

國立交通大學

電機與控制工程學系

碩士論文

以主成分影像作特徵抽取之三維模型檢索系統



A 3D Models Retrieval System Based on Feature Extraction of

Principle Image

研究生：徐道智

指導教授：林昇甫 博士

中華民國九十八年七月

以主成分影像作特徵抽取之三維模型檢索系統
A 3D Models Retrieval System Based on Feature Extraction of
Principle Image

研究生：徐道智

Student : Dou-Chih Hsu

指導教授：林昇甫 博士

Advisors: Dr. Sheng-Fuu Lin

國立交通大學

電機與控制工程學系



A Thesis

Submitted to Department 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

以主成分影像作特徵抽取之三維模型檢索系統

學生：徐道智

指導教授：林昇甫博士

國立交通大學電機與控制工程學系（研究所）碩士班

摘要

本論文提出一個以主成分影像作特徵抽取之三維模型檢索系統，首先藉由主成分分析法(principal component analysis, PCA)找出每個模型的主軸，使相似模型有相似主軸，以此克服部分旋轉變量，之後再以具有旋轉不變性的傅立葉描述子(Fourier descriptors)和 Zernike 矩(Zernike moments)並加上一個三維特徵：主軸投影圖的相關大小做為檢索特徵。其想法來自光場描述子(light field descriptors, LFD)，但藉由主成分分析法改進其需要大量圖形的缺點，並利用新的主軸投影圖相關大小特徵，使對於相似物體的檢索有更大的容忍度。之後由實驗證明本論文對於因相似物體部分區域不同所造成的偏軸情形能有效克服，而且相對於 LFD 檢索速度加快，儲存需求大幅下降。最後使用普林斯頓形狀評判標準(Princeton shape benchmark, PSB)跟四個不同特徵之資料庫做檢索效能評判，證明本論文較其他方法有較好的檢索效能。

A 3D Models Retrieval System Based on Feature Extraction of Principle Image

Student: Dou-Chih Hsu

Advisors: Dr. Sheng-Fuu Lin

Department of Electrical and Control Engineering
National Chiao Tung University

ABSTRACT

In this thesis, a 3D models retrieval system based on feature extraction of principle images is proposed. First of all, find the main axis of all models by principal component analysis (PCA). Let similar models have similar main axis in order to overcome variables of rotation. And then use the rotation invariance features Fourier descriptors and Zernike moments and add a 3D feature: the size relevant of the projection image. The idea is from the light field descriptor (LFD), but use PCA to overcome the shortcomings that it needs a lot of images. And use the new feature: the size relevant of the projection images to let the retrieval system have greater tolerance for similar models. And prove that it can overcome the rotation variables of main axis when similar models have different parts. And it has faster retrieval speed and smaller storage with LFD. In the last, determine the effectiveness of retrieval systems by Princeton shape benchmark (PSB) and four different characteristic databases. And prove this thesis has better performance than other methods.

致 謝

首先要以最誠摯的謝意，感謝我的指導教授林昇甫博士，感謝他在研究上的啟迪與指導，以及細心校閱本論文的謬誤之處。同時也要感謝本實驗室的學長晉嘉、永吉、士哲、國育、逸章、啟曜、俊偉；一起同甘共苦的同學長安、昆義、子航、煜維以及學弟們，有如此龐大的研究團隊，使得我在遇到問題時都能很充裕的能找到人來對我伸出援手，讓論文的研究能夠順利完成，並且在碩士生涯中有好大一票的好伙伴一起成長、一起學習；不只是書本上的知識，還有更多更多有趣有用的技能以及想法。過程中也許有苦有樂，但是在回憶中卻是既燦爛又豐富。

最後要感謝我親愛的父母，感謝他們的養育之恩以及在我求學生涯中給我最大的鼓勵與支持，使我得以在精神與生活上無後顧之憂，順利完成學業。



目錄

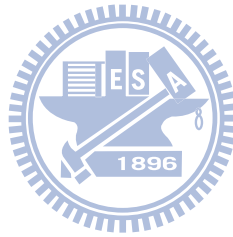
摘要	i
英文摘要	ii
致 謝	iii
目 錄	iv
表目錄	vi
圖目錄	vii
第一章 緒論	1
1.1 研究動機與背景	1
1.2 相關研究	2
1.3 論文架構	4
第二章 相關知識與理論	6
2.1 主成分分析法(principal component analysis , PCA).....	6
2.2 擷取特徵的方法	10
2.2.1 輪廓提取	10
2.2.2 傅立葉描述子(Fourier descriptors)	11
2.2.3 Zernike 矩	14
2.3 相似度分析	16
2.4 檢索評判標準	17
第三章 檢索系統	20
3.1 離線架構	21
3.1.1 前處理	22
3.1.2 特徵抽取	25
3.2 系統在線架構	28
3.2.1 相似度分析	29
第四章 實驗結果與分析	31
4.1 實驗平台	31
4.2 主成分分析法驗證	31
4.2.1 主成分分析法的旋轉變量修正	31
4.2.2 不同求取主成分分析的方法實驗	40
4.3 檢索效能評判	43
4.3.1 PSB 資料庫檢索	43
4.3.2 小型資料庫檢索	45
4.3.3 計算複雜度與空間需求	53
4.4 特徵需求實驗	54
4.4.1 各別特徵重要性	55
4.4.2 大型資料庫的特徵需求	55

4.4.3 小型資料庫的特徵需求	58
4.5 實驗分析	61
第五章 結論與未來發展	64
參考資料	65



表目錄

表 3.1 Zernike 矩每階所含的矩。.....	27
表 4.1 有無旋轉修正比較表。.....	35
表 4.2 不同方法求取主成分分析的效能。.....	43
表 4.3 十四種方法在 PSB 中的效能。.....	45
表 4.4 各方法在動物資料庫的效能比較表。.....	46
表 4.5 各方法在家具資料庫的效能比較表。.....	48
表 4.6 各方法在常用工具資料庫的效能比較表。.....	50
表 4.7 各方法在交通工具資料庫的效能比較表。.....	51
表 4.8 儲存一個模型的空間需求表。.....	53
表 4.9 時間需求表。.....	54
表 4.10 在大型資料庫中不同特徵數檢索效能表。.....	58
表 4.11 在小型資料庫中不同特徵數的效能表。.....	60



圖目錄

圖 1.1 幾何形狀分布的各種形狀函數。	3
圖 1.2 正十二面體的二十個頂點。	4
圖 2.1 輪廓追蹤圖。	11
圖 2.2 不同傅立葉特徵，(a)中心傅立葉、(b)面積傅立葉、(c)角度傅立葉、(d)複數平面傅立葉、(e)極座標傅立葉、(f)綜和角度跟中心距離傅立葉。	13
圖 2.3 抽取 Zernike 矩特徵向量流程圖。	16
圖 2.4 精確度於回傳率代表圖。	18
圖 3.1 檢索系統架構。	21
圖 3.2 離線架構圖。	22
圖 3.3 前處理流程圖。	23
圖 3.4 三維模型旋轉圖。	24
圖 3.5 格子點轉換假碼。	25
圖 3.6 三主軸投影圖。	26
圖 3.7 三主軸輪廓圖。	26
圖 3.8 Zernike 矩每階所形成的不同單位圓。	28
圖 3.9 系統在線架構。	29
圖 4.1 編號第 1119 飛機，左圖為初始位置，右圖經過旋轉跟平移。	32
圖 4.2 編號第 1119 飛機的三軸投影圖：(a)對第一主軸投影、(b)對第二主軸投影、(c)對第三主軸投影。	32
圖 4.3 編號第 1137 的飛機三軸投影圖：(a)原始未經旋轉的圖形(b)對第一主軸投影、(c)對第二主軸投影、(d)對第三主軸投影。	33
圖 4.4 編號第 1160 的飛機三軸投影圖：(a)原始未經旋轉的圖形(b)對第一主軸投影、(c)對第二主軸投影、(d)對第三主軸投影。	34
圖 4.5 編號 1162 的飛機三軸投影圖：(a)原始未經旋轉的圖形(b)對第一主軸投影、(c)對第二主軸投影、(d)對第三主軸投影。	34
圖 4.6 編號 810 的椅子的三軸投影圖(椅背跟椅腳長度標準)：(a)原始未經旋轉的圖形(b)對第一主軸投影、(c)對第二主軸投影、(d)對第三主軸投影。	36
圖 4.7 編號 814 的椅子的三軸投影圖(椅背長度短，椅腳長度標準)：(a)原始未經旋轉的圖形(b)對第一主軸投影、(c)對第二主軸投影、(d)對第三主軸投影。	36
圖 4.8 編號 816 的椅子的三軸投影圖(椅背長度短，椅腳長度長)：(a)原始未經旋轉的圖形(b)對第一主軸投影、(c)對第二主軸投影、(d)對第三主軸投影。	37
圖 4.9 針對編號 816 在家具資料庫做檢索的結果，左上綠色的為檢索物體依序往右再往下為檢索到的順序，藍色代表正確，紅色代表錯誤。	37

圖 4.10 編號第 110 的兔子投影圖:(a)原始未經旋轉的圖形(b)對第一主軸投影、 (c)對第二主軸投影、(d)對第三主軸投影。	38
圖 4.11 編號第 111 的兔子投影圖:(a)原始未經旋轉的圖形(b)對第一主軸投影、 (c)對第二主軸投影、(d)對第三主軸投影。	39
圖 4.12 編號 112 的兔子投影圖:(a)原始未經旋轉的圖形(b)對第一主軸投影、 (c)對第二主軸投影、(d)對第三主軸投影。	39
圖 4.13 PSB 資料庫中編號 3 的原圖。	40
圖 4.14 PSB 資料庫中編號 2 的原圖。	40
圖 4.15 點求主成分分析(上排為圖 4.13 的三軸投影圖，下排為圖 4.14 的三軸 投影圖。)	41
圖 4.16 權重主成分分析(上排為圖 4.13 的三軸投影圖，下排為圖 4.14 的三軸 投影圖。)	41
圖 4.17 格子點化後再以每個格子點求取主成分分析(上排為圖 4.13 的三軸投影 圖，下排為圖 4.14 的三軸投影圖。)	42
圖 4.18 連續主成分分析(上排為圖 4.13 的三軸投影圖，下排為圖 4.14 的三軸 投影圖。)	42
圖 4.19 動物資料庫精準度對回傳率圖。	46
圖 4.20 動物資料庫的類別對 DCG 圖。	47
圖 4.21 家具資料庫精準度對回傳率圖。	48
圖 4.22 家具資料庫的類別對 DCG 圖。	49
圖 4.23 常用工具資料庫精準度對回傳率圖。	50
圖 4.24 常用工具資料庫的類別對 DCG 圖。	51
圖 4.25 交通工具資料庫精準度對回傳率圖。	52
圖 4.26 交通工具資料庫的類別對 DCG 圖。	52
圖 4.27 各別特徵對重要性。	56
圖 4.28 各別傅立葉描述子特徵的效能。	56
圖 4.29 各別 Zernike 矩特徵的效能。	56
圖 4.30 各別傅立葉描述子特徵的效能。	59
圖 4.31 各別 Zernike 矩特徵的效能。	59

第一章 緒論

近年來，由於電腦運算速度越趨快速，電腦視覺(computer vision)的應用也越來越廣泛，而 3D 模型的製作早已不是問題，不論是在醫療上，如 DNA 模型，腫瘤，身體內臟等，或者是工業上，如模具製造，建築設計等，對於 3D 模型的使用越來越常見，且相關的模型也已經都製作完成。而如何快速找尋並使用這些已經製造出來的 3D 模型或其他相似的模型也越來越重要。而實際可應用在三維場景中尋找所要的物體將，如在雷射或超音波掃描後產生的三維場景中尋找目標物體如炸彈，刀槍等危險物品、或是找尋身體內的惡性腫瘤，骨刺等。如今 3D 模型檢索的演算法已經提出不少種類，如何改進這些舊有的方法或提出新的方法是一大課題。

在本章第一節將說明本論文的研究動機與背景並簡單介紹本論文的方法，第二節整理相關已經被提出的演算法並說明一個好的演算法應該具備什麼條件，第三節提出本論文的架構。



1.1 研究動機與背景

如今 2D 圖形檢索早已在網路上發展以久，但 3D 模型卻還只有少數系統可以簡單搜索，且大部分的檢索方式還是以圖形旁的關鍵字搜索而非實際物體的特徵。不像 2D 圖形早就可以使用形狀跟顏色等不同特徵作檢索，而這種以內容為基礎的檢索方法 (content-based retrieval) 和分類系統，要如何延伸到 3D 模型上抑或者是使用不同的演算法來抽取 3D 模型特有的特徵以檢索相似模型，即為本論文的動機。

本篇論文提出一個以基於主軸分析並藉由傅立葉描述子、Zernike 矩和不同角度投影圖形的大小相關性特徵抽取的 3D 模型檢索系統。首先藉由主成分分析法找出三個主軸並分別對三個主軸做投影，藉由投影圖形求取傅立葉描述子跟 Zernike 矩並同時取出除了傅立葉描述子跟 Zernike 矩之特徵之外的第三種特徵：固定整體模型比例大小比較三個投影圖的形狀比例關係。此特徵是直接藉由計算傅立葉描述子跟 Zernike 矩之特徵

而不再做大小比例的正規化而完成，其緊密跟傅立葉描述子和 Zernike 矩結合而成，而不需再獨立出來。最後藉由此三種特徵作為檢索 3D 物體的依據，其稍為類似 D. Y. Chen 等人所提出的光場描述子(light field descriptors, LFD)[7]，但本論文所提出的方法因為多考慮了圖像的相對大小，且不再是以 10 個描述子，每個描述子 10 個面當做旋轉變量的克服，取而代之以主成分分析法大幅降低所需要的投影面，使整體所需時間和儲存容量都大幅降低，最後藉由 P. Shilane 等人所提出的普林斯頓形狀評判標準(Princeton shape benchmark, PSB)[3]，使用其中所分類好的資料庫和其他較小但有不同特徵的資料庫跟其他的方法比較，驗證本論文所提出的演算法的檢索效能。

1.2 相關研究

3D 模型檢索就是從現有的 3D 模型資料庫中，找出分類相同的模型，並回傳給使用者，所以主要的研究就是以各種不同的演算法抽取模型的特徵來檢索物體。而一個好的演算法應該要具備以下條件：(1)運算快速，(2)正確率高，(3)需求空間不大，(4)能忍受小部分的不同，(5)針對同物體但不同大小角度能得到相同的答案，(6)能容忍小量雜訊，(7)對不相似的物體得到不相同的特徵。而 3D 模型檢索的特徵在 J. H. Tangelder 等人[1]提出的論文中，分成基於全域特徵(global features)相似，基於部分特徵(local features)相似，基於空間分布(spatial map)相似，基於圖像(graph)相似，跟基於幾何(geometry)相似來區別。以下針對幾種常見的檢索演算法來做說明：

1. 幾何形狀分布(shape distributions)：

這方法是由 R. Osada 等人[6]在 2002 年提出，其屬於基於全域特徵的方法，其特性就是強健，運算簡易，能容忍雜訊且不需做特別的處理本身就有旋轉不變性，但缺點是鑒別度不高，完全不同的物體卻可能會有相似的特徵。本方法以形狀分佈作為特徵，使用幾何形狀函數(shape function)把 3D 模型中的形狀轉為機率分佈，在以簡單數學比較兩個模型的特徵分佈曲線。其作法首先藉由統計模型中每一個面的面積，並基於其面積大小隨機取出適當的點位置作為整體模型的代表點，然後隨

機以這些點的組成的幾何形狀函式作為特徵。幾何形狀函式可分為：A3、D1、D2、D3、D4，如圖 1.1 所示。A3 是從上述的代表點中每次隨機找三個點統計其所夾的夾角作為特徵；D1 是隨機從代表點中每次取出一點並計算到一固定點的距離作為特徵；D2 是隨機從代表點中每次取兩點以兩點距離作為特徵；D3 是隨機從代表點中每次取三點並計算其面積作為特徵；D4 是隨機從代表點中每次取 4 點計算其體積作為特徵。其中以 D2 效能最佳，A3 次之。

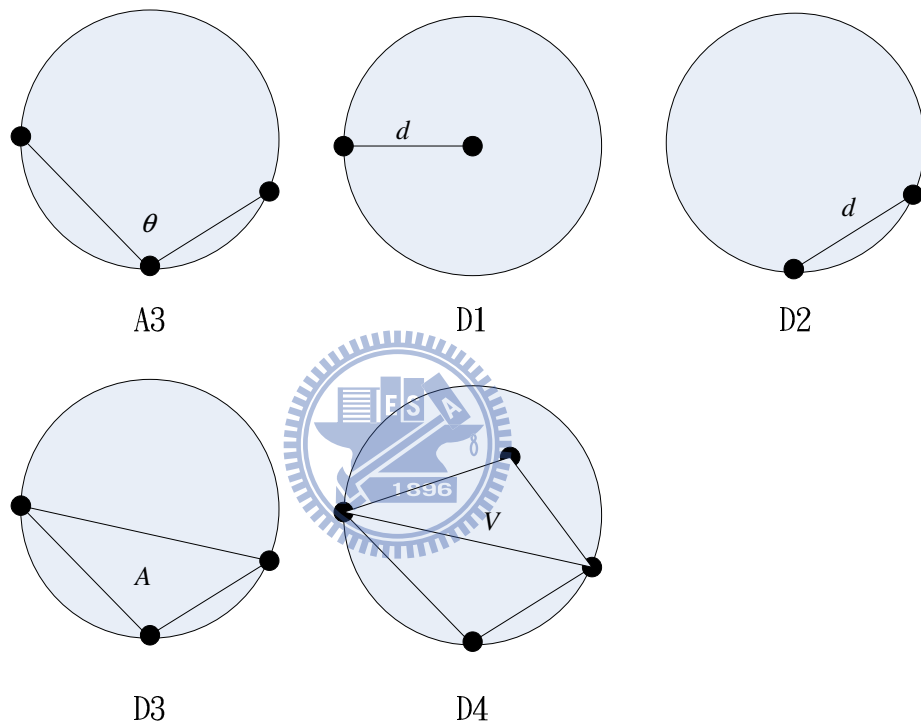


圖 1.1 幾何形狀分布的各種形狀函數。

2. 球形諧波(spherical harmonics, SH)：

這方法是由 M. Kazhdan 等人[8]-[10]在 2003 提出，其屬於基於空間分布的方法，其特性是具有旋轉不變性且鑒別度高，但運算複雜。本方法主要是把物體藉由頻率轉換變成頻率域特徵表示，有如 2D 圖形做傅立葉轉換一樣，不同頻率有不同代表性的形狀。作法是首先把物體從點跟面的資料型態轉為以格子點儲存，再依照點到中心的距離分成幾個同心圓區塊，之後把每一層轉為極座標並作頻率域轉換，

最後取其前幾項特徵向量作為物體的特徵向量，做為檢索依據。

3. 光場描述子(light field descriptors, LFD)：

這方法是由 D. Y. Chen 等人[7]在 2003 年提出，其屬於基於幾何形狀相似比對的方法，特性是使用多張投影圖形並搭配 2D 圖形的檢索方法，其對於形狀的鑒別度高，但如需處理不同角度的投影面，需要大量儲存空間及運算，而在 PSB[3]論文中的實驗其檢索效能為最佳。作法是首先把 3D 物體以正十二面體的 20 個頂點作為不同視角如圖 1.3，從這 20 個頂點往中心投影轉換成 20 個 2D 平面，但其中有 10 個面為另 10 個面的反面故實際上只需找 10 個面，之後再使用 2D 圖形的特徵抽取的方法：藉由輪廓比較的傅立葉描述子跟區域比較的 Zernike 矩求出特徵向量，最後比對其形狀特徵，找出相似物體。但此種方法對於每個模型都需要儲存大量的特徵向量，故如何降低儲存是此種方法的一大課題。

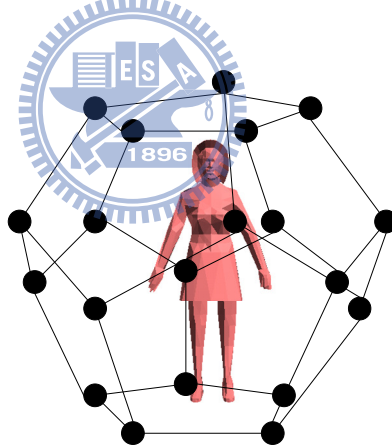


圖 1.2 正十二面體的二十個頂點。

1.3 論文架構

本論文的組織架構如下：

第二章介紹本論文所使用的主成分分析、對物體做特徵抽取的各種方法、相似度分析跟各種檢索好壞的評判標準。

第三章介紹本論文檢索 3D 物體的系統，並詳述其每個步驟。

第四章為實驗並分析本系統對於旋轉變量的修正效果，之後再針對不同資料庫做檢索並分析，最後針對速度與儲存需求加以討論。

第五章總結本論文並指出未來可能研究的方向。



第二章 相關知識與理論

3D 模型檢索系統是從現有 3D 模型資料庫中，尋找跟使用者提供的資訊相符合的 3D 模型，主要的研究目的是希望能有效率且正確的找到相關模型，所以特徵的選定是最為重要，首先選定的特徵最好要能對平移(translation)、旋轉(rotation)和比例大小(scale)具有不變性。使兩個相似或相同的物體就算初始角度，位置跟比例大小都不相同，但也都都能正確檢索找到，故本章第一節首先介紹克服旋轉變量所使用的方法：使用主成分分析法找尋主軸，使相似物體有相同座標軸；第二節是本論文所使用的擷取特徵的方法：傅立葉描述子跟 Zernike 矩並介紹在作傅立葉描述子前的輪廓提取方法；第三節是如何判定相似度，最後第四節則是介紹評判檢索系統效能好壞的方法。

2.1 主成分分析法(principal component analysis, PCA)

