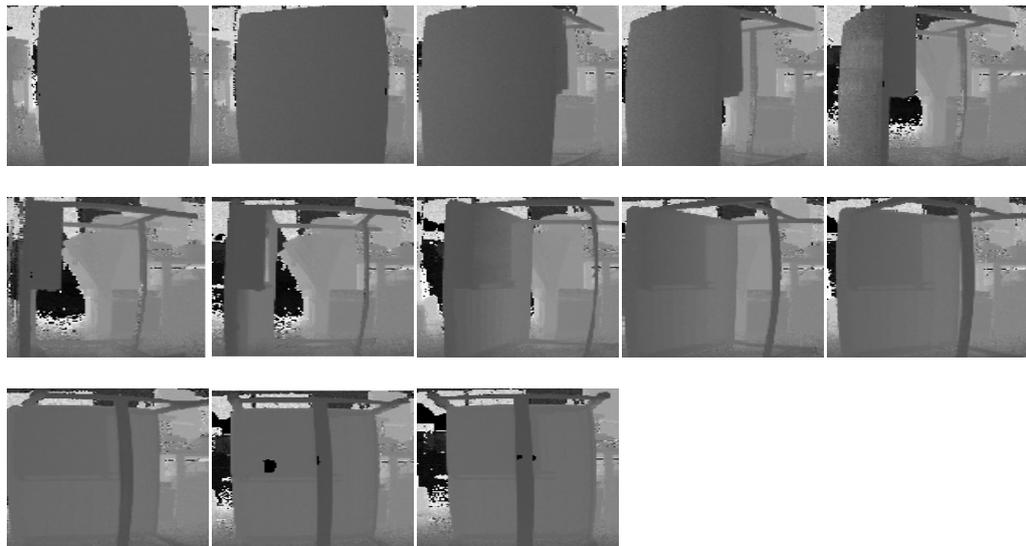
(a)



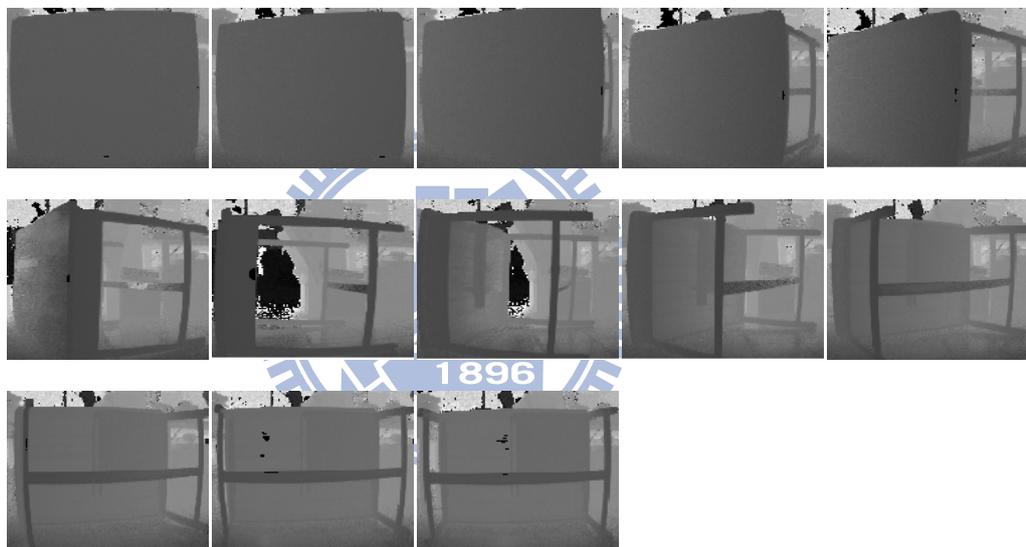
(b)



(c)



(d)



(e)

圖 4-10 桌子(有抽屜)的強度影像與 39 張姿態的深度影像

- (a)俯角 45 度，z 軸固定 (b)仰角 45 度，z 軸固定 (c)水平 0 度，z 軸固定  
 (d)水平 0 度，x 軸固定 (e)水平 0 度，y 軸固定

將對拍攝到每個物體的每個姿態之深度資訊，再以人工的方式進行去除背景之動作，最後就可以獲得每個物體不同姿態的深度資訊，如圖 4-10；此外，將隨機選取資料庫中的物體，擺設於場景中形成複雜場景，以作為測試資料用，並隨機選取拍攝物體的水平角度，作為測試角度用，如表 4-4 所示，而每個測試角度所對應的姿態，將依照表 4-3 的規則來認定該未知物體的姿態。

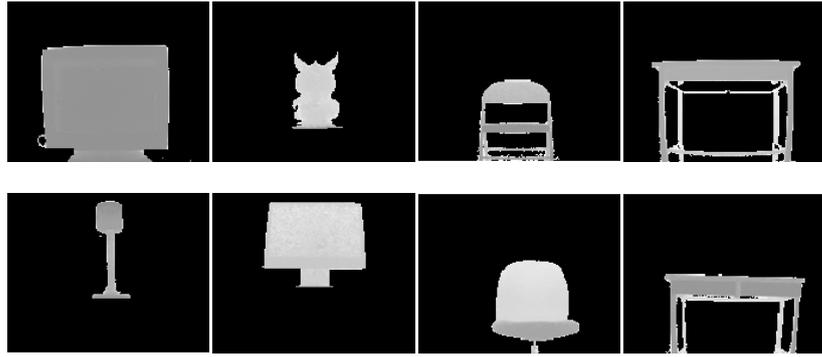


圖 4-11 資料庫中的各別單一物體的深度資訊顯示

表 4-4 未知物體的測試角度與該角度所對應的姿態

測試角度	2	5	13	18	27	33	42	50	57	65	70	77	85
物體姿態	1	1	2	2	3	3	4	4	5	5	6	6	7
測試角度	92	101	110	114	122	130	138	145	153	160	170	175	178
物體姿態	7	8	8	9	9	10	10	11	11	12	12	13	13

測試資料分成兩種狀況，狀況 1 是場景中的物體沒有被遮蔽住的狀況，狀況 2 是場景中的物體有被遮蔽住的狀況，而每一種狀況分別拍攝兩組複雜場景，場景 1 是由物體 1 至物體 4 所組成的場景，其背景為複雜背景，場景 2 是由物體 5 至物體 8 所組成，其背景為複雜背景，而場景中的物體依照表 4-4 上面的測試角度來旋轉，可獲得 26 張場景 1 的深度資訊，圖 4-12 是該場景的強度景像，其場景中深度影相如圖 4-16，同理，由圖 4-13 也可獲得 26 張場景 2 的深度資訊，如圖 4-17 所示，其中未遮蔽的狀況包含兩組場景，總共有 52 張深度資訊作為測試資料的輸入，而遮蔽的狀況也包含兩組場景，場景 1 如圖 4-14 所示，其場景的深度影像如圖 4-18 所示，場景 2 如圖 4-15 所示，其場景的深度影像如圖 4-19 所示，總共有 52 張深度資訊作測試資料的輸入。



圖 4-12 未遮蔽之場景 1 的強度影像



圖 4-13 未遮蔽之場景 2 的強度影像



圖 4-14 遮蔽之場景 1 的強度影像



圖 4-15 遮蔽之場景 2 的強度影像

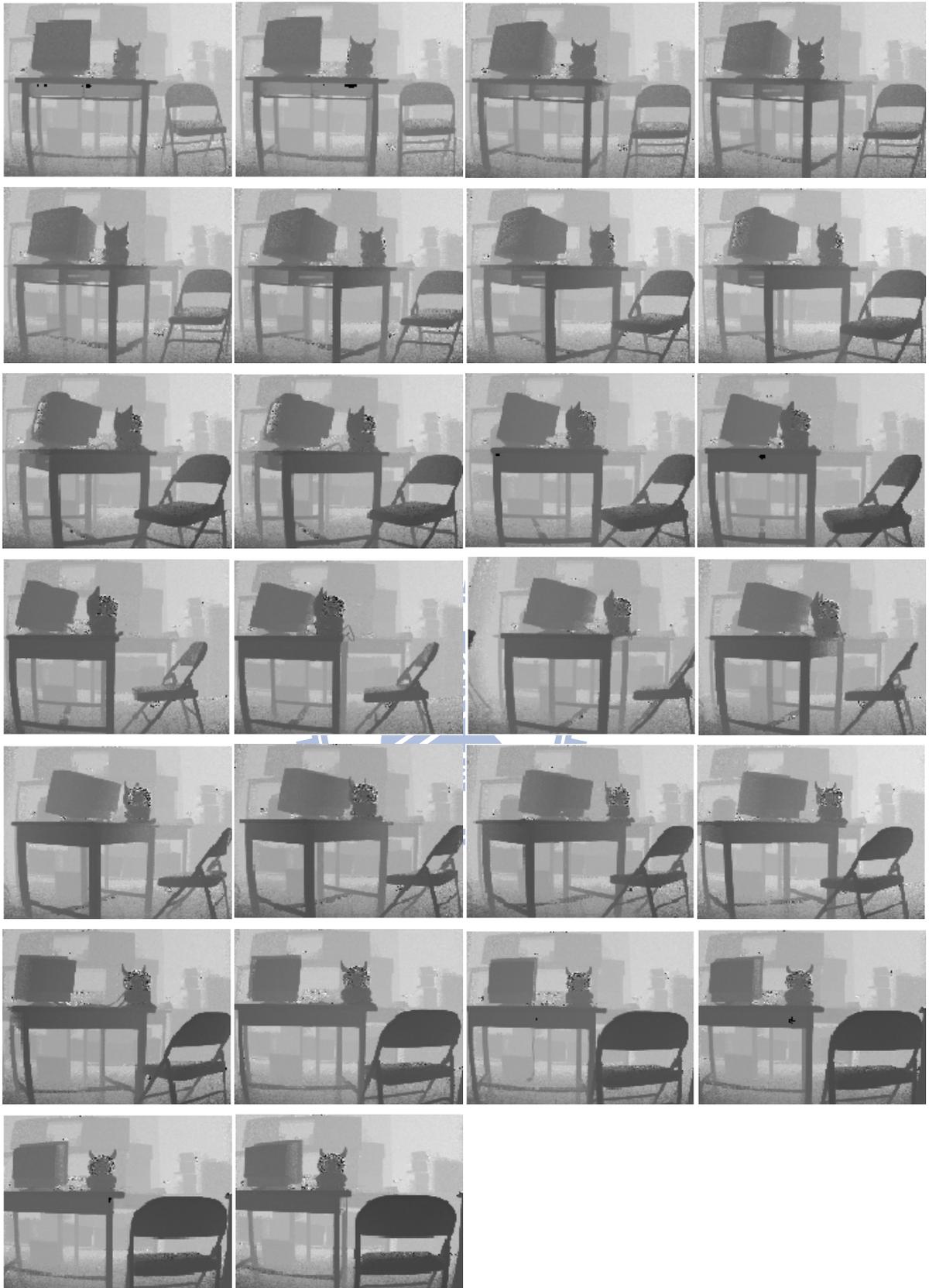


圖 4-16 未遮蔽狀況下之場景 1 的不同姿態之深度影像

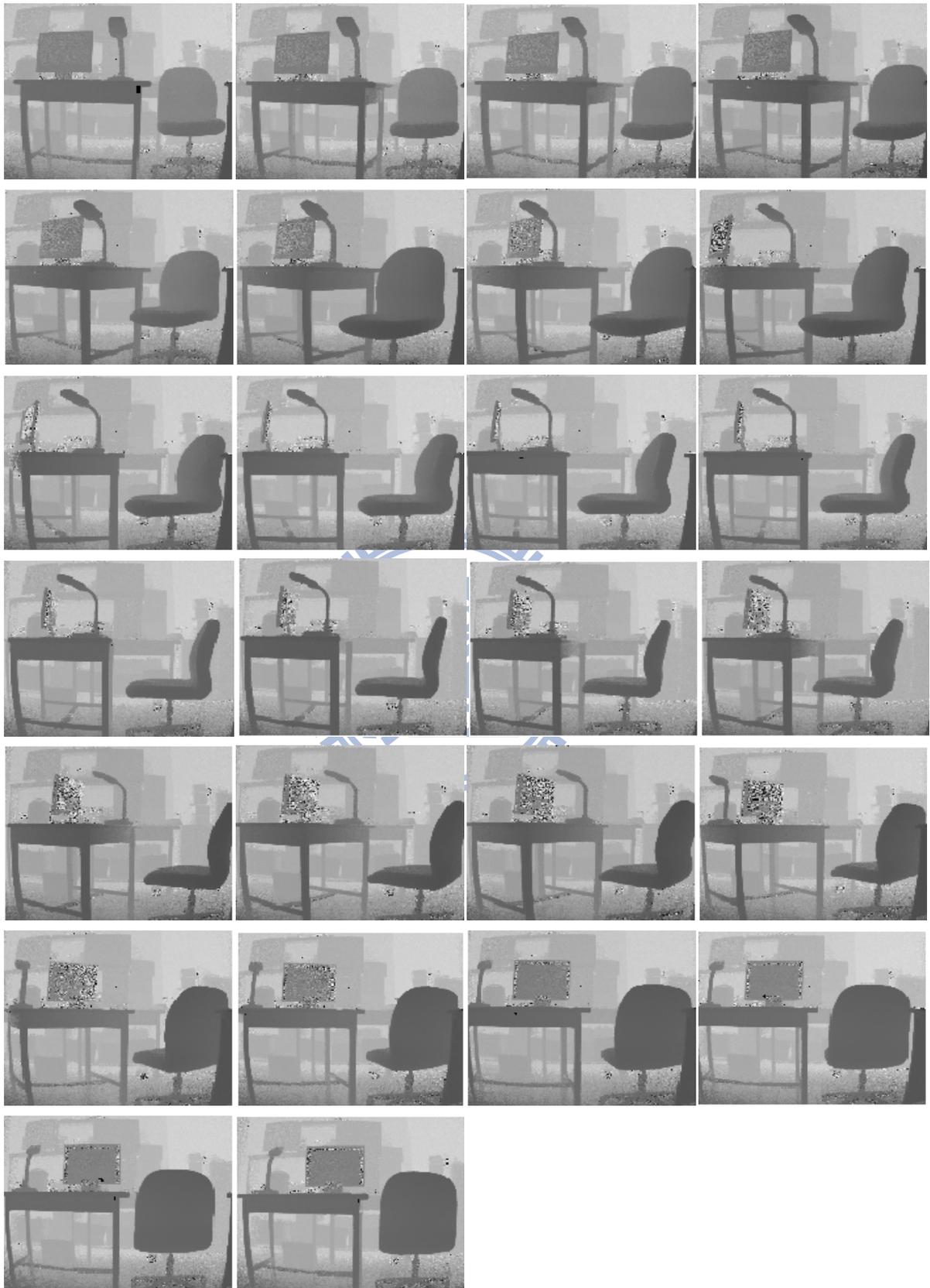


圖 4-17 未遮蔽狀況下之場景 2 的不同姿態之深度影像

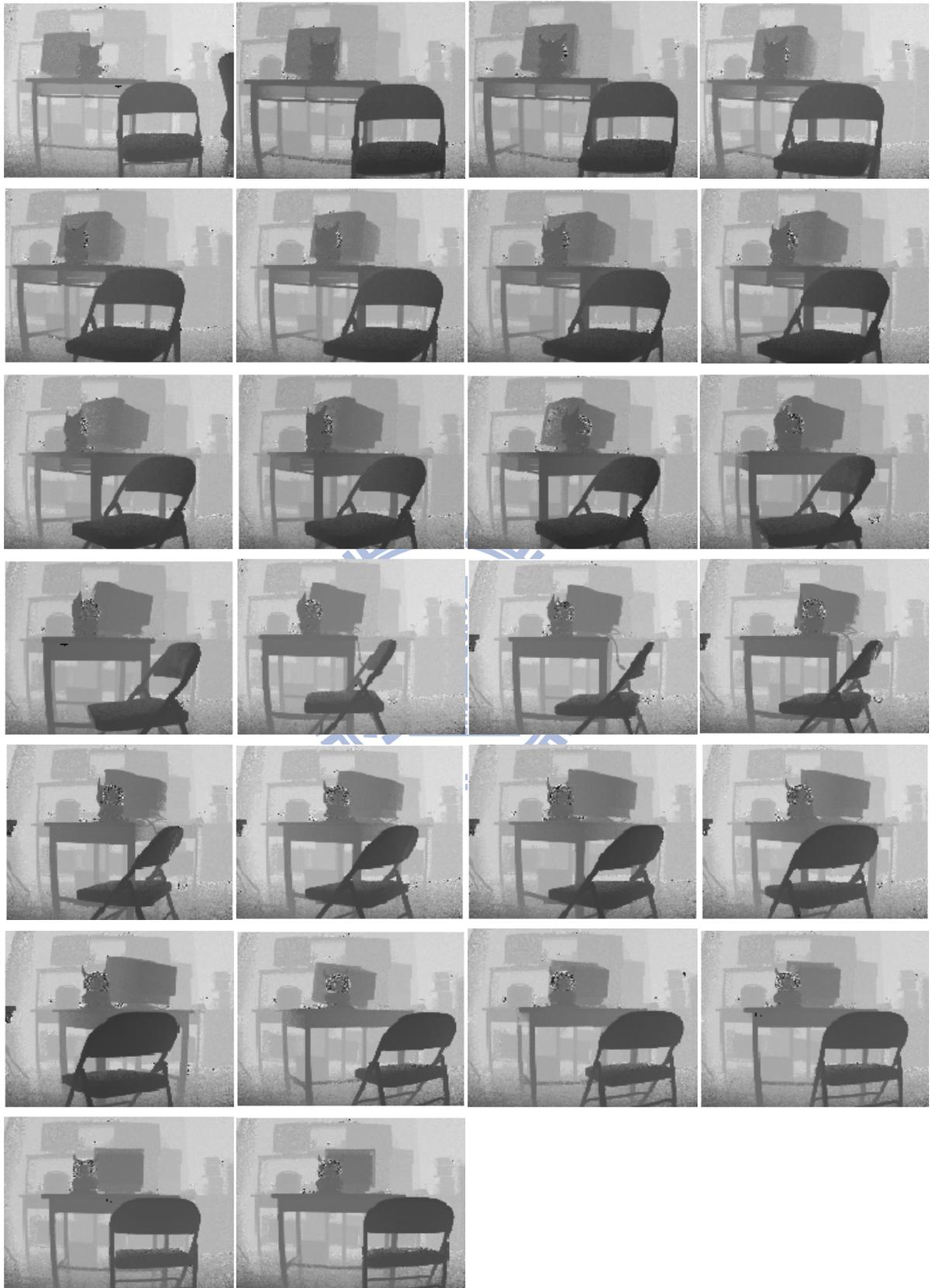


圖 4-18 遮蔽狀況下之場景 1 的不同姿態之深度影像

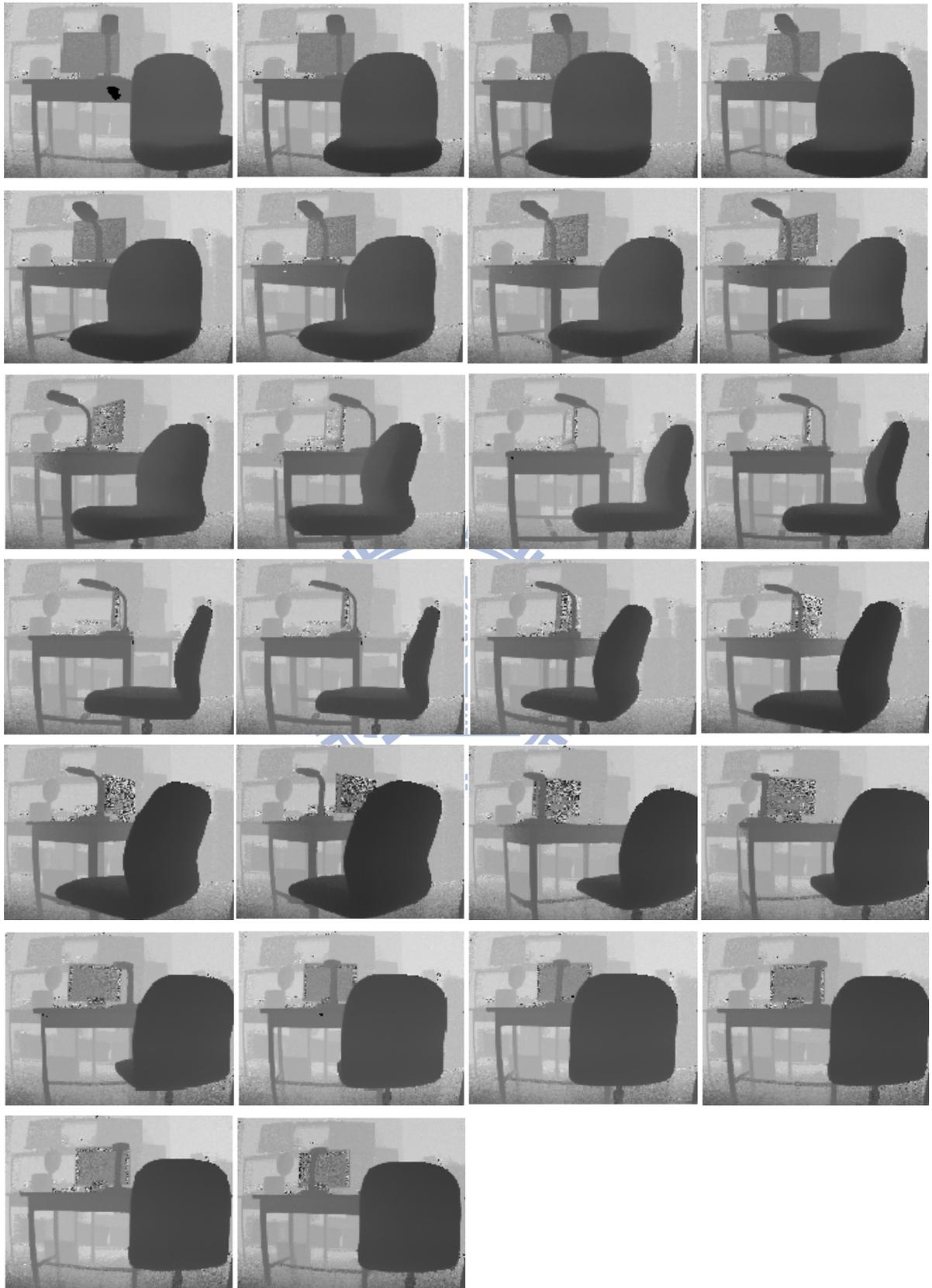


圖 4-19 遮蔽狀況下之場景 2 的不同姿態之深度影像

### 4.1.2 效能評估

在三維物體辨識系統的效能評估上，將比較參考文獻[32]中所列出的辨識方法與本論文所提出的辨識方法作比較，並且比較兩辨識系統對於抵抗雜訊與遮蔽的優劣程度。

針對一個複雜場景，場景中有包括多個物體，

本論文將統計三維物體辨識系統之物體辨識率  $Reg$  隨著雜訊程度以及遮蔽程度的改變，物體辨識率  $Reg$  的計算為(4.1)式。

$$Reg = \frac{\text{Num}_{\text{corret}}}{\text{Num}_{\text{object}}} \quad (4.1)$$

對一個複雜場景的深度資訊而言， $\text{Num}_{\text{corret}}$  為正確辨認場景中的物體個數為， $\text{Num}_{\text{object}}$  為場景中的物體個數。首先，計算完該場景中的第  $j$  個姿態的未知物體  $O_{\text{query}}$  與資料庫中第  $i$  姿態的已知物體  $O_k$  之相異程度  $\chi(O_{\text{query}}^{Pj}, O_k^{Pi})$  的大小後，將對  $\chi(O_{\text{query}}^{Pj}, O_k^{Pi})$  進行排序；若第一名的物體與該未知物體是同一種物體時，即使兩者的姿態不同，但  $O_{\text{query}} = O_k$ ， $\text{Num}_{\text{corret}}$  的個數仍加 1，最後辨識完該場景中所有的未知物體，即獲得該場景的物體辨識率，總共辨識 52 張複雜場景的深度資訊，最後再取平均，即可獲得代表系統的物體辨識率。

此外，為了要驗證本系統對於尋找出正確物體的姿態之強健性，也統計三維物體辨識系統之姿態辨識率，正確辨識出未知物體的正確姿態的物體姿態辨識率  $Reg^{\text{pose}}$ ，如(4.2)式。