主成分分析法最早是由美國心理學家 Charles Spearman 在 1904 年提出，之後由霍特林(Hotelling)加以發展的一套統計系統，其為一種簡化數據集的技術。它是一個線性變換，藉由把數據轉換到另一個新的坐標系統中，使得任何數據投影的第一大均方差在第一個坐標(稱為第一主成分)上，第二大均方差在第二個坐標(第二主成分)上，依次類推。因此主成分分析法經常用在減少數據的維數且同時保持數據的均方差貢獻最大的特徵、或者是用在圖形找最大成分的主軸。在本篇論文中主要藉由主成分分析法找出前三主軸，使相似物體的座標能轉換為相似的座標系統，使其在相似的型態下，投影出三個相似的投影面，以克服旋轉變量。雖然在 T. FunkHouser 等人[8]所提的論文，提出主成分分析法對於部分相似物體會因為部分區域比例形狀大小不同，造成主軸偏斜的問題，使之後檢索錯誤，但在本篇論文最後的實驗可知：在基於 Zernike 矩和傅立葉描述子兩種具備旋轉不變性的特徵抽取方法，並加上在相似度比較時作旋轉比較找最佳視角，使對於這項缺點所受的影響其實很小並不顯著。

計算主成分分析法，首先必須針對模型上所有點 $\mathbf{V} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_N\}$ ，找出模型重

心，以此克服平移變量，其中點的順序並不會影響到最後的結果。

$$\bar{\mathbf{v}} = \sum \mathbf{v}_k, \quad k=1, \dots, N, \quad (2-1)$$

$$\mathbf{v}_k = \{x_k, y_k, z_k\}, \quad (2-2)$$

接著將 \mathbf{v}_k 減去重心 $\bar{\mathbf{v}}$ ，使各點對於重心具有平移不變性：

$$\tilde{\mathbf{v}}_k = \mathbf{v}_k - \bar{\mathbf{v}}, \quad k=1, \dots, N, \quad (2-3)$$

之後產生資料矩陣向量：

$$\mathbf{I} = [\tilde{\mathbf{v}}_1, \tilde{\mathbf{v}}_2, \dots, \tilde{\mathbf{v}}_N], \quad (2-4)$$

其中每列是減去重心之後的點。再建立資料矩陣 \mathbf{I} 的共變異矩陣 \mathbf{C} ：

$$\mathbf{C} = \mathbf{I} \cdot \mathbf{I}^T, \quad (2-5)$$

其中共變異矩陣 \mathbf{C} 的維度是 3×3 。之後再求解共變異矩陣 \mathbf{C} 的特徵值跟特徵向量。

$$\lambda_j \mathbf{e}_j = \mathbf{C} \mathbf{e}_j, \quad (2-6)$$

可以得到共變異矩陣的特徵向量 \mathbf{e}_j 與對應的特徵值 λ_j 。特徵向量 \mathbf{e}_j 也就是經過主成份分析後得到的主軸，之後在依照其特徵值大小，由大而小依序排列形成： λ_1 、 λ_2 、 λ_3 ，而相對應的 \mathbf{e}_1 也就是第一主軸， \mathbf{e}_2 為第二主軸， \mathbf{e}_3 為第三主軸。

$$\tilde{\mathbf{e}} = [\mathbf{e}_1 \ \mathbf{e}_2 \ \mathbf{e}_3], \quad (2-7)$$

轉換座標方程式如下：

$$\mathbf{v}_k = \tilde{\mathbf{e}} \cdot \mathbf{v}_k, \quad k=1, \dots, N, \quad (2-8)$$

之後為了使模型整體方向更一致性，讓有較多點的部份為正方向，較少的為負方向。首先令一對角化矩陣 \mathbf{F} ：

$$\mathbf{F} = \text{diag}(\text{sign}(f_x), \text{sign}(f_y), \text{sign}(f_z)), \quad (2-9)$$

其中 $f_x = \sum_{k=1}^N \text{sign}(x_k) \cdot x_k^2$ ， f_y 跟 f_z 作法也相同， sign 為求函數的正負符號。而為了使之後擁有比例大小的一致性，再令 s 為比例大小參數：

$$s = [(s_x^2 + s_y^2 + s_z^2) / 3]^{1/2}, \quad (2-10)$$

其中 s_x ， s_y ， s_z 為各點到主軸的平均值。其次把每點跟 \mathbf{F} 和 s^{-1} 相乘，則得到最後轉換完成的新座標。

$$\tilde{\mathbf{v}}_k = s^{-1} \cdot \mathbf{F} \cdot \mathbf{v}_k, \quad k = 1, \dots, N. \quad (2-11)$$

但如果只使用模型的點座標而不考慮面的話，所求得的新座標對於相似物體還是不夠一致，所以 D. V. Vranić 等人[18]提出權重主成分分析法，藉由每點相鄰面的面積找出每點的權重值以求出新的主軸方向。首先計算出模型上每個三角形網格的面積：

令 $\mathbf{V} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_N\}$ 為模型上關聯到網格的所有點。

$$\mathbf{v}_k \in \mathbb{R}^3, \quad k = 1, 2, \dots, N, \quad (2-12)$$

藉由相鄰面積計算每點的權重值 w_k ：

$$w_k = \frac{N \cdot S_k}{3S}, \quad k = 1, \dots, N, \quad (2-13)$$

其中 S_k 是每點相鄰面的總面積和， S 為模型的總面積，且 $\sum_{k=1}^N w_k = N$ 。之後再計算增加權重後的新重心 \mathbf{m} ：

$$\mathbf{m} = \frac{1}{N} \sum_{k=1}^N w_k \mathbf{v}_k, \quad (2-14)$$

同樣為了求特徵值跟特徵向量，必須再先以下式求新的共變異矩陣 \mathbf{C} ：

$$\mathbf{C} = \frac{1}{\sum w_k} \sum_{k=1}^N w_k (\mathbf{v}_k - \mathbf{m}) \cdot (\mathbf{v}_k - \mathbf{m})^T, \quad (2-15)$$

最後再以之前求標準主成分分析法相同的方法(2-6)，求其特徵值跟特徵向量，並以特徵值大小遞減排序，即可得到經過加權的主成分分析法的主軸，使經過轉換的座標能更加精確，減少被不均匀分散的點影響的程度能有所降低。

但是權重主成分分析法，對於大小不平均的三角面所做的主成分分析並不夠完備，故之後 D. V. Vranić 等人[19]又提出連續主成分分析法(continuous principal component analysis, CPCA)，藉由計算出三角面內所有點對重心的向量，找出更完美的主軸方向。

首先把模型上的所有面設為一三角面集合，所有點設為一點集合：

$$\mathbf{T} = \{\mathbf{t}_1, \dots, \mathbf{t}_m\}, \quad \mathbf{t}_i \subset \mathbb{R}^3, \quad (2-16)$$

$$\mathbf{P} = \{\mathbf{p}_1, \dots, \mathbf{p}_n\}, \quad \mathbf{p}_i = (x_i, y_i, z_i) \in \mathbb{R}^3, \quad (2-17)$$

並設定好一個對應表記住哪三點對哪個面，之後再令 $\mathbf{I} = \bigcup_{i=1}^m \mathbf{t}_i$ 為所有三角面的點集合。

之後為了找到一個訪射圖(affine map) $\tau: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ，使任意連續的 σ 能具備有平移，旋轉跟比例大小的不變性，如下：

$$\begin{aligned} \tau(\mathbf{I}) &= \tau(\sigma(\mathbf{I})), \\ \sigma(\mathbf{I}) &:= \{\sigma(\mathbf{v}) \mid \mathbf{v} \in \mathbf{I}\}, \end{aligned} \quad (2-18)$$

首先跟權重主成分分析法同樣的令 S_i 為 T_i 的面積，並假設面跟面的交界值為零，使整個模型面積為：

$$S := S_1 + \dots + S_m = \int_{\mathbf{I}} d\mathbf{v}, \quad (2-19)$$

同樣為了克服平移變量，把所有點減去重心 c ，形成新的一組 I_1 ：

$$\mathbf{I}_1 := \mathbf{I} - \mathbf{c} = \{\mathbf{u} \mid \mathbf{u} = \mathbf{v} - \mathbf{c}, \mathbf{v} \in \mathbf{I}\}, \quad (2-20)$$

再來是藉由積分求新的變異矩陣：

$$\mathbf{C} = \frac{1}{S} \int_{I_1} \mathbf{v} \cdot \mathbf{v}^T dv, \quad (2-21)$$

之後再以式(2-6)求解 \mathbf{C} 的特徵值跟特徵向量 $\tilde{\mathbf{e}}$ ，同樣依照特徵值大小由大往小依序排列。為了使方向一致，由式(2-9)求對角矩陣 \mathbf{F} ，並且為了使比例大小有一致性，由(2-10)求比例參數 s 。最後得到訪射圖方程式如下，其也就是座標轉換公式：

$$\tau(v) = s^{-1} \cdot \mathbf{F} \cdot \tilde{\mathbf{e}} \cdot (\mathbf{v} - \mathbf{c}). \quad (2-22)$$

2.2 擷取特徵的方法

擷取特徵是檢索或辨識系統中最重要的一環，良好的特徵不僅能使系統檢索快速，並且也能提高正確率。在檢索系統中通常可將特徵分為全域特徵，區域特徵，全域特徵通常是以統計為主，其特性就是計算簡單，檢索快速，系統強健，但效果通常不佳，如 D2。區域性特徵特性是計算較為複雜，較易受雜訊干擾，容易受物體本身姿態影響，但檢索效果較佳。而本篇論文所使用的特徵，傅立葉描述子跟 Zernike 矩皆為區域性特徵，其在 D. Zhang 等人[17]的論文早已證明其在 2D 圖形檢索上的效能，且合併使用更能得到最佳效能。以下首先介紹為了傅立葉描述子所要先做的輪廓提取的方法，再來是各種傅立葉描述子，最後則是 Zernike 矩的特徵抽取方法。

2.2.1 輪廓提取

輪廓提取的目的是獲得影像的外部輪廓特徵，是為影像形狀分析作準備，在本篇論文是為了給傅立葉描述子作預先的轉換。輪廓提取的方法有很多種，如梯度法，或者是用型態學先對原影像進行腐蝕，在藉由原影像減去腐蝕後的影像，更簡單的方法是尋找圖中每個黑點，且其八個相鄰點皆為黑色時，即刪除該點，但後兩種方法皆可能會得到內部輪廓，而這在傅立葉描述子的運算中並不需要，而本論文所處理的圖形也只是二值影像，並不需要使用梯度法。所以最後採用的是輪廓追蹤法，如此一來也可把輪廓點依序排好不需在做好輪廓圖後重新排序。

輪廓跟蹤的基本方法是：首先找到第一個邊界像素其為最左下方的黑點，可由下往上，並由左往右尋找第一個遇見的黑點，記為 A_0 點，而其右、右上、上、左上四臨點必含有一個邊界點，記為 A_1 。再從 A_1 開始找起，按右、右上、左上、左、左下、下、右下的順序找相鄰的邊界點 A_2 ，若 A_2 就是 A_0 ，則表明已經旋轉一圈，程式結束，否則就繼續尋找下個相鄰邊界點，直到再次找到 A_0 為止。但此方法運算量較大，可藉由稍稍的改變跟蹤準則使整個過程簡化。

首先跟上述方法相同的方法找到第一個起始點 A_0 ，定義初始的搜索方向為沿左上方，如果左上方為黑點則為邊界點，否則搜索方向順時針旋轉45度，直到找到黑點，然後把此點定為 A_1 ，並把當前搜索方向逆時針旋轉90度，繼續尋找下一個黑點，直到返回第一個初始點為止，則程式結束，如下圖。

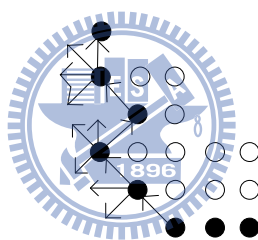


圖 2.1 輪廓追蹤圖。

2.2.2 傅立葉描述子(Fourier descriptors)

傅立葉描述子是一種以輪廓特徵為基礎的特徵抽取工具，在 2D 圖像檢索中已經是很有效的一種方法，其可經由正規化(normalize)使其對平移、旋轉，和比例大小具有不變性。而基於不同方法來取特徵，其正規化的方法也有些許不同，而辨識效果也有顯著的差別，其中依照其採取特徵的不同可分為中心距離傅立葉(radius Fourier)，面積傅立葉(area Fourier)，角度傅立葉(angular Fourier)，複數平面傅立葉(contour complex Fourier)，極座標傅立葉(polar Fourier)和綜和角度跟中心距離傅立葉(angular radius Fourier)等，從 I. Kunttu 和 L. Lepiston[12]所提出的論文，可知綜和角度跟中心距離傅立葉的檢索效果為最佳，其次是中心距離傅立葉，而且這兩種方法明顯勝過其他方法，但

在本論文中比較中心法搭配 Zernike 矩跟中心角度法搭配 Zernike 矩的效能，其中以中心法搭配 Zernike 矩的效能較優，且不必使用更多的參數作最佳化。

1. 中心距離傅立葉：

在平面上輪廓可以以一維的方程式來表示，使傅立葉方程式只需要處理實數部分，可以簡化計算量，其中中心法就是以物體輪廓上每一點到物體中心的距離為特徵方程式，定義如下：

$$r(k) = \sqrt{(x_k - x_c)^2 + (y_k - y_c)^2}, \quad k = 1, \dots, N, \quad (2-23)$$

其中 (x_k, y_k) 為物體上輪廓的一點， (x_c, y_c) 為物體的中心點， N 為邊界點的總數。

2. 面積傅立葉：

這方法主要是以邊界上兩點跟中心所夾的三角形面積作為特徵方程式，其也是一維方程式，故做傅立葉轉換時也只需做實數部分，定義如下：

$$a(k) = \frac{|(x_k - x_c)(y_{k+1} - y_c) - (x_{k+1} - x_c)(y_k - y_c)|}{2}. \quad (2-24)$$

3. 角度傅立葉：

這方法是以兩個邊界點與水平軸的夾角，其也是一維方程式，其中 w 為一固定常數，可因所取的輪廓點總數不同而有所不同，但從論文[13]和[14]可知這種取特徵的方式跟其他方法比起來檢索效能非常低。其定義如下：

$$\varphi(k) = \tan^{-1}\left(\frac{y_k - y_{k-w}}{x_k - x_{k-w}}\right). \quad (2-25)$$

4. 複數平面傅立葉：

這方法主要是把原本座標的 x 當做傅立葉方程式中的實數部份，而 y 當成虛數部，使原本圖型投影到複數平面，形成一維的複數方程式，定義如下：

$$z(k) = (x_k - x_c) + j(y_k - y_c). \quad (2-26)$$

5. 極座標傅立葉：

這方法則是把平面座標系改成極座標系，使距離為傅立葉方程式的實數部，角度則為虛數部，其中距離也就是式(2-23)的 $r(k)$ ，定義如下：

$$p(k) = r(k) + j\theta(k). \quad (2-27)$$

6. 綜和角度跟中心距離傅立葉：

這方法主要是結合第一跟第二種方法，其原因是因為中心距離法並無法考慮空間中個點的關係，而由角度法作為彌補，形成一個一維複數傅立葉方程式，定義如下：

$$\Gamma(k) = r(k) + j\varphi(k) \quad (2-28)$$

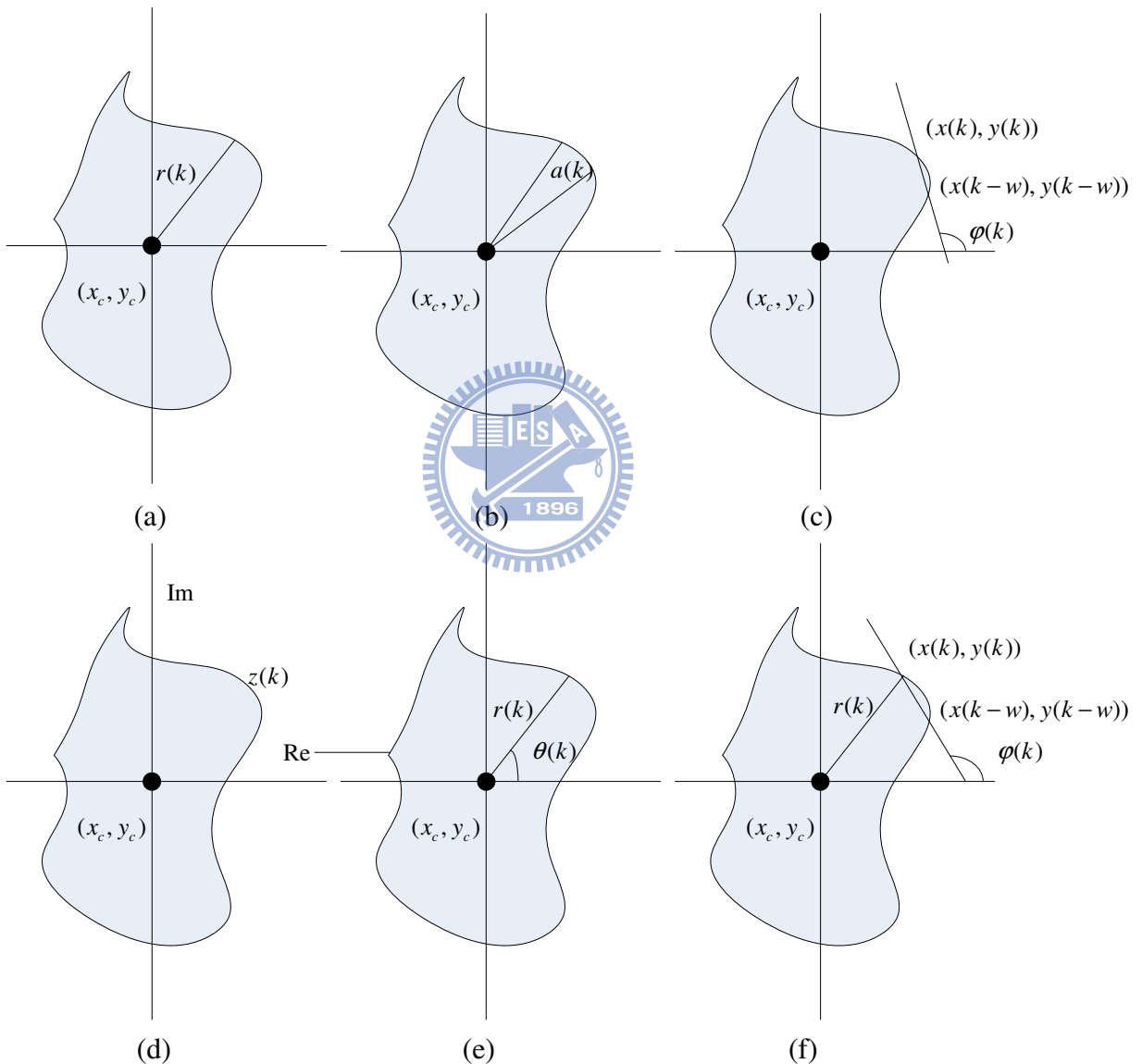


圖 2.2 不同傅立葉特徵，(a)中心傅立葉、(b) 面積傅立葉、(c) 角度傅立葉、(d) 複數平面傅立葉、(e) 極座標傅立葉、(f) 綜和角度跟中心距離傅立葉。

選定好抽取特徵的方法後，對參數作傅立葉轉換，但因為上述所說之特徵參數均為離散訊號，故必須使用離散傅立葉轉換，其定義如下：

$$F(n) = \frac{1}{N} \sum_{k=0}^{N-1} f(k) \exp\left(\frac{-2\pi nk}{N}\right), \quad n=0, 1, \dots, N-1, \quad (2-29)$$

其中 $F(n)$ 也就是 $f(k)$ 的傅立葉轉換後的值。其中為了使每項係數都能具有旋轉不變的性質，故只取 $|F(n)|$ 的大小，而不考慮正負。其次為了使傅立葉描述子具有比例大小不變性，要針對複數特徵時把每項 $F(n)$ 除以 $F(1)$ ，定義如下：

$$\left[\frac{|F_{-(L/2-1)}|}{|F_1|}, \dots, \frac{|F_{-1}|}{|F_1|}, \frac{|F_2|}{|F_1|}, \dots, \frac{|F_{L/2}|}{|F_1|} \right]^T, \quad (2-30)$$

而針對特徵是實數時，每項 $F(n)$ 除以 $F(0)$ ，定義如下：

$$\left[\frac{|F_1|}{|F_0|}, \frac{|F_2|}{|F_0|}, \dots, \frac{|F_L|}{|F_0|} \right]^T, \quad (2-31)$$

其中 L 為一固定值，越大代表特徵的維度越高，且不可超過輪廓總數 $\frac{N}{2}$ ，超過的特徵值為重複的特徵值並不需要。



2.2.3 Zernike 矩

Zernike 矩是一個以區域特徵為基礎的圖像檢索工具，且在 D. Zhang 和 G. Lu[17]所提之論文中證明其在平面空間中的檢索效果比傅立葉還好，主要就是因為其能考慮更多細節，並且因為其為一正交複數矩，藉由正交多項式處理在一單位圓內的形狀特徵。故可以保證其所提供之特徵相關性小，重複度低。並且 Zernike 矩對於平移和旋轉也具有不變性而再經過正規化後也可具有比例大小的不變性，故其為一良好的圖像搜索工具。其低階矩是描述物體整體的形狀特徵，高階矩為描述物體細微的變化，故適當選擇使用其低階矩能得到很好的檢索成果。Zernike 矩的定義如下：

$$Z_{mm} = \frac{m+1}{\pi} \iint_{x,y} f(x,y) [V_{mm}(x,y)]^* dx dy, \quad \text{where } x^2 + y^2 \leq 1, \quad (2-32)$$

其中 m 為 Zernike 矩的階數， $m=0, 1, \dots, \infty$ ， $f(x, y)$ 為圖像亮度， $*$ 代表共軛複數，當 n 為正整數，其為旋轉角度依存，且要滿足 $(m-|n|)=\text{even}$ ， $n \leq |m|$ 。之後把 Zernike 矩轉換成極座標表示：

$$Z_{mn} = \frac{m+1}{\pi} \int_0^1 \int_0^{2\pi} f(r, \theta) V_{mn}(r, \theta) dr d\theta, \quad \text{where } r \leq 1, \quad (2-33)$$

其中 $r = \sqrt{x^2 + y^2}$ ， $\theta = \tan^{-1}(y/x)$ 。而為了方便計算可再把 Zernike 多項式改為：

$$V_{mn}(r, \theta) = R_{mn}(r) \exp(jn\theta), \quad (2-34)$$

其中 $j = \sqrt{-1}$ ， (r, θ) 的範圍定義在單位圓上， $R_{mn}(r)$ 為正交徑項多項式，其定義如下：

$$R_{mn}(r) = \sum_{s=0}^{\frac{m-|n|}{2}} (-1)^s \frac{(m-s)!}{s! \left(\frac{m+|n|}{2} - s\right)! \left(\frac{m-|n|}{2} - s\right)!} r^{m-2s}, \quad (2-35)$$