$$Reg^{\text{pose}} = \frac{\text{Num}_{\text{corret}}^{\text{pose}}}{\text{Num}_{\text{object}}} \quad (4.2)$$

$Reg^{\text{pose}}$  跟(4.1)式的計算方法一樣， $\text{Num}_{\text{corret}}^{\text{pose}}$  為某一複雜場景下正確辨識出物體的姿態的物體個數， $\text{Num}_{\text{object}}$  為該場景中的物體個數，前提是，必須要辨識結果的物體與未知物體為相同物體外，姿態也必須相同時， $\text{Num}_{\text{corret}}^{\text{pose}}$  的個數加 1，最後辨識完該場景中所有的未知物體之姿態，即獲得該場景的姿態辨識率，總共辨識 52 張複雜場景的深度資訊，最後再取平均，即可獲得代表系統的姿態辨識率。

最後，為了要證實本系統對於物體姿態些微變化下，仍可辨識出同一種物體，計算此最後辨識結果的前三名物體之辨識率  $Reg^{1-3}$ ，其中  $Reg^{1-3}$  是計算某一複雜場景下的某一未知物體  $O_{query}$  被正確辨識出正確種類的比例，也就是前三名的物體均是同一種物體的比例值，如(4.3)式。

$$Reg^{1-3} = \frac{\sum_{k=1}^{Num_{correct}} P(O_{query})}{Num_{object}} \quad (4.3)$$

若最後真對某一複雜場景下的某一未知物體，其辨識結果的前三名中只有一名物體種類為正確， $P(O_{query}) = \frac{1}{3}$ ，若最後辨識結果的前三名中只有兩名物體被正確辨識出來，只是姿態略為不同， $P(O_{query}) = \frac{2}{3}$ ；若前三名中的所有物體都是同一種物體，則  $P(O_{query}) = 1$ ，最後累加該場景中所有物體的  $P(O_{query})$ ，並且除以該場景中的物體總數  $Num_{object}$ ，而本論文將計算 52 張深度資訊的  $Reg^{1-3}$ ，最後再取平均，為代表此系統的辨識率。



## 4.2 實驗結果與分析

實驗將會分成四組來比較，依照有無使用本論文所 3.2 節提出的雜訊處理系統，以及拍攝場景是否為遮蔽環境下，這兩個變因來依序組合成如下表 4-5 所示。

表 4-5 實驗組別與實驗變因之關係

	無雜訊處理	有雜訊處理
無遮蔽	組別 1	組別 2
有遮蔽	組別 3	組別 4

例如：組別 1 的測試資料為無遮蔽的測試資料，即為圖 4-10、圖 4-11 所有的深度資訊且這些深度資訊並沒有經過 3.2 節的雜訊處理，直接進行場景分離、物體分割、抽取特徵等步驟，最後再與資料庫中的物體進行比對，以獲得最後辨識結果，而組別 2 與組別 1 的差異就是有經過 3.2 節的雜訊處理步驟，同理，組別 3 與組別 4 的差異在於有

無雜訊處理，組別 2 與組別 4 的差異在於輸入的測試資料中，物體是否有遮蔽的情形。

此外，本論文將會跟七種不同的方法，如表 4-6 所示，並且比較其辨識率之優劣。

表 4-6 各方法與各種樣本區間個數列表

方法	d	n	s	s+d	n+s	n+d	nsd	nsd_region
區間個數	32	8×8	64	16×16	8×8×16	4×4×8	4×4×4×8	4×4×4×8

d 表示的方法是只使用 2.5.1 節物體表面每點的深度變化作為區域特徵，利用 2.6 節所提的直方圖來統計此區域特徵，以進行三維物體比對；n 表示的方法是只使用 2.5.2 節物體表面的每點法向量之夾角作為區域特徵，利用 2.6 節所提的直方圖來統計此區域特徵，以進行三維物體比對；而 s 表示的方法是只使用 2.5.3 節物體表面的每點的曲率指標作為區域特徵，利用 2.6 節所提的直方圖來統計此區域特徵，以進行三維物體比對；而 s+d 是結合了 2.5.1 節與 2.5.3 節這兩個區域特徵作為比對特徵之方法，n+s 也是結合了三個區域特徵作為比對特徵之方法，n+d 也是結合了三個區域特徵作為比對特徵之方法，而 nsd 就是選取 2.5.1 節、2.5.2 節以及 2.5.3 節這四個區域特徵作為整體比對特徵之方法；然而本論文在此稱之 nsd\_region 是使用了 nsd 這四個特徵作為整體之比對特徵外，另外使用了區塊特徵的概念來比對遮蔽物體，結合整體特徵比對與區塊特徵比對形成兩階段物體辨識，其中每個區域特徵將統計成直方圖，而直方圖中區間個數(bin size)的選取，是根據 G. Htzl, B. Leibe, P. Levi, and B. Schiele [32]，所提出的最佳區間個數值設定的。

將每組測試資料以及每個方法所獲得的物體辨識率、姿態辨識率以及前三項物體的辨識率，依序列在表 4-7 至表 4-10。

表 4-7 輸入的測試資料為組別 1 之各種方法的辨識率比較

組別 1	d	n	s	s+d	n+s	n+d	Nsd	nsd_region
<i>Reg</i>	30.81%	35.80%	43.41%	35.80%	40.81%	33.17%	54.10%	47.14%
<i>Reg</i> <sup>pose</sup>	21.55%	24.00%	30.05%	24.00%	24.03%	32.05%	34.04%	48.08%
<i>Reg</i> <sup>1-3</sup>	25.88%	34.88%	35.95%	40.88%	40.25%	43.34%	70.18%	55.40%

表 4-8 輸入的測試資料為組別 2 之各種方法的辨識率比較

組別 2	d	n	s	s+d	n+s	n+d	Nsd	nsd_region
<i>Reg</i>	35.25%	39.26%	41.76%	39.26%	39.26%	39.26%	60.32%	60.32%
<i>Reg</i> <sup>pose</sup>	25.03%	35.03%	30.03%	36.56%	37.56%	40.03%	48.62%	51.12%
<i>Reg</i> <sup>1-3</sup>	35.87%	40.19%	41.98%	31.98%	43.98%	45.98%	75.47%	77.74%

表 4-9 輸入的測試資料為組別 3 之各種方法的辨識率比較。

組別 3	d	n	s	s+d	n+s	n+d	nsd	nsd_region
<i>Reg</i>	21.63%	26.04%	25.04%	20.44%	28.04%	28.04%	32.80%	30.06%
<i>Reg</i> <sup>pose</sup>	12.01%	16.02%	16.02%	13.50%	23.05%	22.02%	27.02%	28.07%
<i>Reg</i> <sup>1-3</sup>	28.96%	30.62%	31.81%	30.81%	35.18%	31.10%	51.81%	45.53%

表 4-10 輸入的測試資料為組別 4 之各種方法的辨識率比較。

組別 4	d	n	s	s+d	n+s	n+d	nsd	nsd_region
<i>Reg</i>	26.42%	32.85%	30.65%	28.59%	32.85%	34.45%	46.12%	53.75%
<i>Reg</i> <sup>pose</sup>	17.62%	20.25%	23.55%	23.25%	25.25%	25.25%	29.62%	28.47%
<i>Reg</i> <sup>1-3</sup>	32.26%	31.78%	31.78%	31.78%	51.78%	51.96%	55.17%	57.28%

首先，要比較的是組別 1 與組別 2，測試資料是物體均沒有被遮蔽的狀況發生且均處於複雜場景中，然後比較沒使用本論文的雜訊處理之深度資訊與有使用雜訊處理系統之深度資訊，使用不同特徵抽取之方法的辨識率比較。