而對於 $f(x, y)$ 只為純實數且為離散時，可將式(2-32)改為：

$$Z_{mn} = \frac{m+1}{\pi} \sum_x \sum_y f(x, y) [VR_{mn}(x, y) + jVI_{mn}(x, y)], \quad \text{where } x^2 + y^2 \leq 1, \quad (2-36)$$

其中 $VR_{mn}(x, y)$ 為 $V_{mn}(r, \theta)$ 的實數部分， $VI_{mn}(x, y)$ 為 $V_{mn}(r, \theta)$ 的虛數部分。

最後為了使 Zernike 矩具有旋轉不變性，故對 Z_{mn} 只取其大小 $|Z_{mn}|$ 不考慮正負，而為了使 Zernike 矩具有比例大小不變性，再對其每一項 Z_{mn} 除以 Z_{00} 作比例大小正規化，但不包含 Z_{00} 本身，而 Z_{00} 也就是其在單位圓所佔的面積，作法如下：

$$\left[\begin{array}{c} |Z_{10}|, |Z_{20}|, |Z_{22}|, \dots, |Z_{\infty\infty}| \\ |Z_{00}|, |Z_{00}|, |Z_{00}|, \dots, |Z_{00}| \end{array} \right]^T, \quad (2-37)$$

整個 Zernike 矩的特徵向量抽取流程如下：

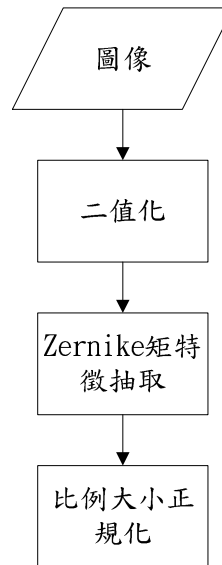


圖 2.3 抽取 Zernike 矩特徵向量流程圖。

首先針對每一張圖像必須先作二值化，使圖形變成只有黑與白的圖像，其次藉由 (2-36) 式，計算每個 Z_{mn} ，之後再對每項 Z_{mn} 做比例大小正規化。當 $m=4$ 時，形成特徵向量集合 $\{Z_{11}, Z_{20}, Z_{22}, Z_{31}, Z_{33}, Z_{40}, Z_{42}, Z_{44}\}$ ，階數越大特徵值越多。但是代表性也將越低，因為越高階屬於越小的變動，容易因雜訊或小部分的不同而有很大的差異，故作檢索時階數要取得適當，而不要過多。

2.3 相似度分析

要如何分辨兩物體的相似度分析，通常主要使用的是兩物體的距離，也就是差異度而非是相似度。當距離越小，相異度也就越小，相似度也就越高。J. H. Tangelder 等人[13] 在論文中提出一個好的距離 (d) 評判標準在相似模型集合 S 中應該要符合以下幾個定義：

1. 一致性：對於所有 $x \in S$, $d(x, x) = 0$ 。
2. 正定性：對於所有 $x \neq y$ in S , $d(x, y) > 0$ 。
3. 對稱性：對於所有 $x \neq y$ in S , $d(x, y) = d(y, x)$ 。

4. 三角不定性：對於所有 $x, y, z \in S$, $d(x, z) \leq d(x, y) + d(y, z)$ 。
5. 轉移不變性：對於一個座標轉換矩陣 G ，對於所有 $x, y \in S$ ，
 $d(x, y) = d(G(x), G(y))$ 。

第一種性質是指對於同物體所抽取出的特徵距離為零，也就是跟自己要完全相似，而第二種性質是不同的物體其一定有所不相似，這是高階的描述子應該要有的強健性，但大部分的描述子皆沒有此項性質，但缺乏此性質也不算是嚴重的缺點，因為可能只是省略了部分不需要的小區塊而造成此問題。

第三種特性也並不是必要的，畢竟就算從人眼中觀看，這項性質也不一定成立，很有可能只會覺得猴子像人，但人不像猴子，兩邊的差距並不對稱。而三角不定性通常在區域形辨識系統上並不遵守，常會有區域相似但整體不相似。

轉移不變性則是是模型檢索一定要具備的條件，其就是前面所講的平移不變性，旋轉不變性，跟比例不變性的綜合體。



最常使用的距離判斷是範數距離(norm)：

$$\|d_{n-norm}\| = (\sum |x_i - y_i|^n)^{1/n}, \quad (2-38)$$

當 $n=2$ 時，就是歐基里德距離，是常用的距離標準。不過在本論文的演算法其距離都以 1-norm 為標準，因實驗結果其檢索效能比使用 2-norm 來的較佳。

2.4 檢索評判標準

當檢索結果出來時，必須判斷其檢索成果的好壞，以知道其整個系統的效能，以下列出，幾種常用方法：

1. 精準度對回傳率圖(precision-recall plot)：

將回傳物體對正確率作一條相關性曲線，對一個分類，其內含 C 個物體，從其

中一個物體尋找最相關的 K 個物體，準確度則為這 K 個最相似物體的正確度， K 必須小於 C 。故當一個物體回傳越多，其準確度通常會越低，而一個完美的系統是不管回傳多少準確度永遠為一。而其中精確度(precision, p)跟回傳率(recall, r)定義如下：

$$p = \frac{TP}{TP+FP}, \quad (2-39)$$

$$r = \frac{TP}{TP+FN}, \quad (2-40)$$

其中 TP 為想要並且檢索到的資料(true positive)，TN 為不想要也沒檢索到的資料(true negative)，FN 為想要但沒檢索到的資料(false negative)，FP 為不想要但檢索到的資料(false positive)，如下圖所示：

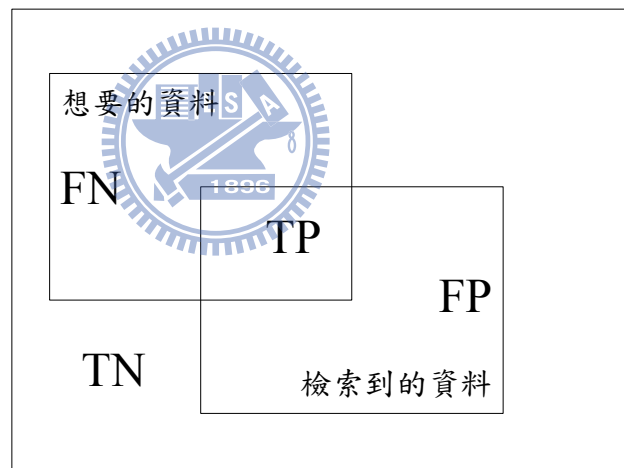


圖 2.4 精確度對回傳率代表圖。

2. 最近相鄰者(nearest neighbor, NN)：

對一物體檢索到最接近的物體為相同分類的正確度，理想值為 100%，其值越高代表系統效果越好。

3. first-tier 和 second-tier：

檢索結果對同一類型內含 C 個物體的分群，其中最相似的 K 個物體的正確率，對 First-tier 而言， $K = |C| - 1$ ，而對 Second-tier 而言， $K = 2 * (|C| - 1)$ 。理想值為 100%，分數越高代表效果越好。

4. E-measure :

一個組合正確與不正確機率的修正參數，通常用來表達在第一頁找到相關物體的準確度，故只計算檢索到的前 32 個物體。其中 p 跟 r 定義同式(2-38)跟(2-39)，E-measure 定義如下：

$$E = \frac{2}{\frac{1}{p} + \frac{1}{r}}. \quad (2-41)$$

5. discounted cumulative gain (DCG) :

統計關於正確結果越早出現分數越高的統計量，是對搜尋引擎常用的一種標準。當一個物體被檢索到時其為相關物體則 G_i 為 1，否則為 0，其定義如下，

$$DCG_i = \begin{cases} G_1, & i = 1, \\ DCG_{i-1} + \frac{G_i}{\lg_2(i)}, & \text{otherwise,} \end{cases} \quad (2-42)$$

當計算完畢後除以最大可能值如下：

$$DCG = \frac{DCG_k}{1 + \sum_{j=2}^k \frac{1}{\lg_2(j)}}, \quad (2-42)$$

其中 k 為資料庫中模型的總數。DCG 值理想結果為一，同樣當分數越高代表系統越好。

第三章 檢索系統

本章提出一個基於主軸分析與區域特徵抽取的檢索系統，其擁有位移、旋轉以及比例大小不變性的特性。檢索系統的架構圖，如圖 3.1 所示。首先對每一個 3D 模型作前處理，定義整體解析度為 $(2R, 2R, 2R)$ ，使每一個物體的中心都在座標 (R, R, R) 上，以此克服位移變量，並把最遠點到中心的距離轉換為解析度的一半 R ，使相似物體具有一固定的比例大小，使其有比例大小的不變量，之後藉由主成分分析定出主軸旋轉物體並輸出成格子點，使相似物體具有相同的座標系，以克服旋轉變量，但由之後實驗可知只做主成分分析法並不能完全克服旋轉變量，故之後所使用的特徵抽取也需具備旋轉不變量，並在最後相似度比對時必須旋轉物體找最佳解，才能完整克服旋轉變量。做完前處理後，抽取三種不同特徵：傅立葉描述子、Zernike 矩跟三軸投影圖形的相關大小，其中圖像相關大小特徵隱含在另兩種特徵之內。且在計算傅立葉描述子之前，必須先找出輪廓的每點位置當作傅立葉方程式的值，並且在計算傅立葉描述子特徵時，不做最後的比例大小正規化這一步驟，而 Zernike 矩特徵同樣也不作比例大小正規化這步驟，如此圖像大小相關性則存在在這兩項特徵之內，最後把此兩種特徵向量權重相加，得到 3D 物體的相似度，把其相似度依序排列也就是檢索結果。本方法跟 D. Y. Chen 提出的論文 (LFD)[7]都均為以 2D 圖形為檢索準則的演算法，但克服了 LFD 所需要的龐大運算量跟儲存空間，其原因就是以主成分分析法等相關步驟克服旋轉變量，並藉由各個圖像大小相關性取代多面特徵比較，使運算簡化並且對於相似物體擁有更好的容忍度，使檢索效果更佳。以下第一節說明離線架構的步驟，第二節說明在線架構的部份。

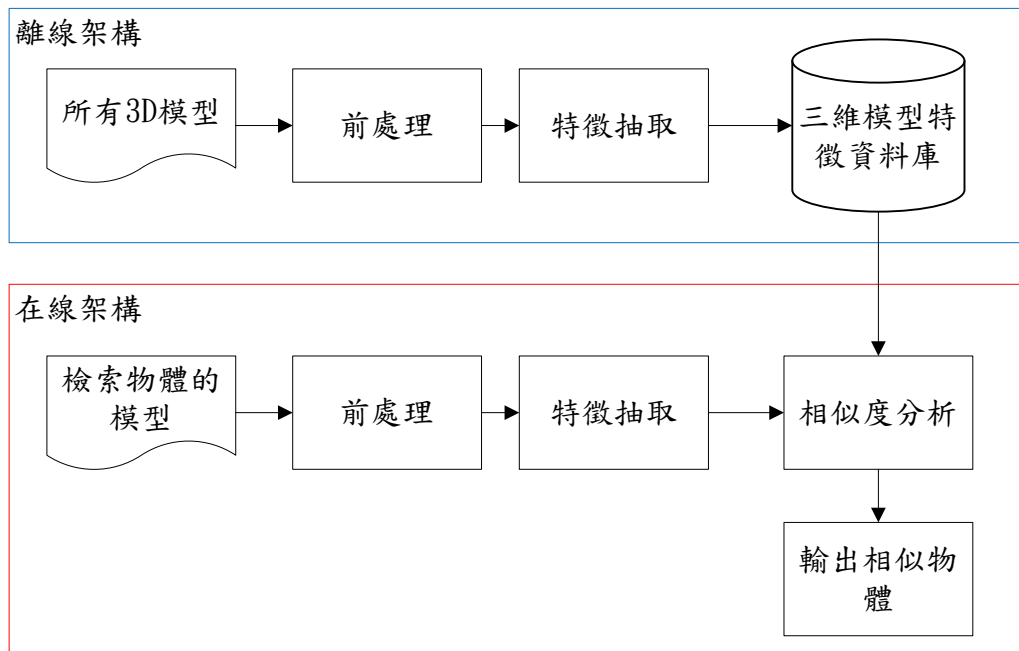


圖 3.1 檢索系統架構。

3.1 離線架構

本節將介紹檢索系統的離線架構，如圖 3.2 所示，此階段是在實際檢索物體前必須先作的準備，首先要針對所有要處理的 3D 模型資料做前處理，平移、旋轉跟比例大小的變換，使相似物體有相同中心位置，姿態跟比例大小並輸出成格子點，使原本點數過多的模型減少太過於細節的部份和補足原本點數過少的模型的點跟點間太過空洞的部份，然後投影輸出 2D 圖型並抽取傅立葉描述子的特徵向量跟 Zernike 矩的特徵向量，最後輸出特徵向量形成 3D 模型的特徵資料庫，之後在使用此系統時則只需在特徵資料庫中尋找相關特徵，則能達到檢索相關物體的效果。

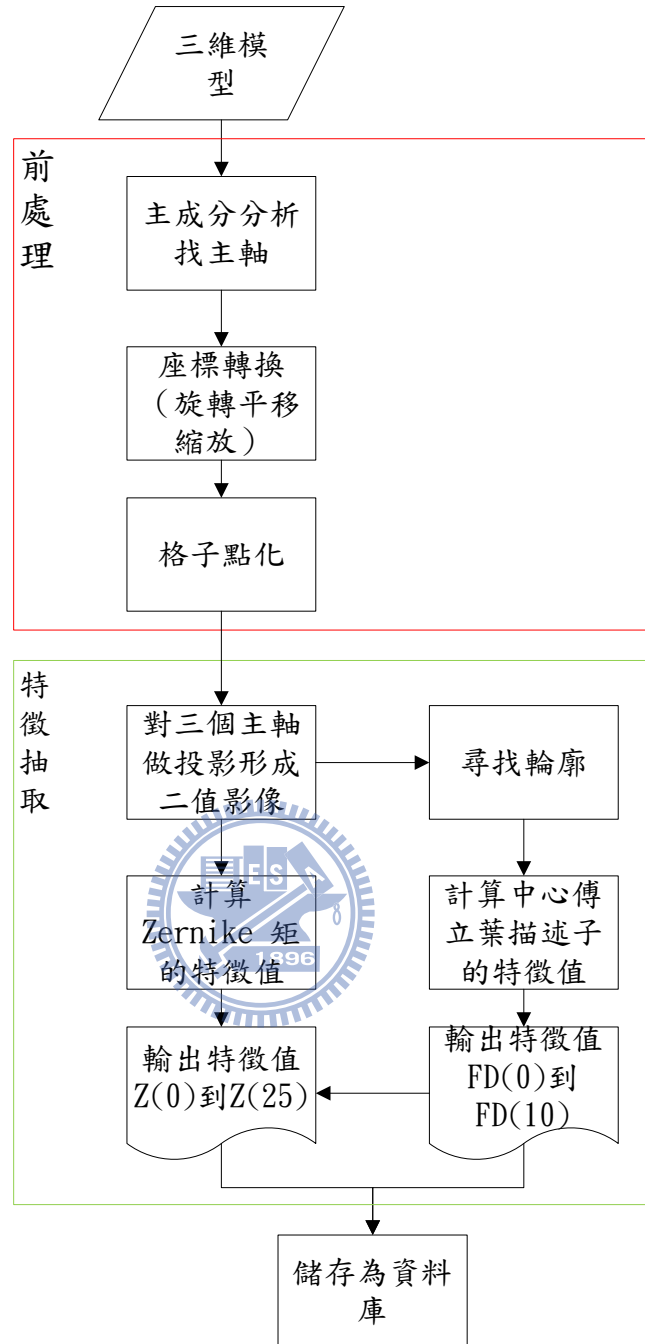


圖 3.2 離線架構圖。

3.1.1 前處理

一個好個檢索系統，對於每一個物體都必須要有適當的前處理，使物體具有位移，旋轉和比例大小的不變性，如此就算輸入的物體初始式不同角度位置，或比例大小，也能有相同的檢索效果，其流程圖如下。

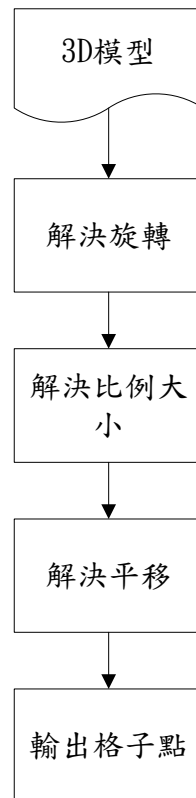


圖 3.3 前處理流程圖。

前處理的第一步是克服旋轉變量，本篇論文是以主成分分析法找出整個物體的三大主軸，並轉換座標系使相似或相同的物體都會有相同的座標軸，如圖 3.4 所示，則之後投影出來的圖型也能具備相似的形狀，在最後也才能有正確的相似度判斷，不過因為主成分分析法還是可能會因為物體部分的變動，使三大主軸顛倒或互換，而最糟的結果是完全沒對齊，但從之後實驗結果可知，主軸完全沒對齊的結果非常少，主要會發生的問題是主軸的顛倒或互換，故在此階段並沒有完全克服旋轉變量，必須在之後搭配傅立葉描述子跟 Zernike 矩並利用其所具備的旋轉不變量加以克服主軸的偏移，並在最後的相似度比對時做三大投影圖旋轉找最佳組合的步驟。

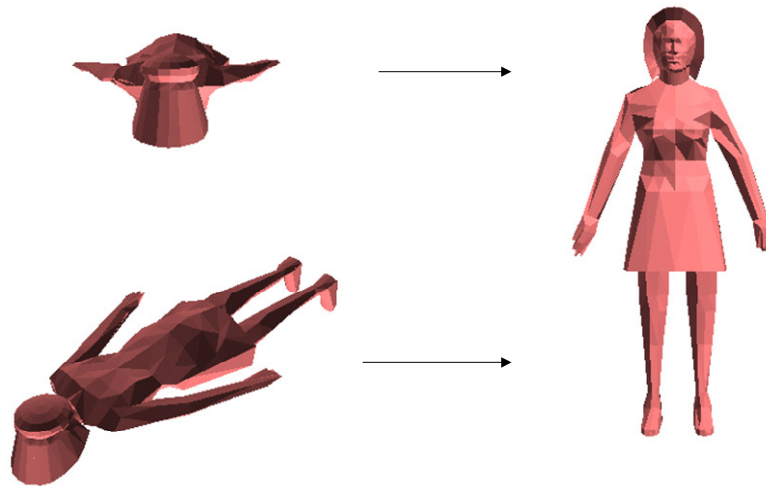
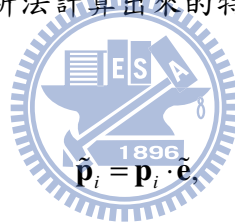


圖 3.4 3D 模型旋轉圖。

為了準確計算主成分分析的主軸，本篇論文使用 D.V. Vranić 等人[19]所提出的連續主成分分析法(CPCA)找出模型的前三大主軸，並開始對原始資料作主軸旋轉。即每個點的 $\mathbf{p}_i = (x_i, y_i, z_i)$ 乘上主成分分析法計算出來的特徵向量 $\tilde{\mathbf{e}} = [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3]$ ，形成新的點為 $\tilde{\mathbf{p}}_i = (\tilde{x}_i, \tilde{y}_i, \tilde{z}_i)$ ，作法如下：



$$\tilde{\mathbf{p}}_i = \mathbf{p}_i \cdot \tilde{\mathbf{e}} \quad (3-1)$$

其次要對模型作比例(S)的固定，使相似模型具有相似的比例大小，且能使之後的特徵向量大小能具有比例相關性，形成一個新的特徵。故以模型上的點到模型中心點 $\mathbf{c} = (x_c, y_c, z_c)$ 的最遠距離(L_{\max})的兩倍作為新格子點的總長度，然後對每點作比例縮放，並且把物體中心移到格子點的中心處(R, R, R)，作法如下：

$$S = R / 2L_{\max}, \quad (3-2)$$

$$\tilde{\tilde{\mathbf{p}}}_i = (\tilde{\mathbf{p}}_i - \mathbf{c}) \cdot S, \quad (3-3)$$

最後把資料庫的物體從點跟面資料轉換成格子點(grid voxel)型態，使原本點數太少的模型之後的投影圖不會產生內部空洞。

對於一個有三個點 A, B, C 的面，其相對位置為 a, b, c 轉換為格子點，必須依

照以下步驟：

```
Step1  voxel(floor(i)) = 1    , i = a, b, c ,
Step2  let ai = a, bi = b, r = 1,
Step3  when  $\overline{a_i b_i} > r$ 
        ai = ai +  $\overline{a_i b_i} \times r$ ,
        voxel(floor(ai)) = 1,
        Until  $\overline{a_i b_i} \leq r$ .
Step4  if  $\overline{ac} > 1 \times r$  &  $\overline{bc} > 1 \times r$ ,
        a = a +  $\overline{ac} \times r$ , b = b +  $\overline{bc} \times r$ ,
        Return Step2
else end.
```

圖 3.5 格子點轉換假碼。

其中 r 為每次檢索的範圍，當 r 越小則所轉換的三角形邊界越準確，但也越耗時間，在本篇論文中 r 定為 1。

完成上述的座標轉換跟格子點化後即完成前處理的動作，接下來則要對格子點作主軸投影並作特徵抽取。

3.1.2 特徵抽取

首先為了要抽取傅立葉描述子跟 Zernike 矩，必須把 3D 模型先轉為 2D 圖型，故先對 X ， Y ， Z 三主軸作垂直投影，形成三張二值化影像，如圖 3.5 所示，在實驗中解析度定為 256×256 像素，更高的解析度會增加整體的運算複雜且會使特徵太專注於細節，但過低的解析度也會造成 Zernike 矩的檢索效果嚴重下降。