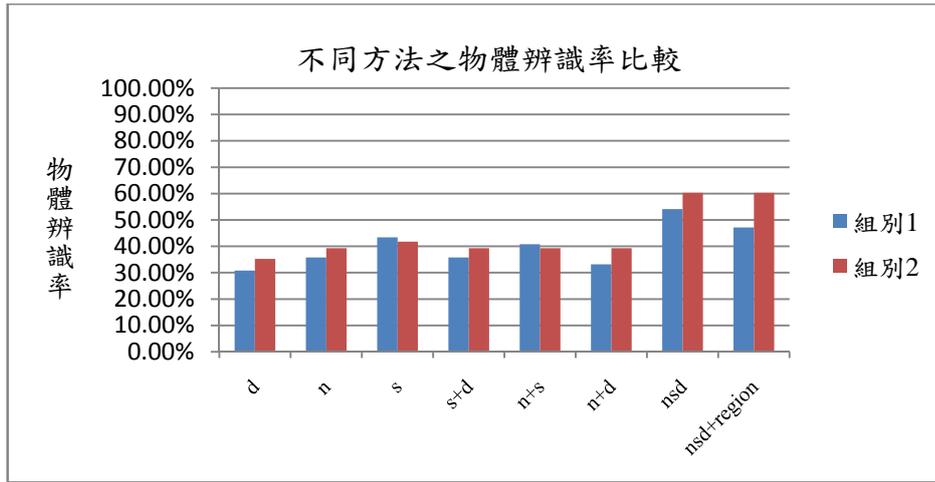


圖 4-20 組別 1 與組別 2 的各種方法之物體辨識率比較

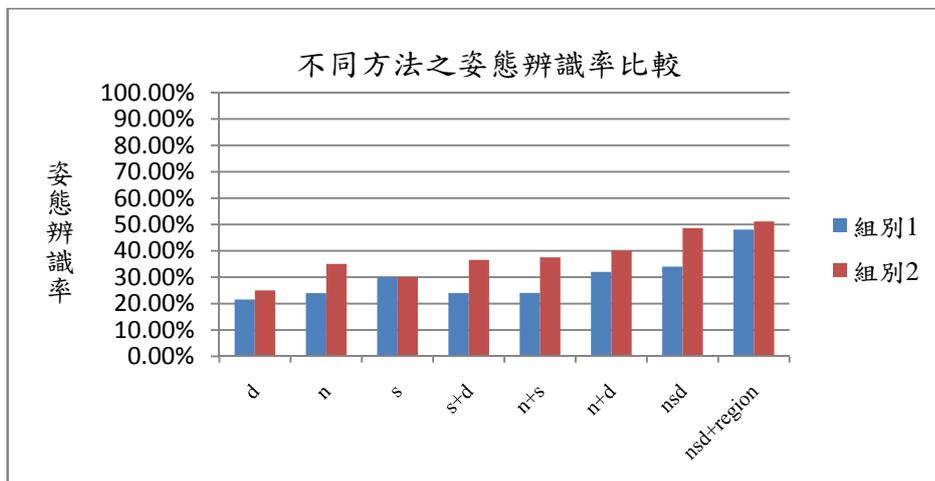


圖 4-21 組別 1 與組別 2 的各種方法之姿態辨識率比較

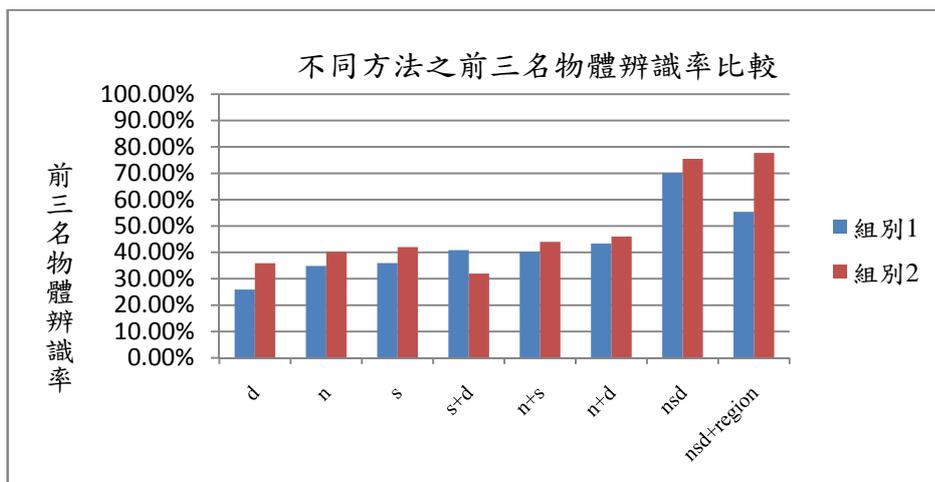


圖 4-22 組別 1 與組別 2 的各種方法之前三名物體辨識率比較

可以看得出來，在有雜訊影響的前提下，本論文所提出的方法 nsd+region，辨識率會略輸給 nsd，主要是因為利用物體的整體特徵作辨識與利用物體的區塊特徵作辨識，整體辨識相對來說比較不會受到雜訊的影響，但是當輸入資料為組別 2，即輸入資料經過雜訊處理後，物體辨識率相對於組別 1 的物體辨識率高，並且跟其他方法相比，本論文在組別 2 的辨識率較高，故本論文所提出的辨識架構較佳。接下來是要比較的是組別 3 與組別 4，測試資料均是物體有遮蔽狀況發生且處於複雜場景之中，然後比較沒使用本論文的雜訊處理之深度資訊與有使用雜訊處理系統之深度資訊，使用不同特徵方法之辨識率比較。

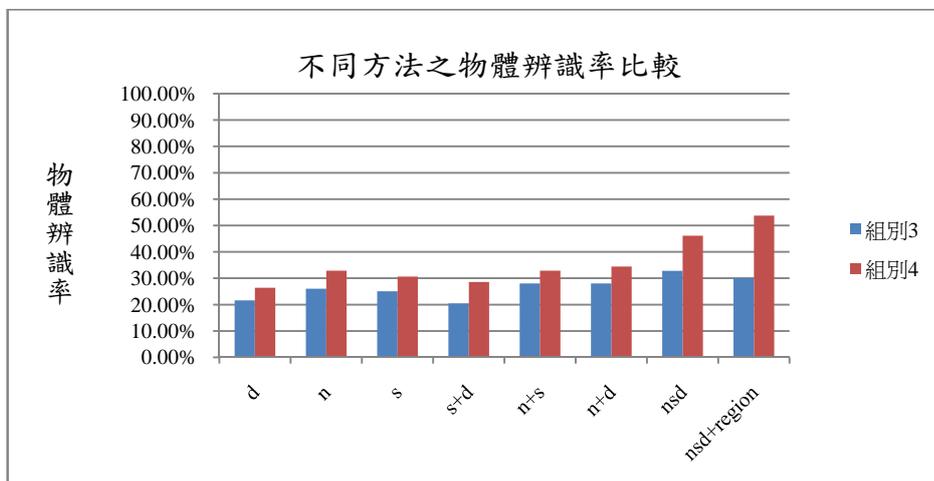


圖 4-23 組別 3 與組別 4 的各種方法之物體辨識率比較

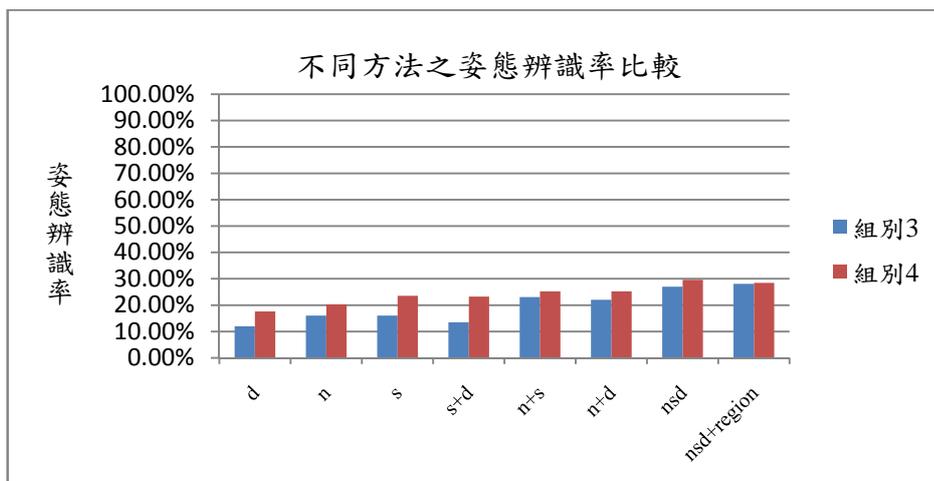


圖 4-24 組別 3 與組別 4 的各種方法之姿態辨識率比較

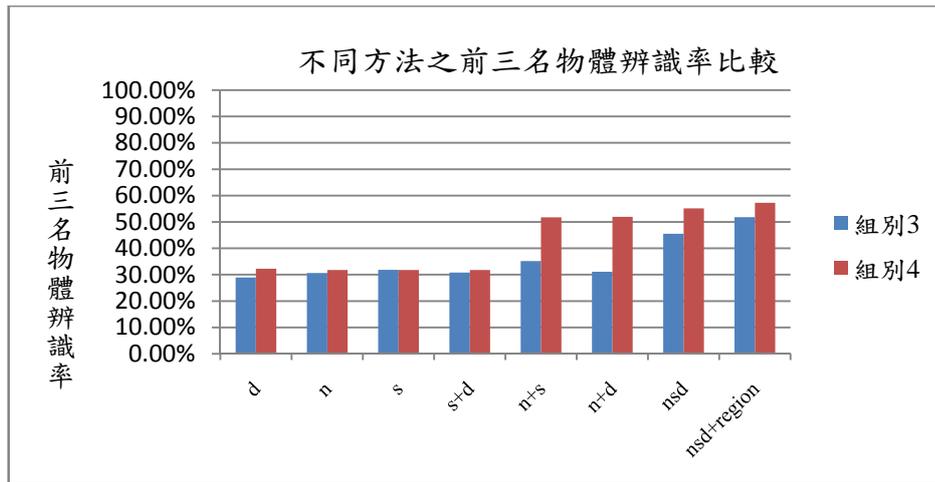


圖 4-25 組別 3 與組別 4 的各種方法之前三名物體辨識率比較

比較組別 3 與組別 4，可以看得出來在物體處於被遮蔽的狀況時，有經過雜訊處理後，明顯的可以增加物體辨識率，但是在姿態辨識率上，依舊沒有太大的改進。

最後可以看得出來在資料都經過雜訊處理後，本論文所提出的辨識架構在更適用當物體有被遮蔽之情形發生時，在沒有雜訊處理下，本論文所提出的辨識系統依舊相對於其他七個方法，有較好的辨識率，然而即使物體沒有被遮蔽的前提下，本論文的辨識架構與 nsd 整體辨識架構相比，在物體辨識率上也略為差不多，故本論文所提出的雜訊處理架構，的確是可以有效的處理雜訊來增加物體辨識率，故本論文所提出的三維雜訊處理方法，將適應性中值濾波器以及移動式最小平方法這兩個方法結合，作為三維雜訊處理用，在三維雜訊處理上有一定的強健性，尤其是針對當雷射測距儀測距誤差很大的時候，更可發揮其處理雜訊的作用。

接下來要比較的是組別 1 與組別 3，測試資料均沒有經過雜訊處理的步驟，比較物體被遮蔽前後，使用不同特徵抽取方法的辨識率之變化。

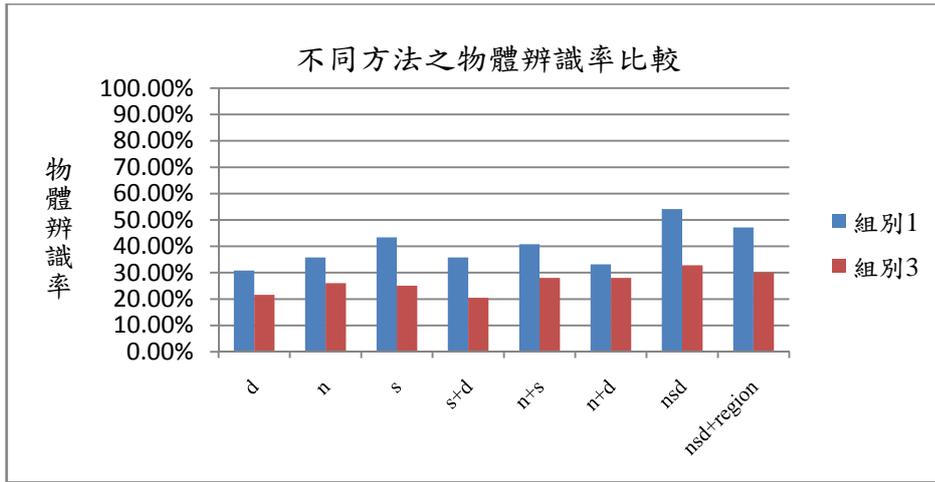


圖 4-26 組別 1 與組別 3 的各種方法之物體辨識率比較

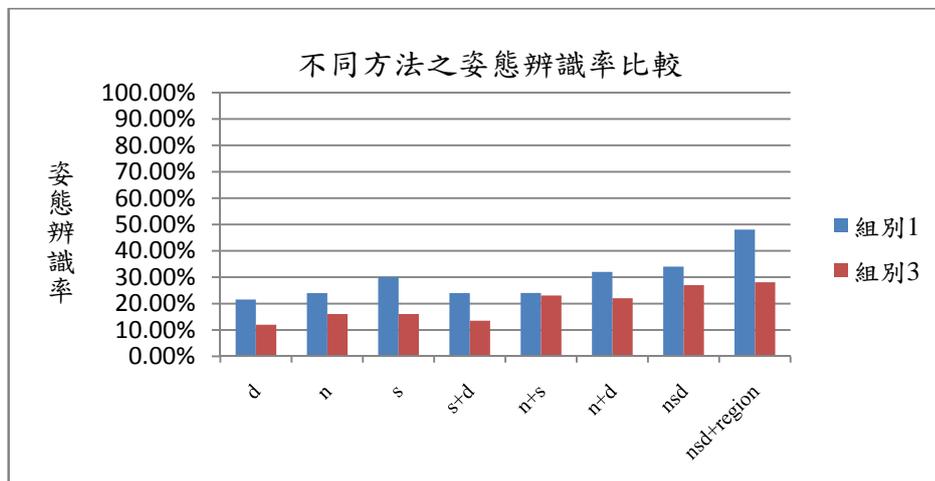


圖 4-27 組別 1 與組別 3 的各種方法之姿態辨識率比較

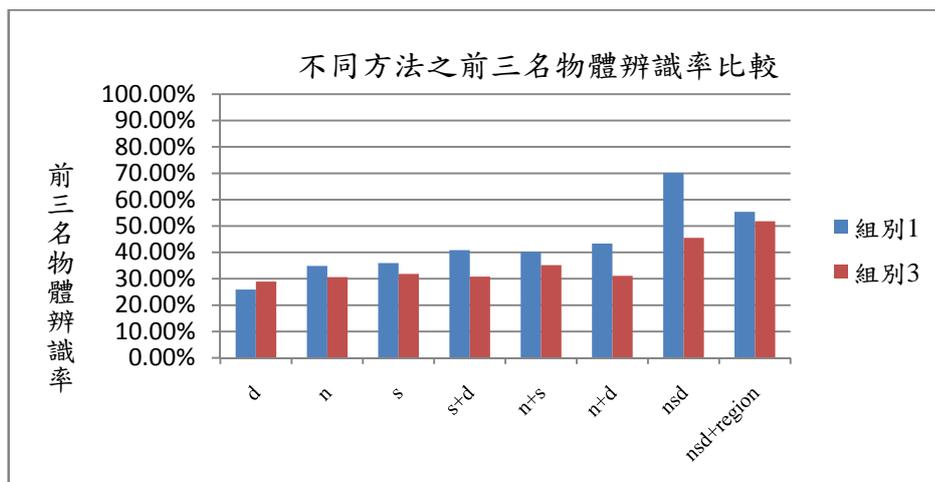


圖 4-28 組別 1 與組別 3 的各種方法之前三名物體辨識率比較

藉由組別 1 與組別 3 的比較，可以看得出來，當場景中的物體被遮蔽的時候，造成辨識率的下降是一定的，但是本論文所提出的方法 nsd+region 在辨識率下降的程度上，比其他方法較少，並且在遮蔽後，跟其他論文所提出的方法比較起來，在前三名物體辨識率上仍為最高，而其他兩個辨識率只略輸給 nsd，但是依舊比其他方法的辨識率好上許多。再來要比較的是組別 2 與組別 4，測試資料均經過了雜訊處理，比較測試資料中的物體沒遮蔽前與被遮蔽後，使用不同特徵方法之辨識率比較。

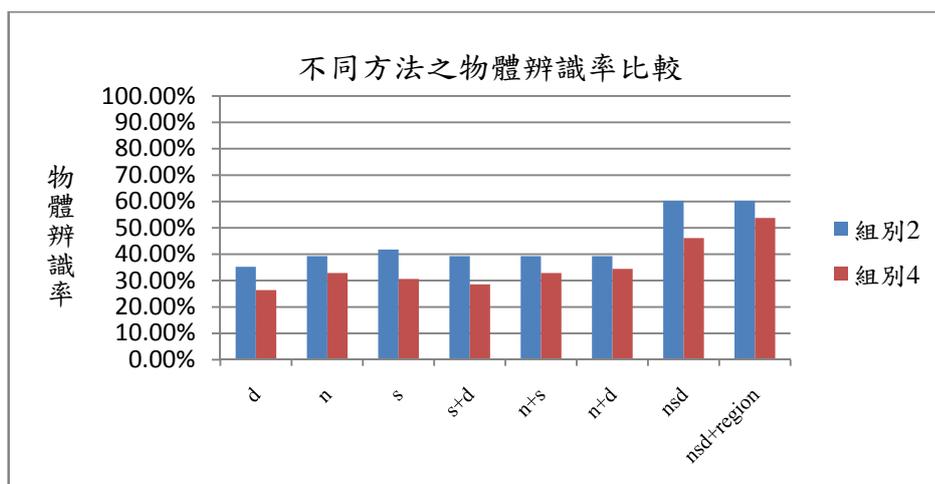


圖 4-29 組別 2 與組別 4 的各種方法之物體辨識率比較

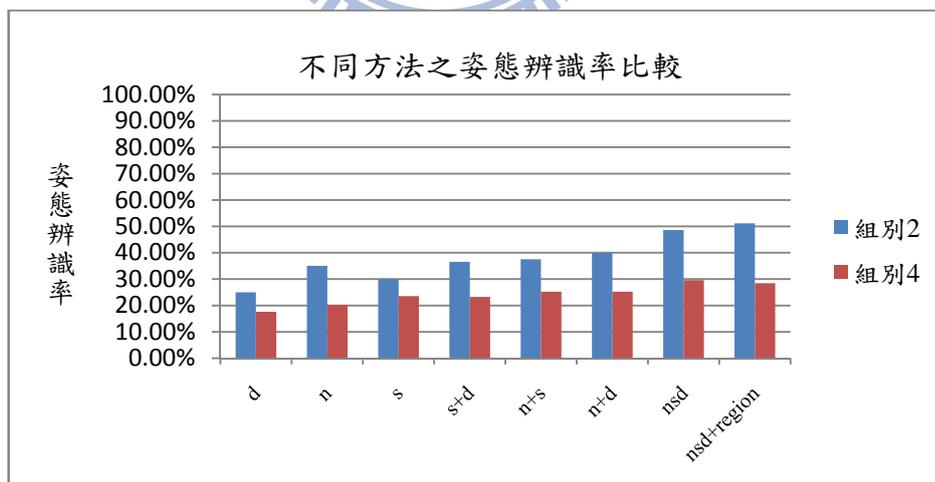


圖 4-30 組別 2 與組別 4 的各種方法之姿態辨識率比較

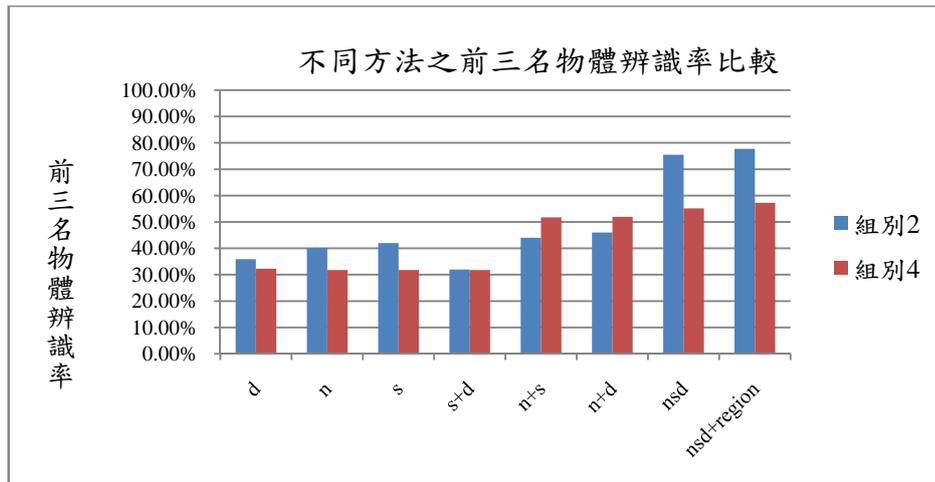


圖 4-31 組別 2 與組別 4 的各種方法之前三名物體辨識率比較

藉由組別 2 與組別 4 的比較，可以看得出來，當物體經過本論文所提出的雜訊處理架構後，在沒有遮蔽前，本論文之辨識架構 nsd+region 的物體辨識率約略跟整體辨識架構 nsd 的物體辨識率差不多，但在姿態辨識率以及前三名物體辨識率上，本論文的辨識架構相對其他方法的辨識架構，有較佳的辨識率；然而，在處理遮蔽物體上，可看得出來，本論文在處理遮蔽物體上，任一辨識率均有最佳的表現，故可以驗證本論文的辨識架構對於處理物體被遮蔽狀況是強健的。