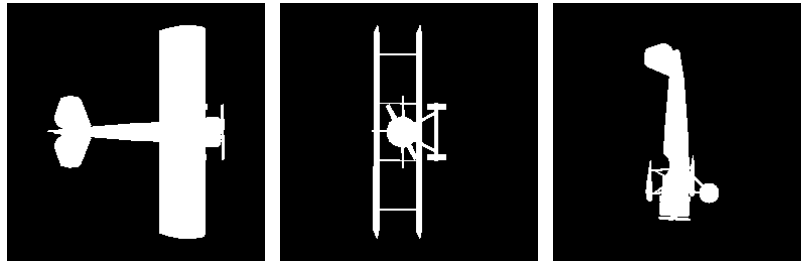


圖 3.6 三主軸投影圖。

其次對兩種特徵抽取方法將分成兩部分：首先對傅立葉描述子作特徵抽取前，必須要先找尋三張影像的輪廓，如圖 3.6 所示：

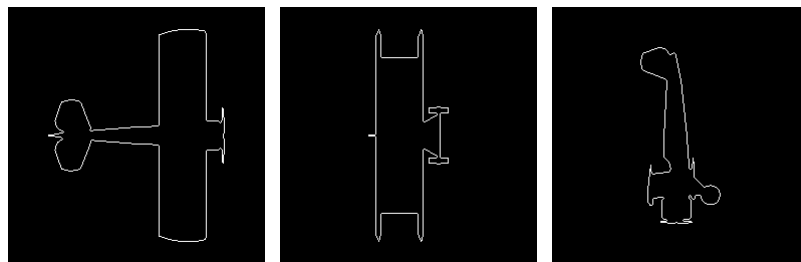


圖 3.7 三主軸輪廓圖。

之後再以輪廓上每點到中心點的距離為傅立葉參數，經過傅立葉轉換，計算出傅立葉描述子的前 11 個特徵值 $F(0)$ 到 $F(10)$ ，這包含了平常用來作為大小正規化的 $F(0)$ ，其代表意義是本張圖型所有輪廓點到中心的平均距離。

而 Zernike 矩則是直接針對三張 2 值化的影像作處理，在本篇論文中選定階數(m)為 8，共 25 個參數，其每階所含的矩個數如表 3.1 所示，更高階的矩經由實驗知道其所含有的資訊已經是太過細節的部份，對於檢索系統的效能並沒有幫助，只會增加更多的運算故不使用。

表 3.1 Zernike 矩每階所含的矩。

m	Moments	No. of Moments
0	Z_{00}	1
1	Z_{11}	1
2	Z_{20}, Z_{22}	2
3	Z_{31}, Z_{33}	2
4	Z_{40}, Z_{42}, Z_{44}	3
5	Z_{51}, Z_{53}, Z_{55}	3
6	$Z_{60}, Z_{62}, Z_{64}, Z_{66}$	4
7	$Z_{71}, Z_{73}, Z_{75}, Z_{77}$	4
8	$Z_{80}, Z_{82}, Z_{84}, Z_{86}, Z_{88}$	5

其次在對每張投影圖形計算 Zernike 矩前，必須先計算好每階所會形成的不同亮度的單位圓圖形，其超過單位圓的部份階設為 0，如圖 3.7 所示。最後投影圖以一對一的方法對每張單位圓圖形相乘能快速得到 Zernike 矩的特徵向量 Z_{00} 到 Z_{88} ，其中也包含了原本作為比例正規化參數的 Z_{00} ，最後在把每個模型依序儲存為特徵向量，形成特徵資料庫。

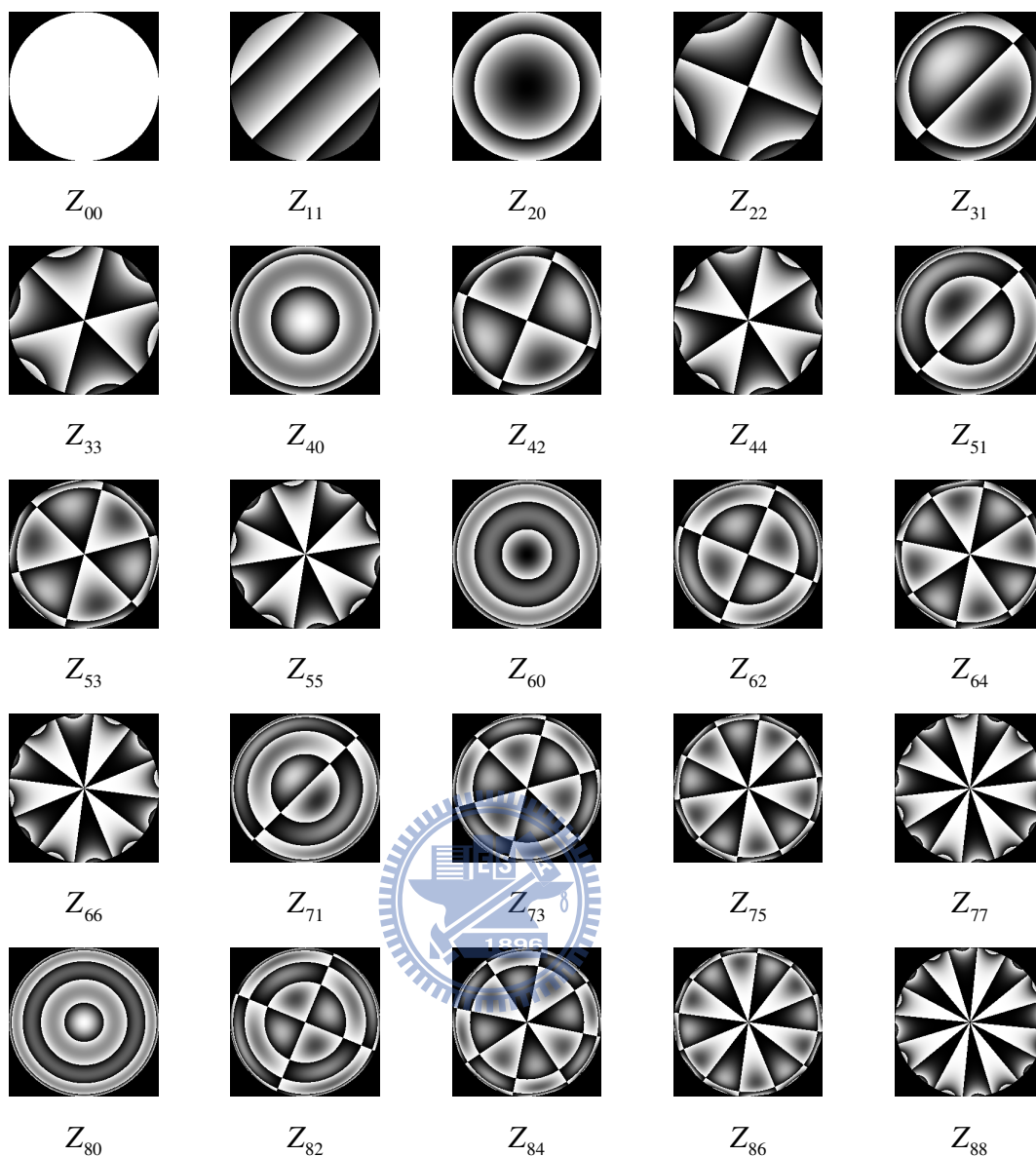


圖 3.8 Zernike 矩每階所形成的不同單位圓。

3.2 系統在線架構

在線架構部分是提供給使用者能快速檢索相關物體的部份，其架構圖如圖 3.8，首先要檢索的物體作前處理跟特徵抽取，再把特徵向量跟資料庫中所有 3D 模型的特徵向量作相似度比較，故如果特徵向量儲存越多，將會使檢索速度越慢，故特徵向量越少但具有足夠代表性會越好。而在線階段的前處理跟特徵處理階段跟離線階段相同，故本節只再介紹相似度分析。

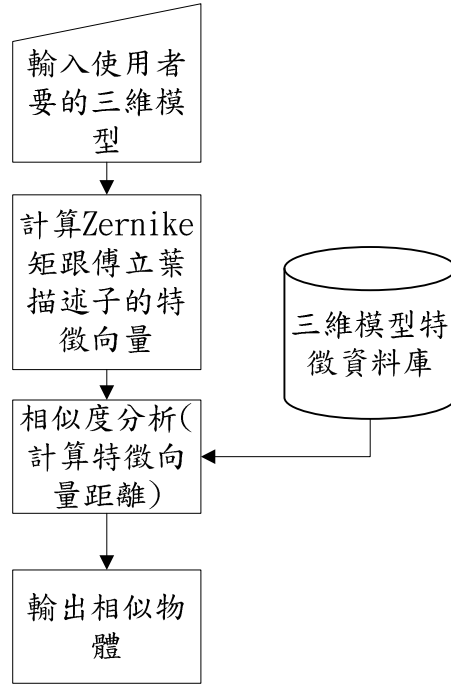


圖 3.9 系統在線架構。

3.2.1 相似度分析

當檢索相似物體時，首先也要對檢索物體作前處理之後再求取檢索物體的 Zernike 矩和傅立葉描述子的特徵向量。最後再針對所求得的特徵向量跟特徵資料庫作相似度比較。首先 Zernike 矩對於一個影像，所求得的特徵向量會有兩組，正面 $Z_1(i)$, $i = 0, \dots, 24$ 和反面 $Z_2(i)$, $i = 0, \dots, 24$ ，而為了處理雖然主成分分析法轉成相似型態，但卻沒有完全克服旋轉變量的物體，其還是會因為不同物體雖然整體相似但因為部分的變化，發生正方向顛倒或主軸顛倒的情形，形成沒有對齊的狀況，故在此最後比對階段時要以一對一面的特徵向量距離和為最小的解為其最後的比對依據，如下式。

$$d_{iz} = \min_k (\min_k (\text{dist}(\tilde{Z}_{q,1,1}, \tilde{Z}_{i,a,k})) + \min_k (\text{dist}(\tilde{Z}_{q,2,1}, \tilde{Z}_{i,b,k})) + \min_k (\text{dist}(\tilde{Z}_{q,3,1}, \tilde{Z}_{i,c,k}))),$$

$$\text{for } \begin{cases} a = 1, 2, 3 \\ b = 1, 2, 3 \wedge b \neq a \\ c = 1, 2, 3 \wedge c \neq a \wedge c \neq b \end{cases}, \quad (3-4)$$

其中 Z 的下標第一項 q 是指檢索模型， i 為資料庫內的任意模型，下標第二項為第幾個投影面，下標第三項為正反面，1 為正面，2 為反面。

而傅立葉描述子針對一個影像，只需求得一組特徵向量 $F(i)$, $i=0, \dots, 10$ ，因為正反面的中心距傅立葉描述子的特徵向量幾乎相等，故省去反面的特徵向量只取正面所得到的特徵向量，其次和 Zernike 矩一樣，為了完全克服旋轉變量，必須取一對一面的最小和，求特徵距離如下式：

$$d_{if} = \min(\text{dist}(\tilde{F}_{q,1}, \tilde{F}_{i,a}) + \text{dist}(\tilde{F}_{q,2}, \tilde{F}_{i,b}) + \text{dist}(\tilde{F}_{q,3}, \tilde{F}_{i,c})),$$

$$\text{for } \begin{cases} a=1,2,3 \\ b=1,2,3 \wedge b \neq a \\ c=1,2,3 \wedge c \neq a \wedge c \neq b \end{cases}, \quad (3-5)$$

其中以 1-norm 當作兩個特徵向量的距離，而又因為使用的 Zernike 矩跟傅立葉描述子是並未大小正規化的特徵向量，所求得的距離會差異很大，為了使兩參數較好結合，再把這兩特徵向量的參數皆除上兩特徵向量的最遠距離，使傅立葉描述子跟 Zernike 矩所求得特徵向量的距離都在 0 到 1 之間，之後在特徵向量前加上權重值讓兩參數相加形成最後的兩物體的差異距離 d_i 。

$$\tilde{d}_{iz} = d_{iz} / \max_i(d_{iz}), \quad i \text{ for all database model}, \quad (3-6)$$

$$\tilde{d}_{if} = d_{if} / \max_i(d_{if}), \quad i \text{ for all database model}, \quad (3-7)$$

$$d_i = \tilde{d}_{iz} \cdot w + \tilde{d}_{if}, \quad (3-8)$$

其中 w 為權重值，針對於不同 database 將有不同的最佳解，預設為 1。

最後兩物體的相似度(SM)的定義如下式：

$$SM = \frac{(1+w)-d}{(1+w)}, \quad (3-9)$$

當其值越大代表兩物體越相似，當值為 1 時代表兩物體完全相同。

第四章 實驗結果與分析

在本章第一節將說明整個實驗平台，其次第二節是實驗主成分分析加旋轉變量修正的效果與不同方法下主成分分析的結果，之後第三節分析在 PSB 資料庫跟不同類別的小型資料庫中檢索物體的成效並比較本方法與其他方法的優劣，之後再比較不同方法的運算複雜度跟空間需求，第四節是針對不同資料庫所需特徵數加以探討，最後第五節是對整個實驗的結果作分析。

4.1 實驗平台

實驗的平台是在 Intel CPU 2.4Ghz 的電腦上，以 Visual Studio 2005 的環境中完成整個演算法。其中有使用到 Open CV 的程式資料庫跟 PSB[3]所提供的程式碼。本實驗中的所有 3D 物體均為 PSB 資料庫中的 3D 物體，其 3D 模型的點數從數十到上萬都有。

4.2 主成分分析法驗證



主成分分析法為一尋找最大分布的正交主軸的工具，本身已具備有平移，旋轉，比例大小的不變性，但在本節將先說明主成分分析本身的旋轉不變性對於在檢索物體時並不完善，需要加以修正才能完整克服旋轉變量，之後再以不同方法求取主成分分析，說明以不同的方法求取主成分分析將對整體檢索結果有很大的影響。

4.2.1 主成分分析法的旋轉變量修正

主成分分析法雖然本身即具有平移跟旋轉的不變性，但是這個特性只對於同一物體而具有不同初始角度時方有作用，即當同一物體輸入時就算是中心位置不同且座標軸也不同，但之後藉由乘上主成分分析法所求得的特徵向量還是能把物體主軸轉成一致的，如圖 4.1 的飛機，對於不同的起始輸入座標，但同樣都是輸出圖 4.2 這三張投影圖。

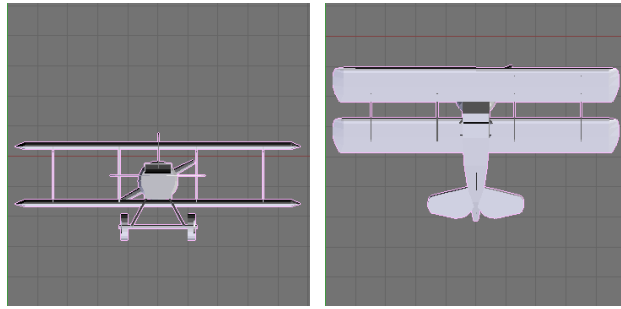


圖 4.1 編號第 1119 飛機，左圖為初始位置，右圖經過旋轉跟平移。

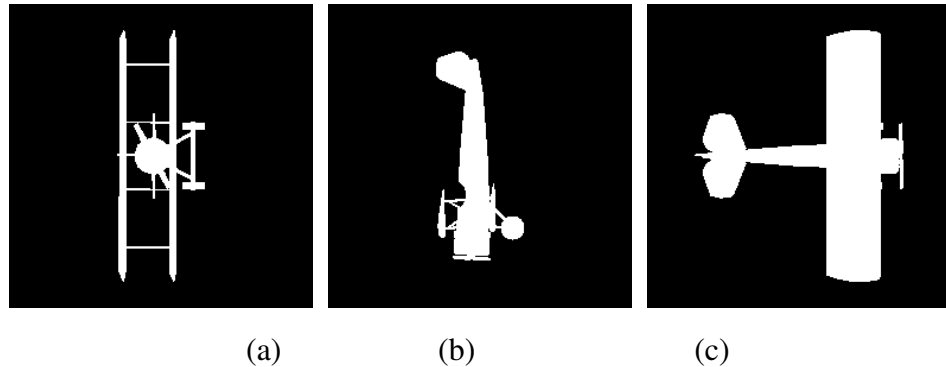


圖 4.2 編號第 1119 飛機的三軸投影圖：(a)對第一主軸投影、(b)對第二主軸投影、(c)對第三主軸投影。

但是對於相似物體，就可能會出現圖樣顛倒的情形，如圖 4.2 跟圖 4.3 為兩架不同的飛機，但都是屬於兩翼型飛機類型。卻明顯可看出每張圖均左右相反。其原因是因為當物體處於極為對稱的情況下，容易因為小部分的變化而造成主軸正向顛倒，而導致圖與圖形成左右顛倒的情況，所以之後必須配合具有旋轉不變性的特徵抽取。

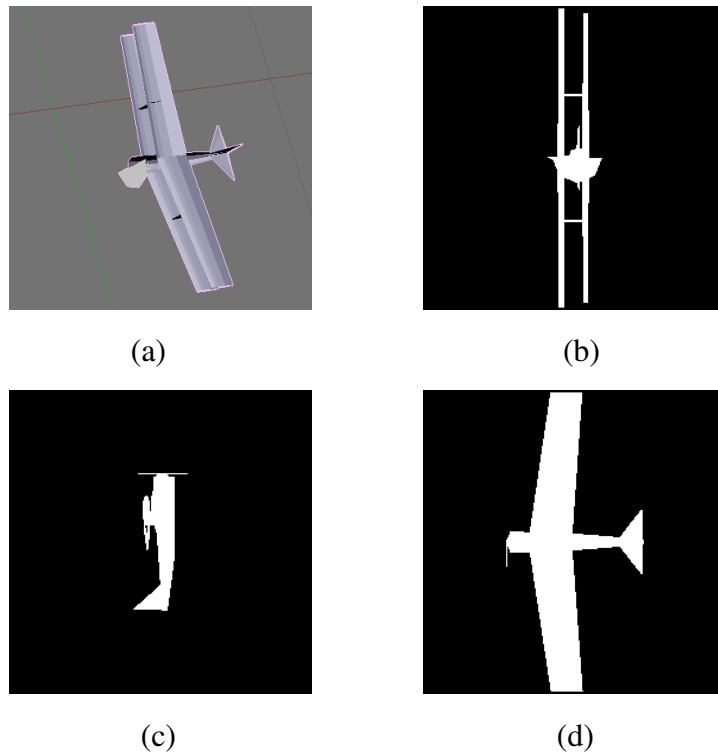


圖 4.3 編號第 1137 的飛機三軸投影圖：(a)原始未經旋轉的圖形(b)對第一主軸投影、(c)對第二主軸投影、(d)對第三主軸投影。

而除了顛倒以外更嚴重的情形是各主軸順序顛倒的情形，如下面兩架飛機的三主軸投影圖，可明顯由圖 4.4 和圖 4.5 可看到兩架飛機的投影圖順序不一致，圖 4.4 的(a)應該要對應到圖 4.5 的(b)，而圖 4.4 的(b)應該對應到圖 4.5 的(a)，且兩張圖的(c)旋轉了 90 度，若只使用主成分分析的旋轉不變性，而之後的特徵抽取沒考慮旋轉變量的話，檢索結果必然會出現錯誤。

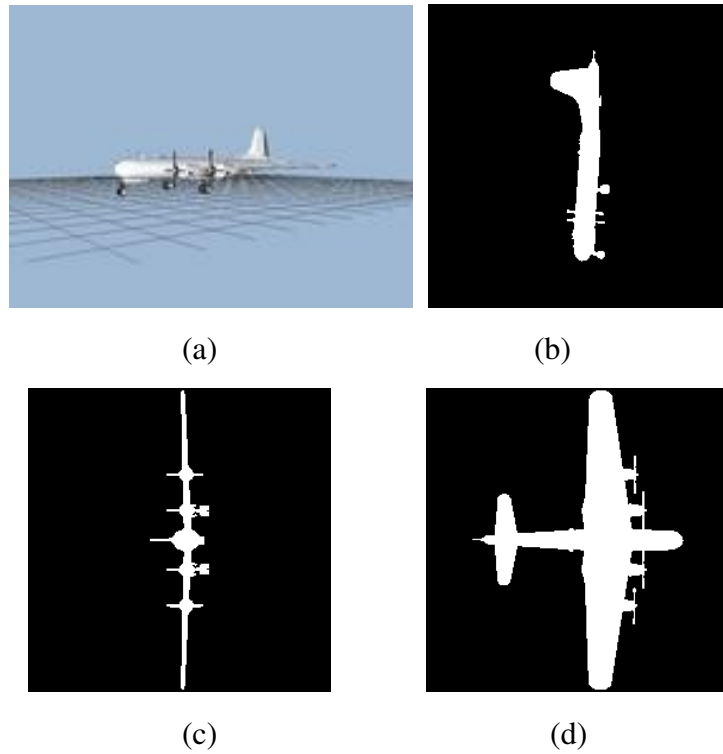


圖 4.4 編號第 1160 的飛機三軸投影圖：(a)原始未經旋轉的圖形(b)對第一主軸投影、(c)對第二主軸投影、(d)對第三主軸投影。

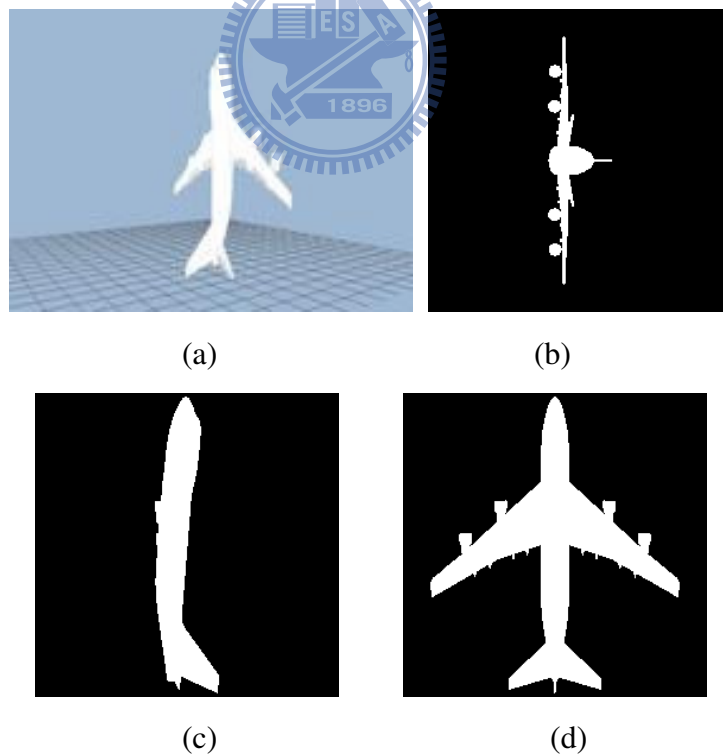


圖 4.5 編號 1162 的飛機三軸投影圖：(a)原始未經旋轉的圖形(b)對第一主軸投影、(c)對第二主軸投影、(d)對第三主軸投影。

從以上兩例可知只作主成分分析並不足以克服旋轉變量，必須在之後抽取特徵時也

要具備有旋轉不變性，而傅立葉描述子跟 Zernike 矩正都具備這一特性。而對於主軸順序顛倒，則必須要在相似度比對時，旋轉找尋最佳的對應方式，也就是旋轉主軸投影圖並找出最相似的一組解。而因為是三對三的判斷，故只需判斷六種組合，其和最小則為最相近，以此值定為這兩個物體的距離。而由以下實驗可看出，有做此修正能改善因主軸順序顛倒而造成錯誤的檢索，而提昇效能。