最後，藉由組別 1 與組別 3 之比較以及組別 2 與組別之比較，本論文所以提出的辨識系統在處理物體被遮蔽時的狀況，在物體辨識率、姿態辨識率以及前三名物體辨識率上，相對於其他方法的辨識系統來說，均有較佳的辨識率，可以證明本論文在處理物體被遮蔽狀況時，有強健的辨識系統。

### 4.3 結果分析

首先，由表 4-7 及表 4-8 可看出沒有經過雜訊處理時，本論文提出的辨識系統其物體辨識率為 47.14%，在姿態辨識率上，有雜訊影響為 48.08%，姿態辨識率會高於物體辨識率，有一部分原因可能是因為拍攝時拍攝角度的誤差所導致的，然而在參考文獻[32]中，使用的測試資料是人工模擬出來的深度資訊，故沒有測距誤差的問題，故姿態辨識

率最高可高達 80.94%，但在本論文中，因為採用人工去移動視角以取得物體的不同姿態，故會有角度誤差以及測距誤差，因為這兩個因素的影響，所以導致在姿態辨識率上最高只可達到 51.12%，但扣除誤差的因素後，每個方法的姿態辨識率，幾乎可以維持一定的程度，故可以說明區域特徵雖然跟視角有關係，但是當視角變化不大的時候，並不會影響到姿態的辨識。利用拍攝物體各個視角下的姿態三維資訊來做為辨識，相對於用整個物體的三維資訊來作辨識，兩者的辨識速度下，前者會快上許多，主要是因為利用每個不同姿態作比對之過程不需要比對所有姿態的變化，只要物體姿態所對應的表面沒有改變即可，在每個姿態所對應的表面沒有完全重疊的情況下，個別比對每個姿態即可，這樣一來在所有辨識架構狀況下，相似姿態辨識率依舊有 51.12%，可證明利用姿態比對相對於整個物體比對，仍可獲得不錯的姿態辨識結果，而藉由比較表 4-7 至表 4-10 的實驗結果，本論文所提出的辨識架構，在所有方法應用於不同的操控變因下，依然有最高的姿態辨識率，故可以說明使用邊緣圖象的概念來抽取區塊特徵，可獲得較好的姿態辨識率。

第二，再由表 4-8、表 4-9 兩者的比較與表 4-10、表 4-11 兩者的比較中，可以看得出來，本論文所提出的辨識架構，在雜訊處理上有一定的作用，在八個不同的方法中，當物體沒有被遮蔽時，有六個方法使用了本論文的雜訊處理架構後，在物體辨識率上均有提高，其餘兩個方法  $s$  與  $n+s$  卻是降低的狀況，主要的原因，利用區域特徵並使用直方圖作統計時，假設深度資訊所受的雜訊程度小，物體的整體直方圖跟沒有雜訊影響的整體直方圖，幾乎是一樣，而雜訊處理架構是結合適應性中值濾波器與移動式最小平方方法來處理雜訊，這個雜訊處理方法主要是利用調整原本曲面上點之位置，去近似原本的物體曲面，而適應性中值濾波器的優點是當雜訊變化很大的時候，但是受雜訊影響的程度不大時，可以有很好的效果，缺點是可能某點並非雜訊，也會被重新更新過；反之，移動式最小平方方法的優點是當雜訊變化不大的時候，但是受雜訊影響的程度很大時，可以有很好的效果，缺點是會使物體表面本身平滑化，所以當物體受到雜訊影響程度不大時，未經雜訊處理之物體辨識率會可能高於經過雜訊處理後之物體的辨識率，但是當雜訊影響很大的時候，經過雜訊處理後的物體會比較接近原本的物體，而本論文所使用的

SR3000 雷射測距儀，測量出來的深度資訊均有蠻大的誤差，而實驗證明對於大部分方法使用此雜訊處理架構是可以提升物體在雜訊環境中的辨識率，故雜訊處理架構在抗雜訊能力上有一定的強健性。

第三，比較表 4-8 與表 4-10 時，可以看得出來，本論文提出的三維物體辨識系統利用區域特徵來進行兩階段辨識，在雜訊的影響下並且考慮物體被遮蔽的情形，先使用了本論文所提出的雜訊處理架構後來消除雜訊，可以有效的提高辨識率；然而，本論文所使用的遮蔽方法是利用一個物體任意的去遮住另外一個物體，當遮蔽百分比介於 20% 與 50% 時，物體辨識率有 53.75%，而沒有遮蔽前，物體辨識率仍有 60.32%，物體辨識率下降後依舊是所有方法中最高的，主要是因為使用了四個區域特徵作為辨識用之特徵，並用多維度直方圖統計，故只需考慮物體的表面變化之統計，並且使用了邊緣圖象的概念將物體分割成許多區塊，即使物體有某部分被遮蔽時，可比對沒被遮蔽住的區塊特徵，也可找出對應的該物體區塊。

第四，在使用實際拍攝的深度資訊作三維物體辨識系統的輸入時，並且用原本拍攝之單一物體的表面資訊作為資料庫時，由表 4-8 及表 4-10 可看出當物體沒被遮蔽與有被遮蔽的物體辨識率大約都是 50%，並沒有相差太大，然而姿態辨識率卻下降很多，從原本的 51.12% 下降至 28.47%。主要原因是因為資料庫的物體的形狀過於相似所導致的，另外一個原因是在拍攝雜亂場景時，物體與背景過於接近，使得再作場景分割時，無法將物體與背景分離；此外，用於實際取像的拍攝儀器 SR3000 之距離測量誤差過大，使得即使用適應性中值濾波器以及移動式最小平方法來處理雜訊，物體的表面依舊是凹凸不平，並且這台儀器無法精準的拍攝到複雜物體表面變化，對於過於複雜的物體，只能獲得大略的表面變化，局部的細微變化是無法取得的，導致物體的辨識率只有 50%，而姿態辨識率更下降到 20% 附近。

## 第五章 總結與未來方向

本論文所提出的針對複雜場景的三維物體切割與辨識系統，利用深度資訊來獲得此場景中所有物體的三維資訊，並考慮了因為測量誤差所形成的三維雜訊，進而實際拍攝來驗證本系統的雜訊處理架構之強健性，由第 4.2 節的數據指出，本系統對於三維雜訊程度在一定比例以上，可以有很好的消除雜訊的功用，辨識有雜訊但無遮蔽物體時，消除其三維雜訊後可大幅提升其物體辨識率，達到 60.32% 的辨識率，故在抗雜訊上有一定的強健性。而本論文所使用的區域特徵，在物體辨識上有高的辨識度，首先是當深度資訊是未處理雜訊且物體只有些許被遮蔽的情況下，物體辨識率為 47.14%，姿態辨識率為 48.08%，前三名物體辨識率為 55.40%，在處理雜訊後，物體辨識率提高至 60.32%，姿態辨識率提高至 51.12%，前三名物體辨識率提高至 77.74%，並相對於其他方法的所有辨識率，本論文所提出的辨識架構之辨識率均有最佳表現，故本論文所提出的雜訊處理架構在物體辨識上有很一定的強健性。

此外，本論文所使用了邊緣圖像的概念有效地分割物體，使得物體即使在遮蔽的環境下也可以有效地被辨識出來，並且在不過度犧牲的辨識速度的前提下，本論文提出了兩階段的三維物體辨識系統，利用第一階段選出可能的候選物體，再逐一細部比對，最後找出該物體，而由第 4.2 節的數據指出，本系統對於物體遮蔽在一定程度以下，例如：物體被遮蔽的面積為 20% 到 50% 時，物體辨識率可達到 53.75%，跟其他方法比較起來有最佳的辨識率，而姿態辨識率卻只有 28.47%，約略跟使用 nsd 整體辨識的姿態辨識率 29.62% 差不多，而其他方法的姿態辨識率也只有 20%-25% 之間，未來如何提高物體在被遮蔽的環境下之姿態辨識率，將是一大重點。

未來可改良的部分，首先是場景切割系統，由於本論文所提出的場景切割系統是針對輸入是深度資訊所設計的，如果遇到物體邊緣的深度變化與背景的深度變化幾乎是連續時，就很難分割開；此外，可能必須要依照經驗或是預先知道拍攝場景的大小來設定分割背景時的單一場景個數，才能有效的分割場景，因為當此參數過小時，可能無法有效的分割物體，過大時，可能會導致同一個物體被分割在不同的場景，未來要如何改善此場景分割演算法，也將是一個重點。

其次在於如何找出物體被遮蔽的部分，並且自動辨識出物體的遮蔽關係。首先，可以根據先找出可能的候選物體，再比對候選物體每個區塊特徵，然後找出被遮蔽的區塊

是哪個區塊；此外，由於使用邊緣圖像來分割物體，導致如果物體過於複雜，分割的區塊數可能過多，影響到辨識的速度，未來可以針對如何將相鄰的過小區塊合併，使之成為一個大區塊來減少區塊數目，進而改善原本辨識的速度。

最後在於如何將三維辨識與二維辨識結合，並且真正應用在機器人視覺中，因為二維辨識容易受到外在光源的影響，導致辨識率下降，但好處是藉由強度影像與二維辨識，可有效且快速地分割物體與背景，並且找出物體的輪廓去進行辨識，相對於，深度資訊而言，在某些情況下，只利用深度資訊是無法將物體跟地面或是背景分離的，由其是在某些視角下，目標物體與背景的深度變化幾乎是連續的，故無法有效地分離；然而，三維辨識的優點是，只利用二維影像來獲得物體的輪廓以進行辨識時，會明明表面形狀不同的物體，但可能投影在二維影像上後，有相同的輪廓，導致利用輪廓辨識的二維辨識之失敗。

未來可以將本論文應用到使用大型的雷射測距儀所量測到複雜深度資訊，如圖 5-1 所示，來辨識出複雜場景中的大型物體。

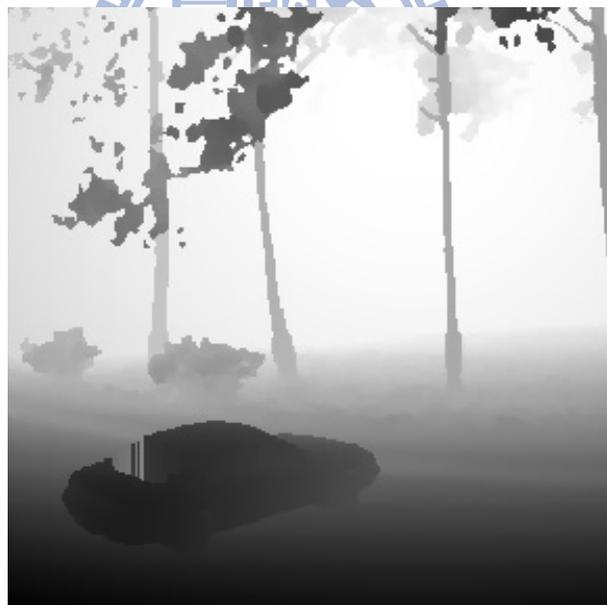


圖 5-1 雷射測距儀所測得的大型複雜深度影像

資料來源：中科院

## 參考文獻

- [1] O. Carmichael, D. Huber, and M. Hebert, "Large data sets and confusing scenes in 3-D surface matching and recognition," *Int. Conf. 3-D Digital Imaging and Modeling*, pp. 358-367, Ottawa, Ont., Canada, Oct. 1999.
- [2] G. C. Sharp, S. W. Lee, and D. K. Wehe, "Maximum-likelihood registration of range images with missing data," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 30, no. 1, pp. 120-130, Jan. 2008.
- [3] H. Gharavi and S. Gao, "3-D motion estimation using range data," *IEEE Trans. Intelligent Transportation Systems*, vol. 8, no. 1, pp. 133-143, March 2007.
- [4] S. Hussmann and T. Liepert, "Three-dimensional TOF robot vision system," *IEEE Trans. Instrumentation and Measurement*, vol. 58, pp. 141-146, Jan. 2009.
- [5] A. Hoover, G. Jean-Baptiste, X. Jiang, P. J. Flynn, H. Bunke, D. B. Goldgof, K. Bowyer, D. W. Eggert, A. Fitzgibbon, and R. B. Fisher, "A experimental comparison of range image segmentation algorithms," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 18, no. 7, pp. 673-689, July 1996.
- [6] G. Medioni and B. Parvin, "Segmentation of range images into planar surfaces by split and merge," *IEEE Conf. Computer Vision and Pattern Recognition*, pp. 415-417, 1986.
- [7] R. L. Hoffman and A. K. Jain, "Segmentation and classification of range images," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 9, no. 5, pp. 608-620, Sept. 1987.
- [8] P. J. Flynn and A. K. Jain, "BONSAI: 3D object recognition using constrained search," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 13, no. 10, pp. 1066-1075, Oct. 1991.
- [9] J. M. Jolion, P. Meer, and S. Bataouche, "Robust clustering with applications in computer vision," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 13, no. 8, Aug. 1991.

- [10] A. Bab-Hadiashar and N. Gheissari, "Range image segmentation using surface selection criterion," *IEEE Trans. on Image Processing*, vol. 15, no. 7, July 2006.
- [11] T. U. Fan, G. Medioni, and R. Nevatia, "Recognizing 3-D objects using surface descriptions," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 11, pp. 1140-1157, Nov.1989.
- [12] M. Kazhdan and T. Funkhouser, "Harmonic 3D shape matching," *Int. Conf. on Computer Graphics and Interactive Techniques*, pp. 191, 2002.
- [13] F. Stein and G. Medioni, "Structural indexing: efficient 3D object recognition," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 14, pp. 125-145, 1992.
- [14] A. E. Johnson and M. Hebert, "Using spin images for efficient object recognition in cluttered 3D scenes," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 21, no. 5, pp. 433-449, May 1999.
- [15] Y. Li, W. A. P. Smith, and E. R. Hancock, "Face recognition using patch-based spin images," *Int. Conf. Pattern Recognition*, vol. 1, pp. 408-411, 2006.
- [16] J. Assfalg, M. Bertini, A. D. Bimbo, and P. Pala, "Content-based retrieval of 3-D objects using spin image signatures," *IEEE Trans. on Multimedia*, vol. 9, no. 3, pp. 589-599, 2007.
- [17] A. E. Johnson and M. Hebert, "Using spin images for efficient object recognition in cluttered 3D scenes," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 21, no. 5, pp. 433-449, May 1999.
- [18] C. Dorai and A. K. Jain, "Cosmos-a representation scheme for 3D free-form object," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 19, no. 10, pp. 1115-1130, Oct. 1997.
- [19] C. Chua and R. Jarvis, "3-D free-form surface registration and object recognition," *Int. J. Computer Vision* 17, pp. 77-99, 1996.
- [20] J. Thirion, "New feature point based on geometric invariants for 3D image registration,"

- Int. J. Computer Vision*, vol. 18, pp. 121-137, 1996
- [21] R. L. Hoffman and A.K. Jain, "Segmentation and classification of range images," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 9, no. 5, pp.608-620, 1987.
- [22] R.C. Gonzalez and R. E. Woods, *Digital Image Processing 2<sup>nd</sup>*, Upper Saddle River, Prentice-Hall Inc., NJ, 2002.
- [23] P. Lancaster and K. Salkauskas, "Surfaces generated by moving least squares methods," *Mathematics of Computation*, vol. 37, no. 155, Jul. 1981.
- [24] M. Alexa, J. Behr, D. Cohen-Or, S. Fleishman, D. Levin, and C. T. Silva, "Computing and rendering point set surface," *IEEE Trans. Visualization and Computer Graphic*, vol. 9, pp.3-15, Jan. 2003.
- [25] R. M. Haralick and L. G. Shapiro, *Computer and Robot Vision*, vol. 1, Addison-Wesley, pp.28-48, 1992.
- [26] I. Sobel and G. Feldman, "A 3x3 isotropic gradient operator for image processing", (presented at a talk at the Stanford Artificial Project), 1968, unpublished.
- [27] J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. PAMI-8, pp. 679-698, Nov. 1986.
- [28] O. R. P Bellon, A. I. Direne, and L. Silva, "Edge detection to guide range image segmentation by clustering techniques," in *1999 Proc. IEEE Int. Conf. Image Processing*, vol. 2, pp. 725-729, 1999.
- [29] N. Yokoya and M. Levine, "Range image segmentation based on differential geometry: a hybrid approach," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 11, pp. 643-649, 1989.
- [30] T. Yung-Kong and Azriel Rosenfeld, *Topological Algorithms for Digital Image Processing*, NY: Elsevier Science Inc., 1996.
- [31] L. Vincent, "Morphological grayscale reconstruction in image analysis: applications and efficient algorithms," *IEEE Trans. Image Processing*, vol. 2, no. 2, pp. 176-201, April

1993.

- [32] G. Htzzel, B. Leibe, P. Levi, and B. Schiele, “3D objects recognition from range image using local feature histograms,” in *2001 Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition*, vol. 2, pp. II-394-II-399, 2001.
- [33] S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry*, Chap. 2 and 3, vol.1, John Wiley & Sons, NY, 1969.
- [34] H. W. Guggenheimer, *Differential Geometry*, chap. 10, McGraw-Hill, NY, 1977.
- [35] J. Koenderink and A. J. van Doorn, “Surface shape and curvature scales,” *Image and Vision Computing*, vol. 10, pp. 557-565, 1992.
- [36] A. Gray, E. Abbena, and S. Salamon, *Modern Differential Geometry of Curves and Surfaces with Mathematica 3<sup>rd</sup>*, CRC press, p342-p377, NY 2006.
- [37] P. Krsek, G. Lukacs and R. R. Martin, “Algorithms for computing curvatures from range Data”, *Mathematics of Surfaces VIII*, pp. 1-16, 1998.
- [38] M. J. Swain and D. H. Ballard, “Color indexing,” *IJCV*,7(1):11-32, 1991.
- [39] S. Kullback, *Information Theory and Statistics*, Wiley, NY, 1959.
- [40] B. Schiele and J. Crowley, “Recognition without correspondence using multidimensional receptive field histograms,” *IJCV*, 36(1):31-52, 2000.