在此直接針對 PSB 資料庫作整體的檢索判斷，以下兩種方法只差異在有作旋轉修正跟不作旋轉修正直接作比對。由下表可知有作修正的方法可克服原本對於旋轉變量處理不佳的物體做修正，進而提昇檢索的正確率。

表 4.1 有無旋轉修正比較表。

方法	NN	first-tier	second-tier	E-measure	DCG
有做修正	0.648	0.387	0.494	0.287	0.643
直接比對	0.634	0.386	0.493	0.286	0.641

而在 T. Funkhouser 等人的論文[8]提出主成分分析的另一問題，對於部分物體具有相似型態，但因為部分區域的比例或形狀不同形成主軸偏移的情況。但在實驗中發現對於這種狀況，影響其實不大。在 PSB 資料庫中的基礎分類(base classification)中的測試(test)分群中的 907 個模型中，發生主軸偏移的現象發生在不同長度或形狀的椅背，其椅子的主軸發生偏移如圖 4.7 到圖 4.9，但因為傅立葉描述子跟 Zernike 矩都具備旋轉不變性，所以還是都能完整的檢索到正確的相似模型，如圖 4.9 對編號 816 的椅子在家具資料庫做檢索，其檢索結果為完全正確，而整體來看在家具資料庫實驗時發現，餐廳椅的檢索效果還是各個方法中最佳，如圖 4.16。

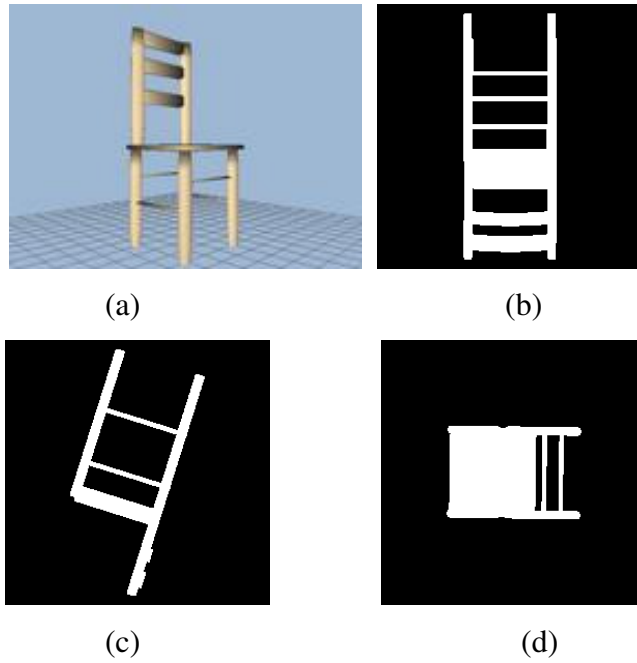


圖 4.6 編號 810 的椅子的三軸投影圖(椅背跟椅腳長度標準)：(a)原始未經旋轉的圖形(b)對第一主軸投影、(c)對第二主軸投影、(d)對第三主軸投影。

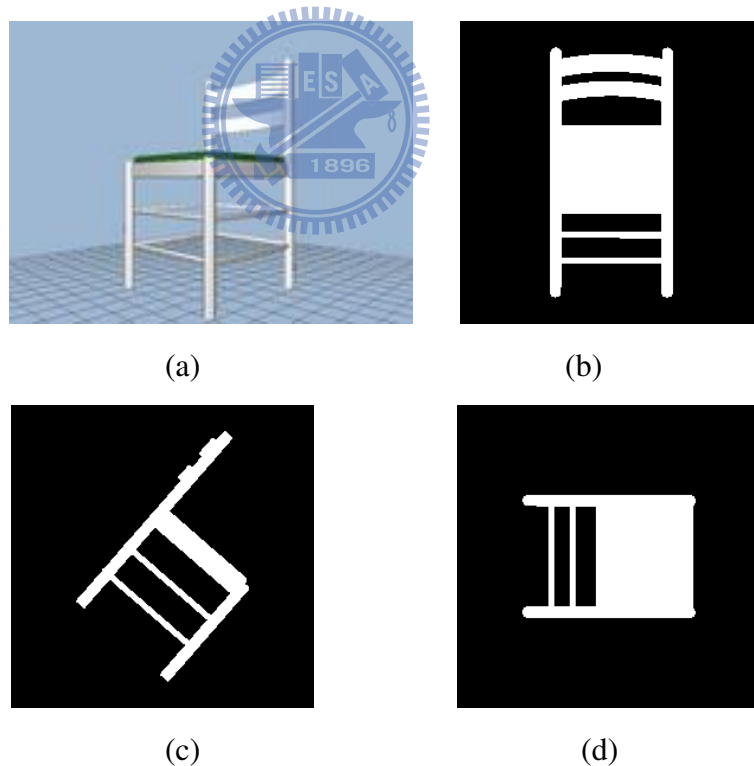


圖 4.7 編號 814 的椅子的三軸投影圖(椅背長度短，椅腳長度標準)：(a)原始未經旋轉的圖形(b)對第一主軸投影、(c)對第二主軸投影、(d)對第三主軸投影。

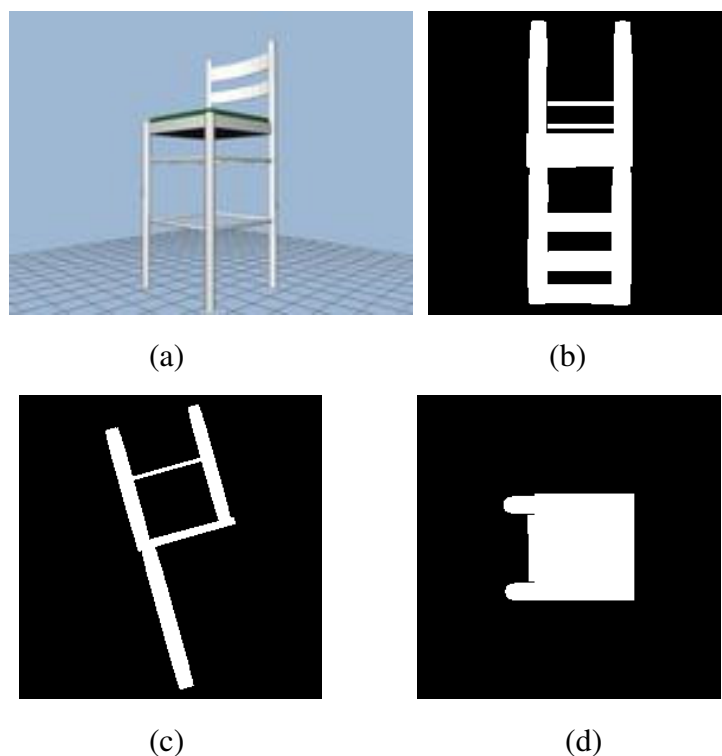


圖 4.8 編號 816 的椅子的三軸投影圖(椅背長度短，椅腳長度長)：(a)原始未經旋轉的圖形(b)對第一主軸投影、(c)對第二主軸投影、(d)對第三主軸投影。

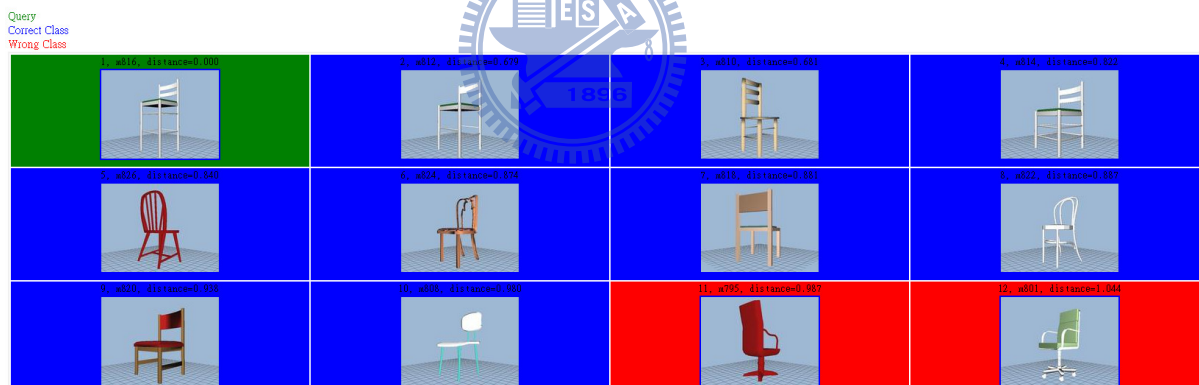


圖 4.9 針對編號 816 在家具資料庫做檢索的結果，左上綠色的為檢索物體依序往右再往下為檢索到的順序，藍色代表正確，紅色代表錯誤。

而發生更為嚴重的主軸未對齊現象是在第 35 號模型的胖鳥，第 112 號模型的胖兔，跟第 389 號模型教堂。但其發生主軸無法對齊的原因主要是因為部分區域差異度太大，或是整體比例已經完全不同。

以下以兔子的模型做例子，其中下面三個皆為兔子模型，但經過前處理後對主軸投影，前兩隻兔子還是能正確辨識，各個投影圖都還是相似型態，但第三隻兔子卻產生主軸跟前兩隻兔子的主軸沒對齊的情形，造成之後檢索錯誤。這種錯誤通常只會發生在基

本形狀已經有所改變，如圖 4.12 的兔子，其 4 隻腳已經跟身體黏合在一起了，跟前兩隻兔子皆有明顯的四隻腳有所差異，而頭頸的部位也跟其他兩隻兔子有所差異，且身體明顯比前兩隻圓滾跟胖，因此在計算主成分分析時才會出現錯誤的主軸，若要修正這種錯誤，必須要找尋更佳的主軸評判工具，或者是藉由整個物體包含體積內部所有點都作主成分分析，而不單只有表面作主成分分析，如此才能避免這種錯誤。

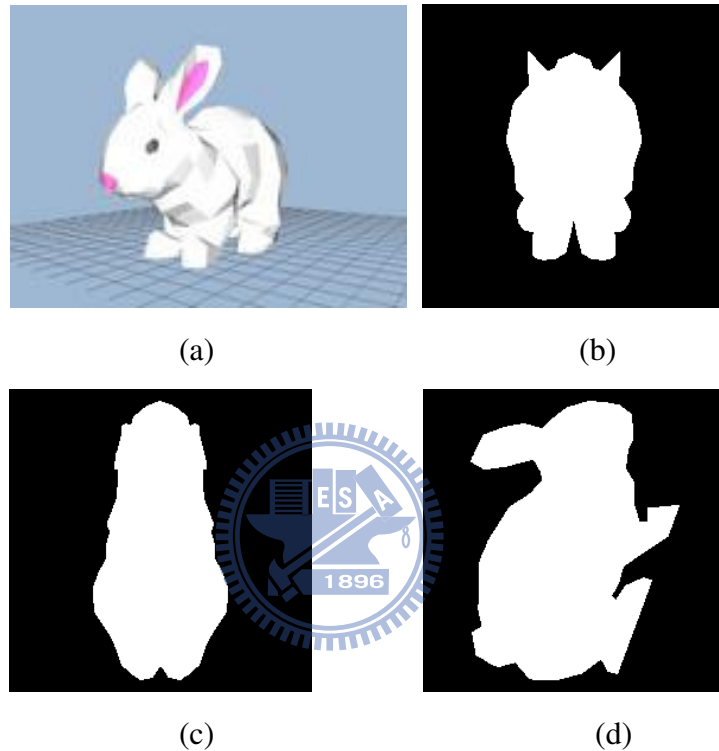


圖 4.10 編號第 110 的兔子投影圖：(a)原始未經旋轉的圖形(b)對第一主軸投影、(c)對第二主軸投影、(d)對第三主軸投影。

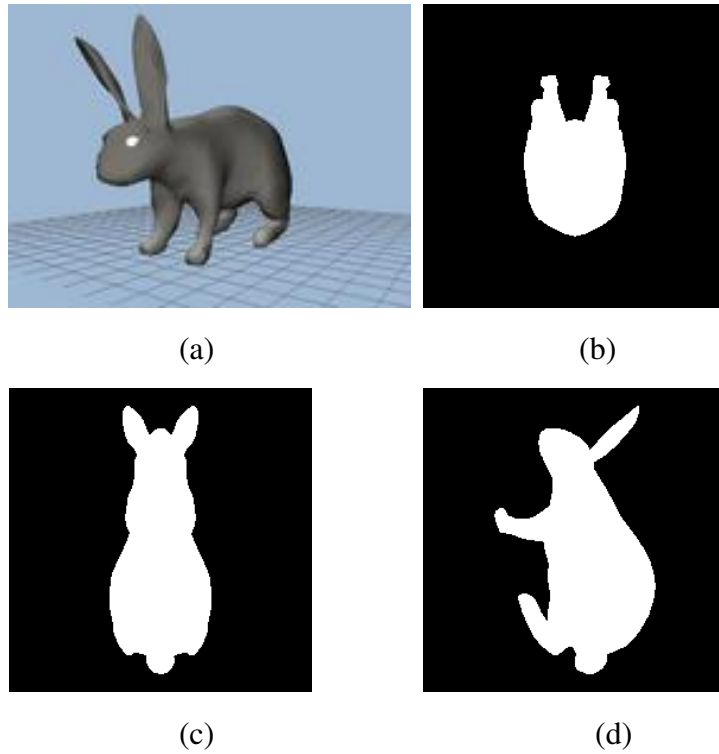


圖 4.11 編號第 111 的兔子投影圖：(a)原始未經旋轉的圖形(b)對第一主軸投影、(c)對第二主軸投影、(d)對第三主軸投影。

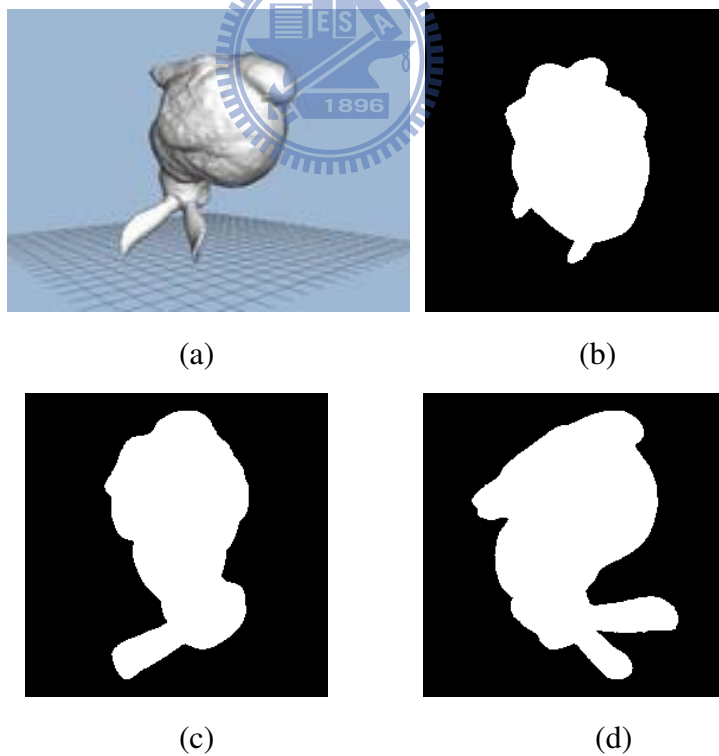


圖 4.12 編號 112 的兔子投影圖：(a)原始未經旋轉的圖形(b)對第一主軸投影、(c)對第二主軸投影、(d)對第三主軸投影。

4.2.2 不同求取主成分分析的方法實驗

主成分分析法對於本論文的成效佔有很大的影響，故在本節分析直接從點求主成分分析、權重主成分分析、格子點化後再以每個格子點求取主成分分析跟連續主成分分析。針對在 PSB 資料庫中編號 3 跟編號 4 的物體做評判，其中這兩個物體在 PSB 的基礎分類中都分為同一個分群：螞蟻，如圖 4.13 跟圖 4.14，首先由圖 4.15 可看出由點求主成分分析的第一主軸投影圖都會往一方偏斜，這將會使得之後檢索時因偏斜角度不同造成便是錯誤；而由圖 4.16 可看出權重主成分分析有可能造成突然的歪斜，使得辨識結果出錯；而從圖 4.17 和圖 4.18 可看出格子點化後再以每個格子點求取主成分分析跟連續主成分分析都能有效的使投影面展現出最具特徵的一面，較不會有偏斜產生，兩物體能有最相似主軸的投影圖，其整體成效如下表 4.2 所示：



圖 4.13 PSB 資料庫中編號 3 的原圖。

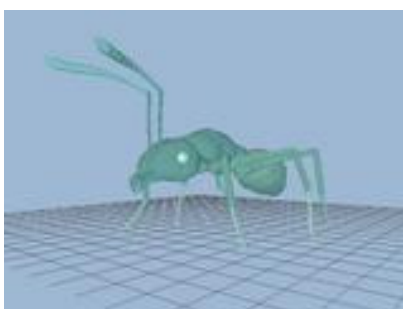


圖 4.14 PSB 資料庫中編號 2 的原圖。

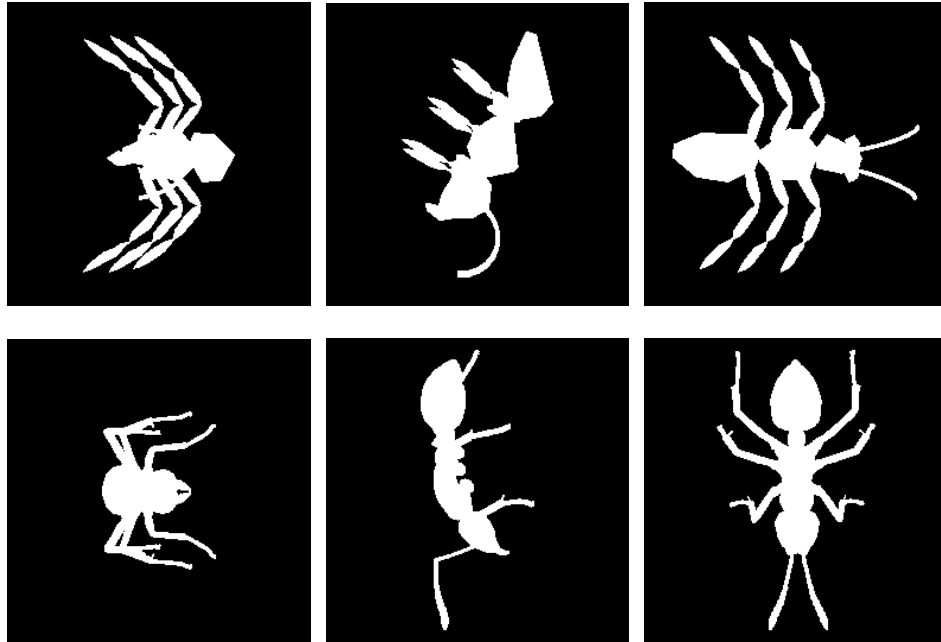


圖 4.15 點求主成分分析(上排為圖 4.13 的三軸投影圖，下排為圖 4.14 的三軸投影圖。)



圖 4.16 權重主成分分析(上排為圖 4.13 的三軸投影圖，下排為圖 4.14 的三軸投影圖。)

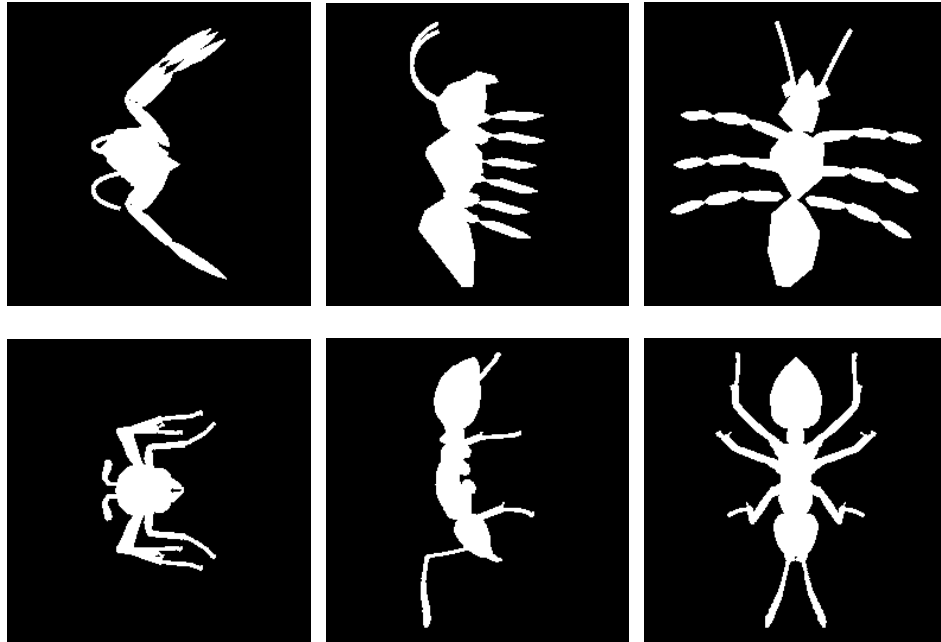


圖 4.17 格子點化後再以每個格子點求取主成分分析(上排為圖 4.13 的三軸投影圖，下排為圖 4.14 的三軸投影圖。)

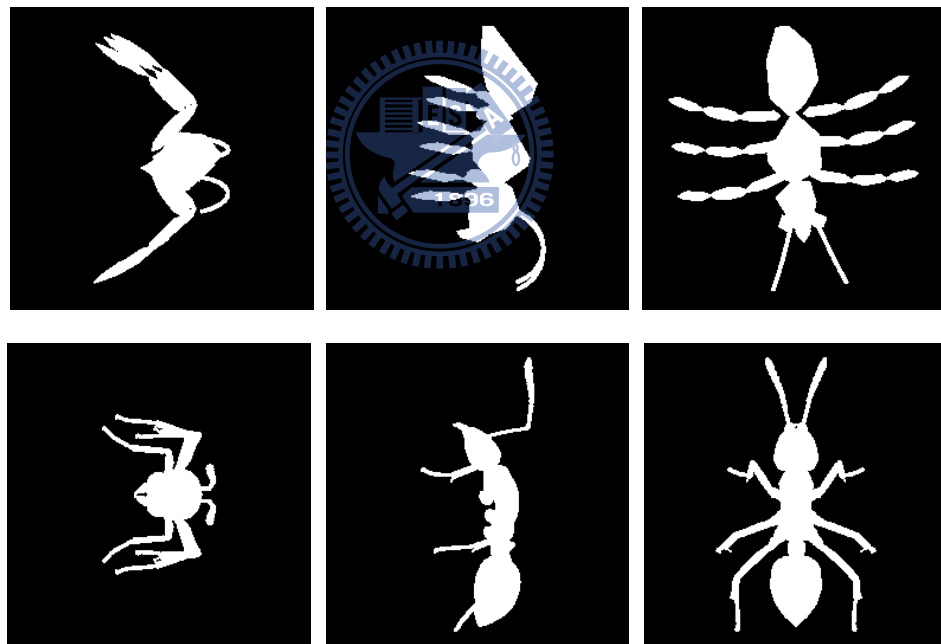


圖 4.18 連續主成分分析(上排為圖 4.13 的三軸投影圖，下排為圖 4.14 的三軸投影圖。)

表 4.2 不同方法求取主成分分析的效能。

方法	NN	first-tier	second-tier	E measure	DCG
連續主成分分析[19]	0.648	0.387	0.494	0.287	0.643
格子點主成分分析	0.63	0.384	0.494	0.288	0.643
權重主成分分析[18]	0.612	0.368	0.469	0.276	0.627
點主成分分析	0.598	0.352	0.453	0.263	0.618

而從表 4.2 明顯看出若只單用 3D 模型的點求取主成分分析的主軸，將造成最嚴重的判別錯誤，主要就是因為只用點的話其代表性不足，分散的點跟密集的點其權重值皆相同，使得整體模型容易因為部分密集點而歪斜，而權重主成分分析明顯就提昇了不少正確率，因為他對每點考慮到相鄰的面的面積大小，以此作為權重，故每點較有代表性，但此種權重考量只適合均勻大小的三角面，當三角面的大小不一致時所計算出之權重值也將不準確而造成辨識錯誤，而為了更加提高正確率而先把所有面改用點來表示形成格子點，使每個點的代表性都一致，則使主軸判斷更為正確但也更為耗時，最後以連續主成分分析把三角面上所有點均有考慮到，使得結果為最佳，因此可以得出把物體的點分佈考慮的越完全則主軸也將越精準的結論。但對於像兔子整體都有所變化，且連腳都合併在身體內時，若無法考慮整個體積所包含的點，則勢必會判斷出錯，而要在加以改進的話，則必須尋找更佳的主軸評判方法。

4.3 檢索效能評判

本論文使用的物體主要是 PSB 資料庫中的物體，首先分析 PSB 資料庫基礎分類(base classification)中的測試(test)分群的檢索效能，其次在使用較小且較少分群的資料庫做檢索效能測試並分析在不同的資料庫中本篇論文的優勢。

4.3.1 PSB 資料庫檢索

在 PSB 資料庫中的基礎分類中的測試分群內總共有 907 個物體和 92 種分群。最大的群體佔整體的 6%。在 PSB[3]中有列出其他 12 種方法以及其在此種環境的效能：

D2 Shape Distribution (D2)：藉由統計兩點間的距離分布作為特徵的檢索方法。

Extended Gaussian Image (EGI)：B. Horn[11]所提出的方法，以球函數統計物體表面方向向量的檢索方法。

Complex Extended Gaussian Image (CEGI)：以複數球函數統計物體表面方向和距離的分布作為檢索特徵的方法。

Shape Histogram (SHELLS)：M. Ankerst 等人[4]、[5]所提出的方法，以統計點到中心的距離分布作為特徵的檢索方法。

Shape Histogram (SECTORS)：M. Ankerst 等人[4]、[5]所提出的方法，以統計空間中每個角度的點分佈作為特徵的檢索方法。

Shape Histogram (SECSHEL)：M. Ankerst 等人[4]、[5]所提出的方法，以統計在不同角度和距離的點分佈作為特徵的檢索方法。

Voxel：以二值表示物體在空間中的位置作為特徵的檢索方法。

Spherical Extent Function (EXT)：以球座標的角度統計從重心到面的最遠距離作為特徵的檢索方法。

Radialized Spherical Extent Function (REXT)：以球座標的角度和距離統計從重心到面的最遠點位置作為特徵的檢索方法。

Gaussian Euclidean Distance Transform (GEDT)：以一個三維方程式作為特徵的檢索方法，其值是藉由把點作高斯歐基里德轉換得到。

Spherical Harmonic Descriptor (SHD)：以旋轉不變的球面諧波表達物體，以其不同頻率的值作為特徵的檢索方法。

Light Field Descriptor (LFD)：藉由各個角度的投影圖作為特徵，尋找相似物體的檢索方法。

其中除了 D2，SHELLS，SHD，和 LFD 皆必須以主成分分析法旋轉主軸。下面是使用 PSB 資料庫第一種分類方法的效能比較表。

表 4.3 十四種方法在 PSB 中的效能。

方法	NN	first-tier	second-tier	E measure	DCG
本論文	0.648	0.387	0.494	0.287	0.643
LFD	0.657	0.38	0.487	0.28	0.643
本論文沒加 3D 特徵	0.574	0.335	0.435	0.253	0.605
REXT	0.602	0.327	0.432	0.354	0.601
SHD	0.556	0.309	0.411	0.241	0.584
GEDT	0.603	0.313	0.407	0.237	0.584
EXT	0.549	0.286	0.379	0.219	0.562
SECSHEL	0.546	0.267	0.35	0.209	0.545
VOXEL	0.54	0.267	0.353	0.207	0.543
SECTORS	0.504	0.249	0.334	0.198	0.529
CEGI	0.42	0.211	0.287	0.17	0.479
EGI	0.377	0.197	0.235	0.139	0.434
D2	0.311	0.158	0.235	0.139	0.434
SHELL	0.227	0.111	0.173	0.102	0.386

其中可看出只有 LFD 在 NN 這項標準下，比本論文的略好，代表其找第一項相似物體的效能比本論文稍好，其原因主要就是因為他採用 20 個面來描述物體，使兩個物體的各個角度的投影特徵都能加以參考，但也因此需要大量的運算與儲存，而且也讓一些相似但並沒有每個投影面均相似的物體變得較難檢索，這也使得其在其他考慮較多個相似物體的評價上都輸給本論文，原因也就是其對於小變化的容忍度較本論文低，在具有差異的相似物體上並不如本論文來的好。而其他方法的所有評價皆落後本論文跟 LFD。

而從本論文跟本論文沒增加 3D 特徵的結果比較，只單使用正規化的傅立葉描述子跟 Zernike 矩，也就是使用跟 LFD 相同的 2D 特徵，但是只使用三個投影面而非 10*10 個投影面時，整體檢索效率將大幅下降，代表本論文所使用的 3D 特徵極具有意義，且並不需要特別運算，反而是減少運算，又能得到比 LFD 還要好的效果。

4.3.2 小型資料庫檢索

在本實驗中主要分成 4 種常用且具有代表性的資料庫作為檢索資料庫：動物資料

庫、家具資料庫、常用工具資料庫、交通工具資料庫。而比較的方法有 LFD[7]、SECTORS[4]、[5]跟 D2[6]三種方法。

動物資料庫中含有螞蟻、蝴蝶、人、飛行鳥、狗、馬、兔、蛇和魚，其相似物體的形狀變異度大，如兔子跟蛇，且相似物體的姿態也會有所差異，尤其是蝴蝶，翅膀角度形狀皆有所不同，故本資料庫最難正確檢索。以下為各個方法對動物資料庫的效能比較圖表。

表 4.4 各方法在動物資料庫的效能比較表。

方法	NN	first-tier	second-tier	E measure	DCG
本論文	0.667	0.51	0.676	0.287	0.736
SECTORS	0.571	0.405	0.516	0.249	0.649
D2	0.349	0.28	0.489	0.266	0.566
LFD	0.683	0.457	0.623	0.272	0.72

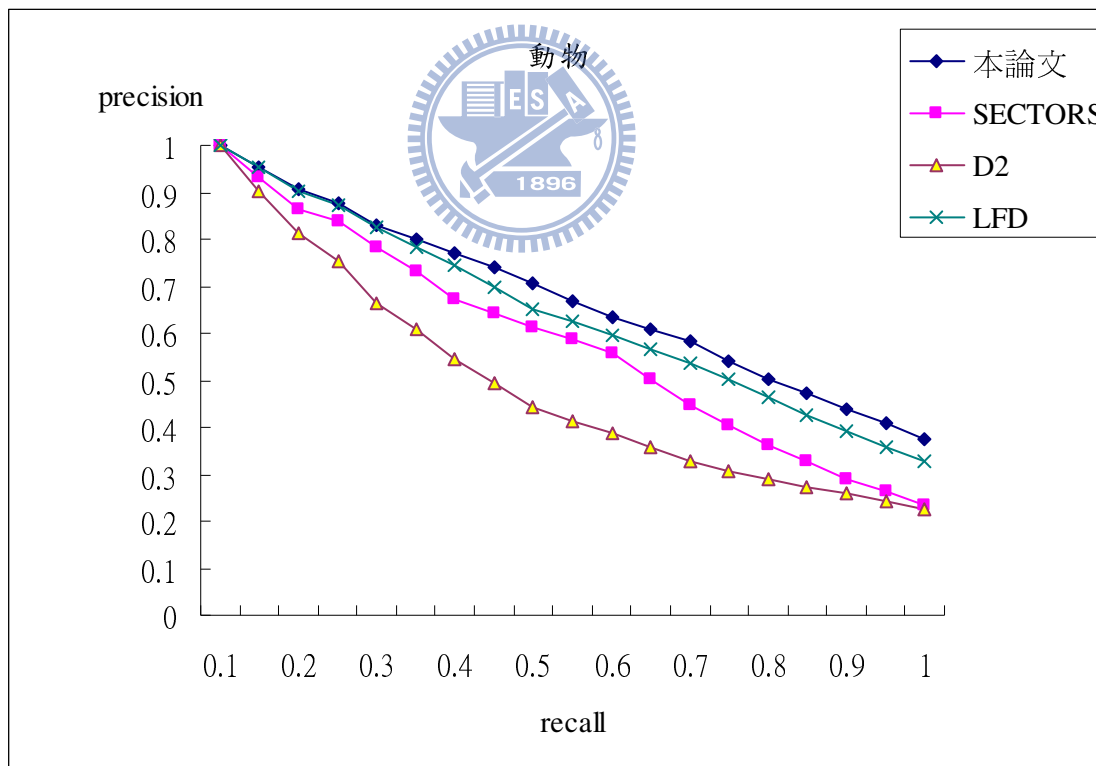


圖 4.19 動物資料庫精準度對回傳率圖。

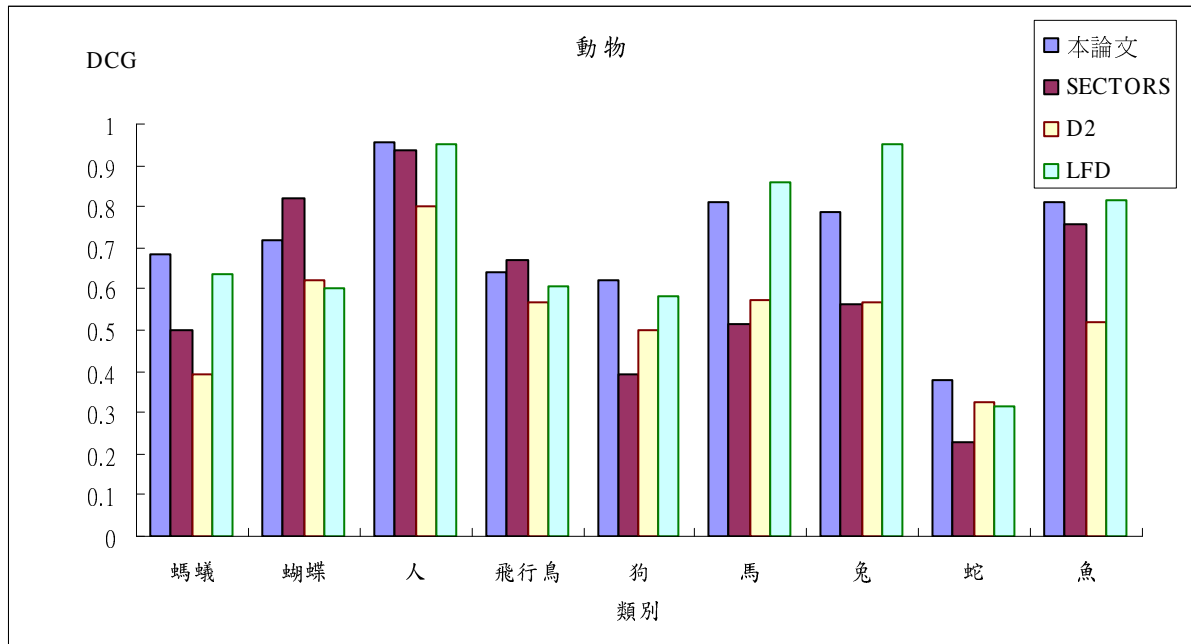


圖 4.20 動物資料庫的類別對 DCG 圖。

首先從表 4.4 可看到跟原本 PSB 資料庫一樣的是，LFD 對於第一個相似判別稍高於本論文，但考慮多個回傳時則輸給本論文。其次從圖 4.18 的 Precision 對 recall 圖中可明顯看出本論文和 LFD 整體上表現最佳，但當回傳增多時，本論文的準確度為最佳。而由圖 4.19 可知在各個分類上的優劣則稍有不同，尤其在蝴蝶分類中 SECTORS 比其他方法都來的優秀，其原因是蝴蝶翅膀形狀差異大，則投影圖的輪廓差異也大，因此使依靠輪廓作為特徵的本論文跟 LFD 的檢索效能下降，而在兔子的分類中 LFD 效能比本論文為佳的原因則就是因為他是對每個角度都做投影，而相似的兔子會發生主軸不一致的問題，造成辨識錯誤，但本論文還是優於天生就有旋轉不變性的 D2 跟同樣有作主成分分析法的 SECTORS，而像蛇這種不定形狀之物體，每種檢索方法對其效能都表現不佳，且就算由人眼判斷此 3D 模型也非常困難，因其不具有鱗片膚色等特徵只有卷曲的形狀。

第二個資料庫為家具資料庫，其中含有螢幕、書桌、長板凳、餐廳椅、電腦椅、櫃子、方桌跟單腳桌。其中部分不同物體，其整體形狀卻類似，如不同樣式的椅子跟桌子，但部分相似物體卻會有部分區域比例不同，因此容易產生相似物體的主軸偏斜的問題，如各種椅子的椅背形狀大小不同，椅腳長短不一等。

表 4.5 各方法在家具資料庫的效能比較表。

方法	NN	first-tier	second-tier	E measure	DCG
本論文	0.771	0.625	0.753	0.35	0.815
SECTORS	0.657	0.397	0.542	0.289	0.685
D2	0.486	0.331	0.494	0.293	0.611
LFD	0.8	0.525	0.691	0.347	0.779

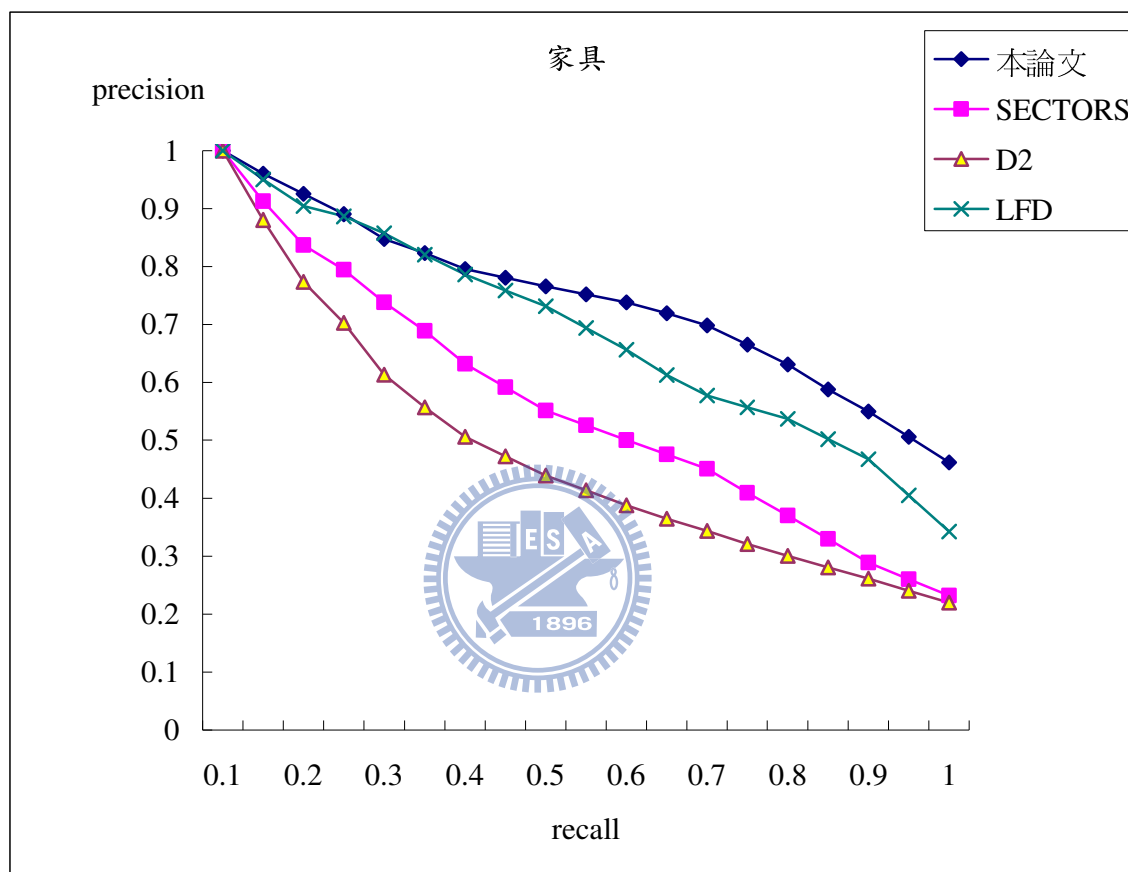


圖 4.21 家具資料庫精準度對回傳率圖。

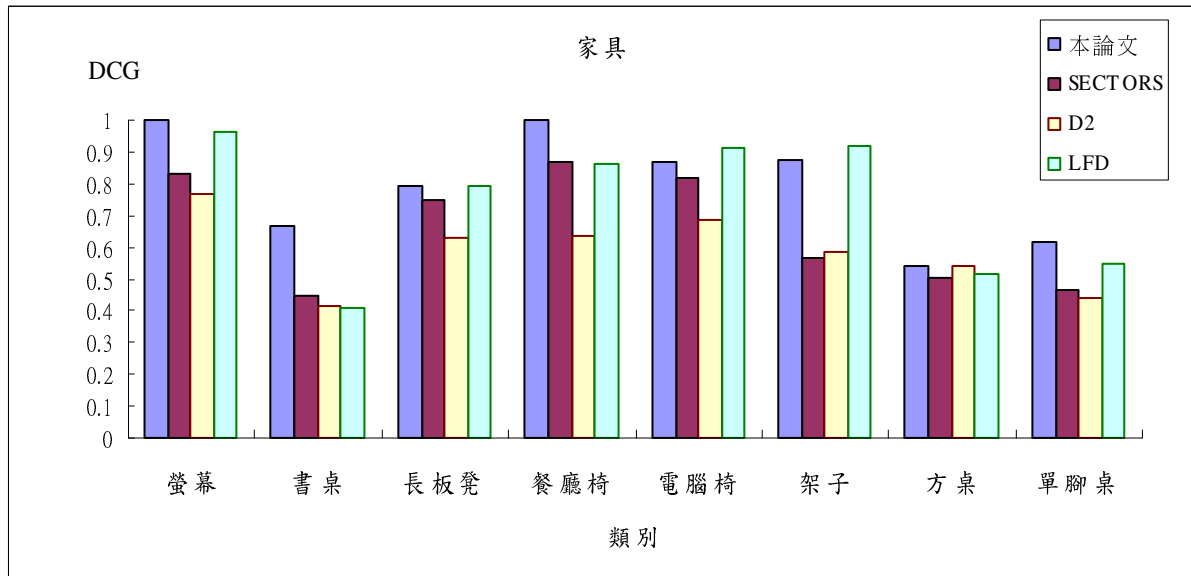


圖 4.22 家具資料庫的類別對 DCG 圖。

首先從表 4.5 可看出，只有 LFD 在第一相似物回傳準確度略贏本論文，但在找尋多物體時本論文皆為最佳。其次從圖 4.20 可看出本論文在精準率對回傳率圖中表現非常優異，當回傳七成的物體時都還有七成的正確率，而 LFD 當回傳不到六成物體時，其正確率就已經不到七成，明顯本論文對於多物體的回傳正確率較好，而另兩種方法更是遠落後本論文。最後從圖 4.22 可看出本方法特別在書桌有最好的 DCG 評判值而 LFD 卻只跟另兩種方法差不多，其原因就是本方法考慮到物體的比例，而對於三種桌子能有較大的差異度，是因為我取的是三個互相垂直的面，其特徵重複性較小，而 LFD 取多個角度投影面，雖然可以完美克服旋轉變量，但也容易造成部分特徵重複，使得判斷相似度時，雖然有少部分角度差異大，但因為其他角度的投影圖都很類似，造成其他物體每面都有部分差異，但整體的相似度卻跟只有一面很大差異的相似度差不多，造成誤判。且在本資料庫中對於各種可能造成主軸偏斜的物體，還是有很好的 DCG 值，可看出因為部分區域不同所造成的主軸偏斜情形，對整體檢索影響並不大。

第三個資料庫是常用工具，其種類有斧、長刀、劍、眼鏡、帽子、砂漏、杯子、郵箱、搥子、剷子、雨傘、輪子、齒輪。此資料庫的特徵是不同的物體差異只在部分區域，如：斧、搥子跟剷子都只有最前端有不同，刀跟劍只有劍刃的差異，輪子跟齒輪只有邊緣的差異。

表 4.6 各方法在常用工具資料庫的效能比較表。

方法	NN	first-tier	second-tier	E measure	DCG
本論文	0.71	0.557	0.738	0.283	0.765
SECTORS	0.538	0.461	0.622	0.252	0.684
D2	0.43	0.331	0.487	0.23	0.577
LFD	0.656	0.491	0.674	0.278	0.73

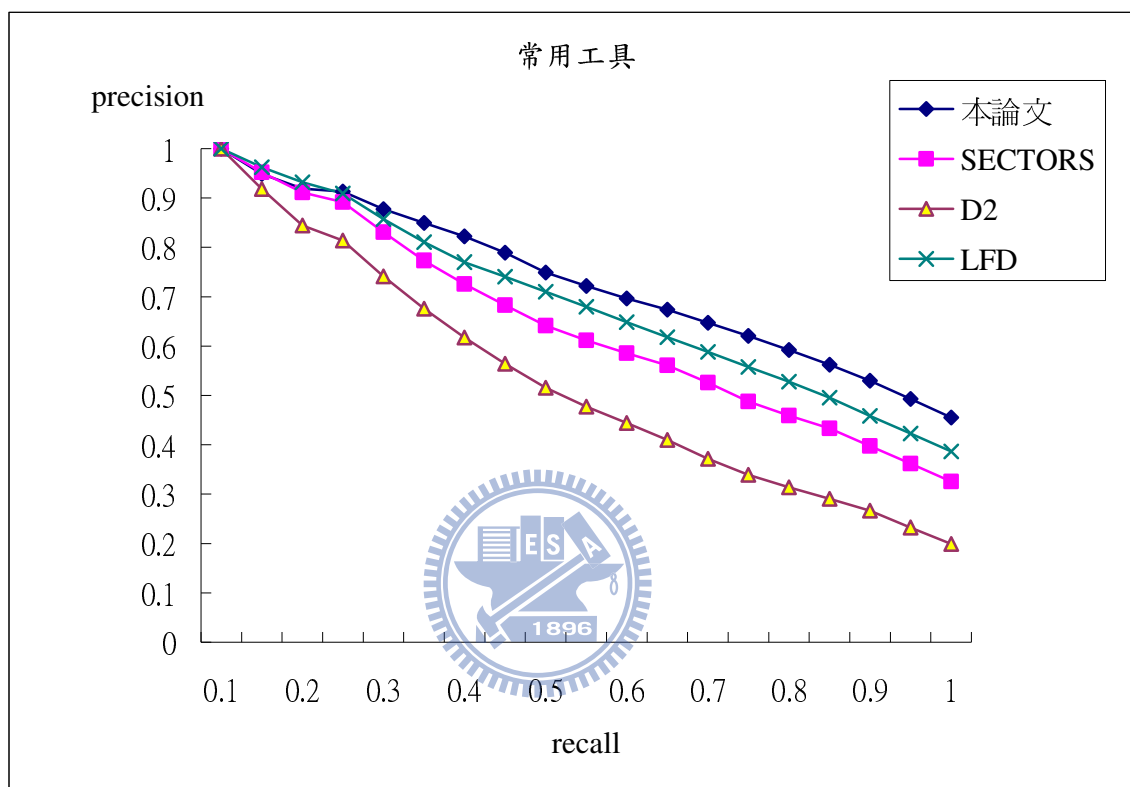


圖 4.23 常用工具資料庫精準度對回傳率圖。

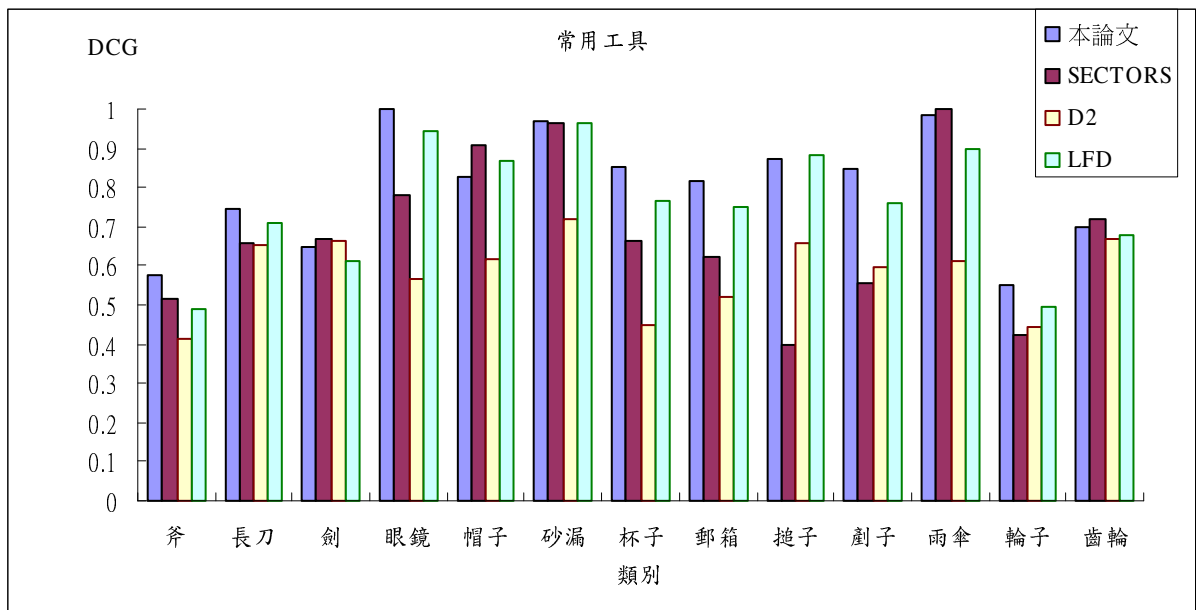


圖 4.24 常用工具資料庫的類別對 DCG 圖。

從表 4.6 中可看出本論文在此種分類上全部評比都最佳。而由圖 4.23 可知本論文在精準率對回傳率圖的檢索效果也是最佳，就算物體差異只在部分區域但都還是能正確檢索，不像 SECTORS 在差異較大的物體才能有正確回傳，當只有部分差異時，效果就變得極差，如圖 4.24 的搥子跟剗子，其原因就是 SECTORS 的方法是以空間分佈作為特徵，而搥子跟剗子在同區域分佈皆相似，主要是差在區域形狀，也就使得 SECTORS 在此部分鑒別度不足。

第四種資料庫是交通工具，其種類有飛機、熱氣球、太空船、船、潛艇、賽車、轎車、摩托車、卡車。此資料庫的特徵是個個模型均為剛體，表面平滑且相似物體形狀較少差異，可說是最容易分類的資料庫。

表 4.7 各方法在交通工具資料庫的效能比較表。

方法	NN	first-tier	second-tier	E measure	DCG
本論文	0.836	0.665	0.817	0.327	0.858
SECTORS	0.672	0.446	0.601	0.292	0.728
D2	0.478	0.388	0.608	0.285	0.662
LFD	0.851	0.639	0.849	0.335	0.852

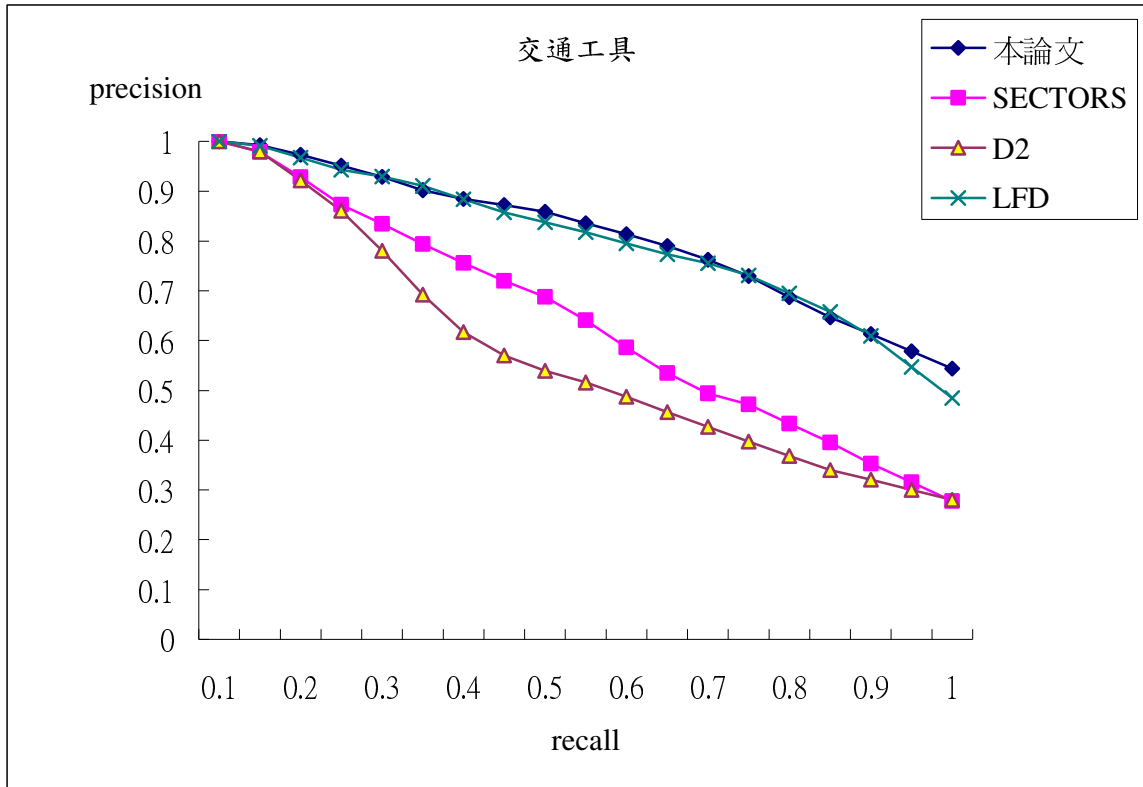


圖 4.25 交通工具資料庫精準度對回傳率圖。

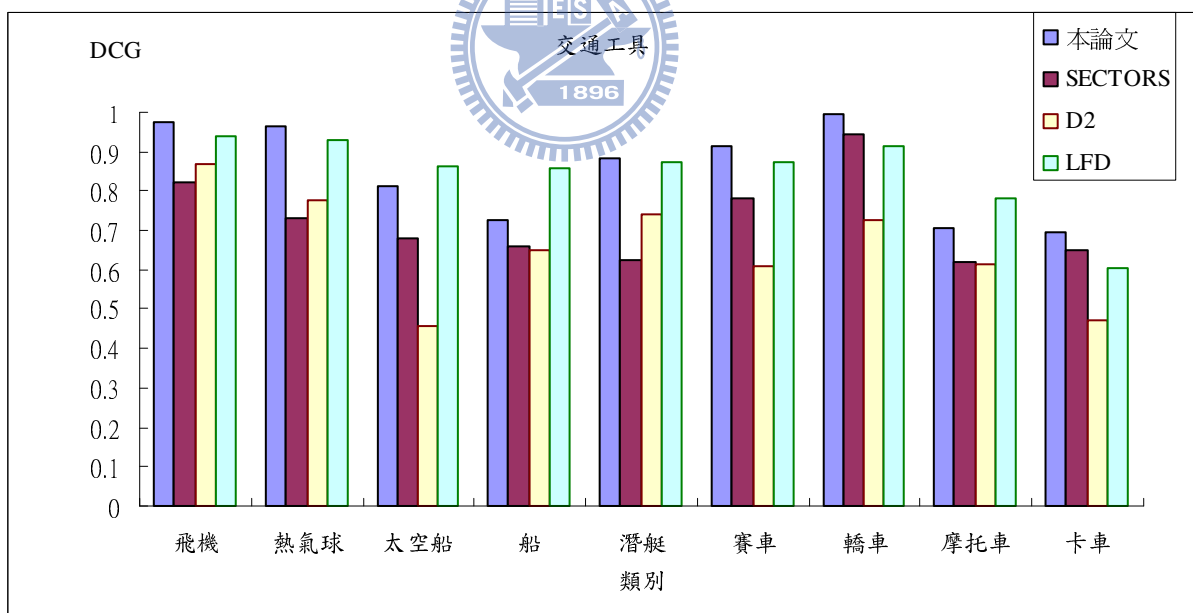


圖 4.26 交通工具資料庫的類別對 DCG 圖。

從表 4.7、圖 4.25 跟圖 4.26 可看出本論文在極容易分類的資料庫中跟 LFD 幾乎對於所有評比都不相上下，而對於精準率對回傳率圖，當回傳八成的物體時也都還有七成正確率，比前三個資料庫都來得好。雖然 LFD 效果根本論文差不多，但其所需要花費

的時間跟儲存卻是本論文的好幾倍，等於儲存跟計算了一堆不必要的資訊，使得在比對時更花時間，以下將說明其所需時間跟空間需求。

4.3.3 計算複雜度與空間需求

本節將比較各方法的運算所需時間跟儲存所需的空間，其中 LFD、SECTORS、D2 跟本論文的儲存空間皆實際計算，其他方法的儲存空間大小皆參考自 PSB[3]如下表：

表 4.8 儲存一個模型的空間需求表。

方法	空間(bytes)
本論文	183
LFD	9,000
REXT	17,416
SHD	2,184
GEDT	32,776
EXT	552
SECSHELL	32,776
VOXEL	32,776
SECTORS	552
CEGI	2,056
EGI	1,032
D2	136
SHELL	136

從表 4.8 可看出本論文所需求的空間，遠小於 LFD，明顯降低空間需求，且所需要的空間也僅次於 D2 跟 SHELL，但本論文在實驗中物體檢索的正確度卻遠高於 D2 跟 SHELL，如在 PSB 資料庫實驗中本論文對整個資料庫的 DCG 值為 0.643 遠比 D2 的 0.434 跟 SHELL 的 0.386 好得多。

計算所需時間將比較 LFD、SECTORS、D2 跟本論文，如下表：

表 4.9 時間需求表。

方法	產生特徵時間(s)	比對時間(s)
本論文	1.1	0.001
LFD	4.3	0.002
SECTORS	0.9	0.001
D2	1.2	0.001

要比對一 3D 模型必須從產生特徵再到跟資料庫內的特徵向量比對，故要比對物體則必須把產生特徵時間跟比對時間相加。從表 4.9 可看出各個方法的比對時間都遠小於特徵產生時間，而本論文產生特徵所需的時間遠比 LFD 來的要低，其原因就是省去製作 100 個影像而只用三個影像取代，且新的特徵並不需要再花更多的時間去求取。而其他兩種方法跟本論文的速率也相差不多。

4.4 特徵需求實驗

在本節將探討各別單獨特徵的效能並比較其重要性，並使用不同數量的傅立葉描述子跟 Zernike 矩特徵互相結合做檢索以得到最佳檢索效果，而從實驗可知使用過多的特徵並不一定會增加整體的正確度，而如何以較少的特徵又維持高正確率則為本實驗的目的。

4.4.1 各別特徵重要性

本節針對傅立葉描述子跟 Zernike 矩作特徵選取(feature selection)的重要性分析，藉由每個特徵的算數平均數算出其分別的重要性，越大的算數平均數代表其變動越大，越難精準的分類，而越小的算數平均數則代表有較大的重要性，而傅立葉描述子 $F(0)$ 到 $F(10)$ 跟 Zernike 矩的 $Z(0)$ 到 $Z(35)$ 的特徵的重要性如圖 4.27，可針對使用者所需要多少程度的檢索正確度加以選擇。其中可看出 $F(0)$ 相當重要也就是大小比例是很重要的特徵。

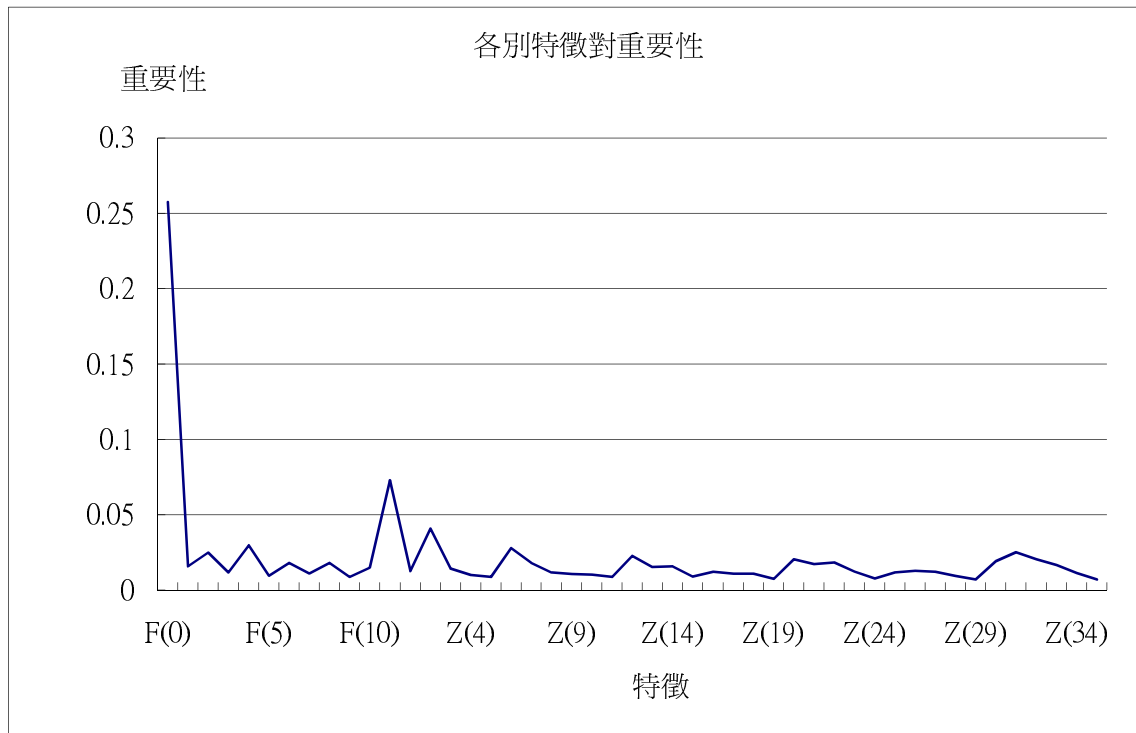


圖 4.27 各別特徵對重要性。

4.4.2 大型資料庫的特徵需求

本節以大型資料庫 PSB 資料庫作單獨特徵跟不同特徵數的檢索效能分析，首先單獨各個特徵來分析，其中 $F(n)$ 代表傅立葉描述子的第 n 個特徵，而 $Z(m)$ 是代表 Zernike 矩的第 m 階特徵，稍有不同。從圖 4.28 觀察傅立葉描述子的各個特徵效能，從中可知 $F(0)$ 最為重要，而 $F(0)$ 也就是單純的比例大小關係，從此也可看出本論文所選定的第三特徵有其重要性，但接下來的 $F(1)$ 的效能卻差上許多，主要原因就是這特徵太過簡單，並非主要的判別特徵，在接下來的特徵效能則是依序減少，差異性不高；而從圖 4.29 觀察 Zernike 矩的各個特徵效能，其中 $Z(0)$ 還是顯得很重要，而 $Z(1)$ 則顯得較為不重要，而且從圖中可發現奇數的特徵都表現較差，其原因可從圖 3.8 推測出是因為奇數的特徵較容易被旋轉變量影響到，而偶數特徵皆有同心圓較不會被旋轉變量影響。

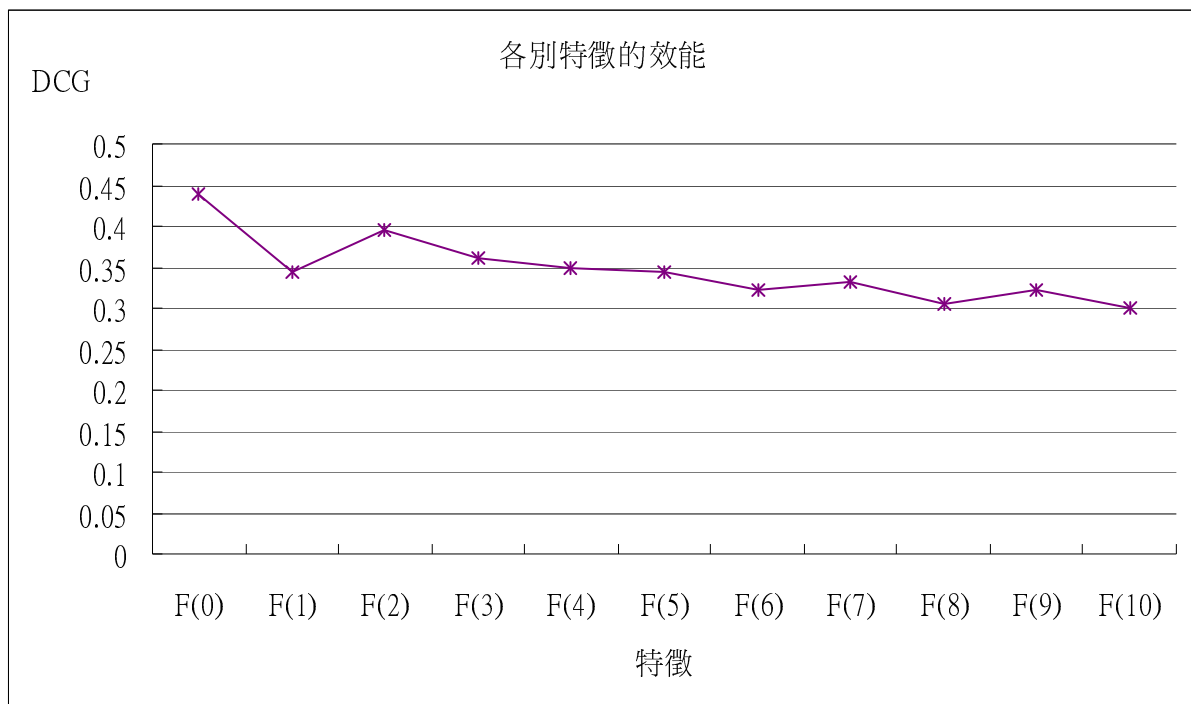


圖 4.28 各別傅立葉描述子特徵的效能。

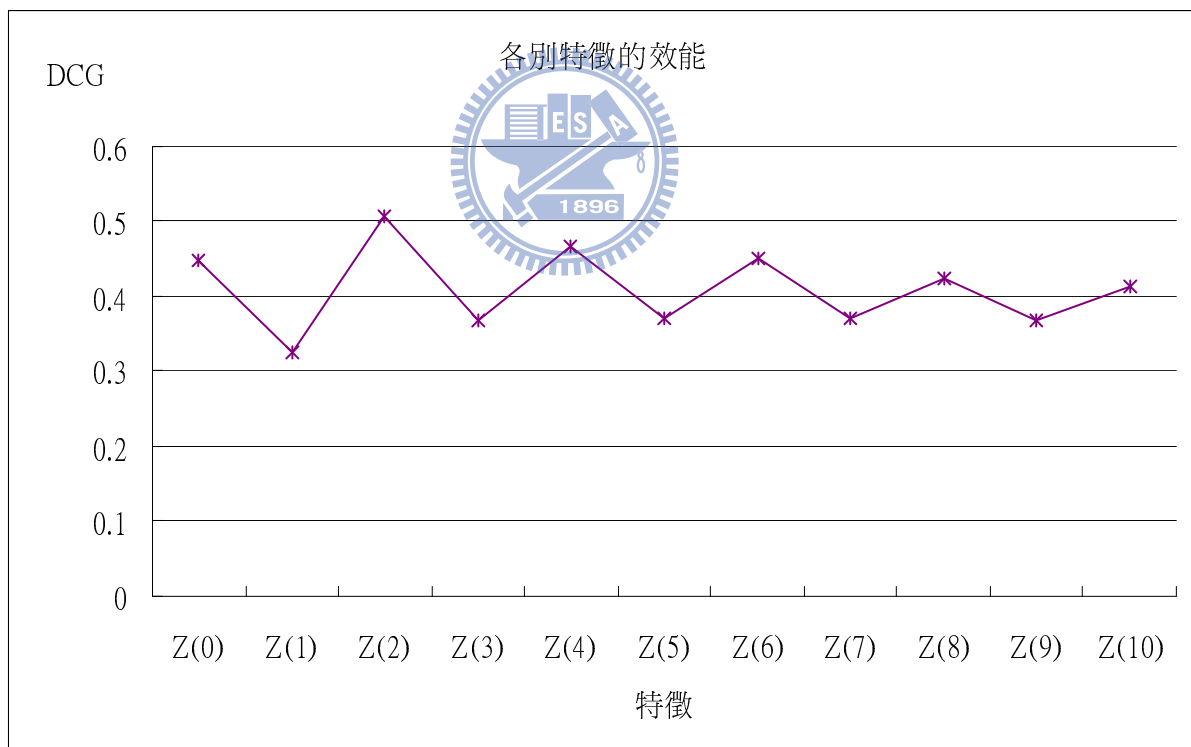


圖 4.27 各別 Zernike 矩特徵的效能。

其次以不同特徵數作檢索效能分析，如下表 4.10，首先是單獨傅立葉描述子跟單獨 Zernike 矩作檢索效能，最後則是兩種結合的效能。首先從前項實驗可知 F(1)跟 Z(1)的檢索效果都較差，但如果直接把他們去掉所造成的影響卻比刪除掉較高階其單獨效能較好的特徵來得大，如表 4.10 的編號 1、2、3 跟 8、9、10，其原因應該是各階特徵依序

有其相關性，低階特徵主要是較籠統的形狀或邊界，而高階特徵主要是邊界或形狀的細微變化，所以使用高階跟低階所檢索出來的結果雖然都有好的效能，但兩者檢索的答案卻可能相衝，兩者的檢索結果可能一個是對的一個是錯的，合併起來變成都錯，使得結果反而更糟。故如果要減少特徵數量，盡量以高階特徵先做刪除為佳，因為低階特徵整體表現較佳。

以下以傅立葉描述子作為特徵，從表 4.10 可看出當傅立葉描述子特徵取到 F(7)即為最小特徵數，再取更少特徵將造成正確率下降，而從 Zernike 矩特徵單獨檢索來看，取到第 8 階最佳，再更少特徵會使 NN 下降許多。而因為 Zernike 矩的一階將會影響數個特徵值，故先決定 Zernike 矩階數再決定傅立葉描述子的特徵數量，而從表 4.10 的編號最後三項可知取 Zernike 矩到 8 階加傅立葉描述子到第 10 項特徵值為最佳，故本論文以這些特徵作為之前實驗的特徵數，不過在此隱含者一個問題，傅立葉描述子跟 Zernike 矩相加的權重值，在下表中傅立葉描述子跟 Zernike 矩相結合的效能都有找出最佳權重值，但都是以嘗試錯誤法來找最佳效能。而且在 PSB 資料庫中，權重值影響較小，原因是整體模型多，有些較容易以傅立葉描述子辨識有些較容易以 Zernike 矩辨識互有影響，使得整體效果差異不大，但在下面小型資料庫中，權重值的變化影響就較為激烈。

表 4.10 在大型資料庫中不同特徵數檢索效能表。

編號	所含特徵	NN	first-tier	second-tier	E measure	DCG
1	F(0)+F(2)~F(10)	0.513	0.284	0.381	0.226	0.561
2	F(0)~F(10)	0.527	0.292	0.392	0.233	0.568
3	F(0)~F(9)	0.526	0.292	0.393	0.232	0.568
4	F(0)~F(8)	0.530	0.291	0.391	0.232	0.568
5	F(0)~F(7)	0.526	0.290	0.392	0.231	0.568
6	F(0)~F(6)	0.523	0.286	0.388	0.230	0.565
7	F(0)~F(5)	0.521	0.283	0.386	0.228	0.561
8	Z(0)+Z(2)~Z(10)	0.58	0.336	0.44	0.259	0.606
9	Z(0)~Z(10)	0.579	0.341	0.446	0.261	0.607
10	Z(0)~Z(9)	0.582	0.341	0.449	0.260	0.607
11	Z(0)~Z(8)	0.585	0.340	0.443	0.259	0.606
12	Z(0)~Z(7)	0.571	0.343	0.445	0.260	0.606
13	Z(0)~Z(6)	0.573	0.336	0.438	0.257	0.599
14	F(0)~F(10)+Z(0)~Z(8)	0.648	0.387	0.494	0.287	0.643
15	F(0)~F(9)+Z(0)~Z(8)	0.643	0.388	0.494	0.286	0.642
16	F(0)~F(8)+Z(0)~Z(8)	0.620	0.384	0.493	0.286	0.642

4.4.3 小型資料庫的特徵需求

本節以之前所提出的家具資料庫作為小型資料庫的特徵需求實驗範例，首先單獨各個特徵來分析，首先從圖 4.30 可看出傅立葉描述子的 F(0)特徵還是跟在大型資料庫中一樣重要，F(1)同樣檢索效果比較差，而越後面的特徵檢索效果也較差，不過有較為激烈的震盪，主要原因是因為模型較少時各物體整體相似度較高，故比較可能會有單一特徵較為重要的情形。再從圖 4.31 可看出 Zernike 矩的 Z(0)也還是很重要，Z(1)同樣下降很多，但之後較高階的特徵，不再像大型資料庫時一樣有較劇烈的變化，原因是在家具資料庫中，在做 PCA 之後都能很有效的轉為相似主軸，故在此階段已經比較不再需要旋轉不變性的特徵，也就是旋轉變量所能影響的更小，故不管是奇數或偶數階效能不再差異巨大。

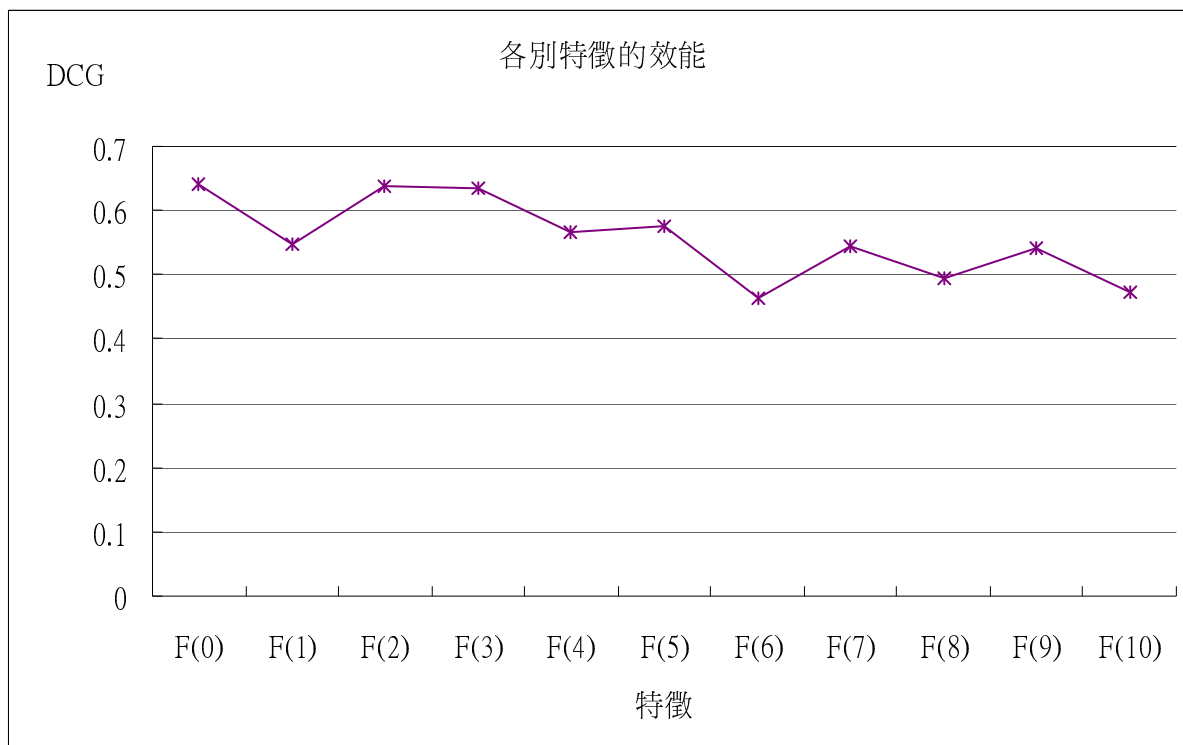


圖 4.28 各別傅立葉描述子特徵的效能。

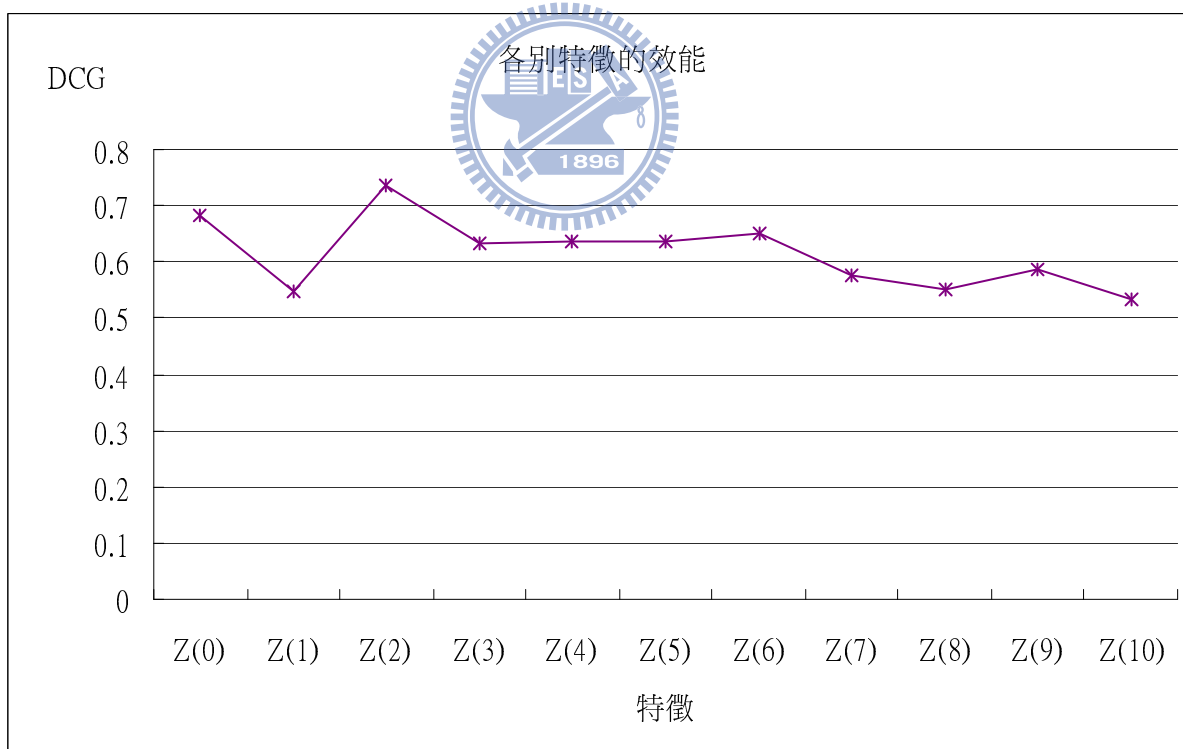


圖 4.29 各別 Zernike 矩特徵的效能。

其次以不同特徵數作檢索效能分析，如下表 4.11，首先是單獨傅立葉描述子跟單獨 Zernike 矩作檢索效能，最後則是兩種結合的效能。同樣從前項實驗可知 F(1)跟 Z(1)的檢索效果都較差，而如果直接把 F(1)去掉所會造成 NN 嚴重下降遠比刪除後面高階特徵

來的影響更大，如表 4.11 的編號 1、2、3，但是 Zernike 矩卻跟大型資料庫檢索時不一樣，刪除 Z(1)後是所有不同數量特徵實驗中表現最佳的，原因是當資料庫較小時，所需的特徵較少，過多的特徵只會帶來更多的錯誤，故對小型資料庫作檢索時，能依靠不同特徵的重要性再加以減少特徵數量。而且從表 4.11 可看出就算整合 Zernike 矩跟傅立葉描述子其效能還是比不上單獨 Zernike 矩刪掉 Z(1)的效能。而在從編號 10 的實驗 PSB 檢索特徵數跟其結合權重值跟編號 11 的實驗 PSB 檢索特徵數再最佳化權重值，是有一段效能差距的，故如何決定最少特徵數跟決定如何相結合兩特徵的權重值都是非常重要的問題。不過以上這些特徵還是含有之前所提出的比例大小相關性特徵，也就是對比例大小不正規化，當具有比例大小不變性時則效能會明顯下降，如編號 16 的數據其作比例大小正規化，相對於編號 8 的數據同樣的特徵數，但整體效能差異甚巨。

表 4.11 在小型資料庫中不同特徵數的效能表。

編號	所含特徵	NN	first-tier	second-tier	E measure	DCG
1	F(0)+F(2)~F(10)	0.729	0.551	0.701	0.344	0.775
2	F(0)~F(10)	0.757	0.558	0.689	0.342	0.775
3	F(0)~F(9)	0.757	0.551	0.686	0.342	0.774
4	F(0)~F(8)	0.743	0.558	0.692	0.342	0.773
5	Z(0)+Z(2)~Z(10)	0.814	0.61	0.744	0.36	0.823
6	Z(0)~Z(10)	0.800	0.611	0.750	0.359	0.820
7	Z(0)~Z(9)	0.800	0.595	0.762	0.360	0.820
8	Z(0)~Z(8)	0.800	0.608	0.742	0.357	0.820
9	Z(0)~Z(7)	0.771	0.592	0.742	0.354	0.811
10	F(0)~F(10)+Z(0)-Z(8)	0.771	0.625	0.753	0.35	0.815
	最佳化					
11	F(0)~F(10)+Z(0)~Z(8)	0.800	0.633	0.756	0.351	0.820
12	F(0)~F(10)+Z(0)~Z(7)	0.829	0.630	0.756	0.350	0.818
13	F(0)~F(10)+Z(0)~Z(6)	0.757	0.623	0.748	0.348	0.815
14	F(0)~F(9)+Z(0)~Z(7)	0.829	0.623	0.756	0.350	0.819
15	F(0)~F(8)+Z(0)~Z(7)	0.814	0.622	0.756	0.350	0.819
16	Z(0)~Z(8)+正規化	0.729	0.483	0.663	0.328	0.738

4.5 實驗分析

在本章第二節提出對於使用主成分分析法的問題：主軸順序顛倒，主軸方向不一致跟主軸偏斜。首先對於第一項問題，如果只像其他演算法不作主軸順序重新判別的動作，則由表 4.1 可看出整體效能將會損失 1~5%的檢索正確率，因此可知其實主成分分析法對於主軸順序顛倒的影像不是很大，但從之後的實驗可知本論文明顯比其他使用主成分分析法的效果來的要好。主要原因就是第二項問題跟第三項問題，主軸方向不一致跟主軸偏斜。本論文因為是使用也具備旋轉不變性的傅立葉描述子跟 Zernike 矩，故對於主軸旋轉 45 度，90 度等不同角度時，所產生的特徵向量還是非常相似，而其他方法不能在此階段使用具備旋轉不變性的特徵抽取，故在此會有大量的錯誤檢索。而對於主軸偏斜是本篇論文相較之下略弱於 LFD 的部分，對於稍微偏斜的椅子資料庫還是能正確檢索，如圖 4.9 所示，其對於因部分區域比例不同而造成特徵影像偏斜，還是能完全正確的檢索。而在大量偏斜的情況如圖 4.10-4.12 的兔子，從之後的動物資料庫中的圖 4.14 中對於兔子檢索可看出，本論文確實在此會稍弱於 LFD，但對於其他方法不管有無使用主成分分析法都較好，可知這部份也就是 LFD 最適合檢索之部分，而對於本論文是相對比較不適合的。

在 4.2.2 節的部份，明確顯示當主成分分析法取的點過少時會對整個檢索效能有大幅度的下降，可知主成分法的正確性對整個檢索效能影響很大，這也就代表了要有好的姿態判定，才能良好的克服旋轉變量，故如果要改良整個演算法，找出更好的姿態判定方法將能使整體檢索效能更為上升。

從第三節的各種不同的資料庫實驗可知本論文在每種資料庫中各個評判標準中均有最佳或次佳的檢索效能。首先在 PSB 資料庫中跟其所列之方法比較，本論文之效能跟 LFD 約略相同，但都遠優於其他方法，證明使用 2D 檢索方法還是比單純只用 3D 特徵抽取的方法，其檢索效能來得好，這在 PSB[3]中的論文也有提到，但在 PSB 論文中提出 LFD 需要過多的空間跟時間為一大缺點，而為了改進此一缺點，我增加了一簡單的

3D 特徵，刪除了過多的投影面，使空間時間大幅下降，而且從 4.3.3 節可明顯看出本論文改善此一大缺點又不失去整體的檢索正確率；其次在動物資料庫的實驗中證明本論文在較複雜的資料庫中也能擁有最佳的檢索效果，對於狗馬等相似物體還是能具有足夠鑒別能力；而在家具資料庫實驗中證明對於部分改變所造成的偏軸情形影響不大，例如其中的各項椅子檢索都能有最好的 DCG 評判值；而從常用工具資料庫實驗中，可知本論文對於區域部分變化的鑒別力還是非常好，例如斧、槌子和剷子，雖然只有區域不同，但還是能分辨出其不同的部位，不像 SECTOR 完全鑒別不出差異；而在最簡單分類的交通資料庫實驗中，LFD 跟本論文幾乎都相差不多，原因就是因為很好分類，而本論文之特徵跟 LFD 又相似故相差不多，但本論文所需時間空間都遠小於 LFD，故再加以考慮時間空間因素，本論文再此資料庫實驗中效率還是最好。

而從整體性來比較本論文跟其他方法，雖然 LFD 跟 D2 都擁有最完備的旋轉不變性，但是 LFD 卻要犧牲極大的時間跟空間，來做特徵抽取儲存跟判斷，相對於其他所有方法，明顯要花上更多時間跟空間，且當資料庫更為龐大時，其所要花的時間也就越是增多，且因為其對一個物體要作二十個面的投影影像取特徵，使得特徵重疊性高，代表性不足，容易造成前面所講的當只有少數投影面差異大，而其餘差異小時會跟全體投影面都有差異時相似度卻差不多，造成之後檢索錯誤。而 D2 因為只是統計距離，造成其鑒別度完全不足，在每個資料庫中效能都墊底，在各個評判標準完全比不贏本論文，而 SECTORS 因為跟本論文一樣必須做主成分分析法，但卻沒有像本論文一樣有具有旋轉不變性的特徵抽取方法與旋轉比對方法，故其對於旋轉變量的影響甚大，容易因為有偏軸或主軸顛倒的情況造成判斷錯誤，而且其對於部分區域有所不同的模型鑒別度不足，因是考慮空間分佈，而非形狀的判別，使得只要在那區塊有相同的分佈就算形狀不同也會被判定為相似，不過也因其不考慮區域形狀，使得其對於相似物體但區域有所變化的容忍度最高，但在各評判標準的效能上還是落後於本論文。

最後再從特徵需求分析特徵的重要性，首先在大型資料庫中，單獨分析傅立葉描述子跟 Zernike 矩的一項特徵，可看出代表大小的 $F(0)$ 跟 $Z(0)$ 有很高的檢索正確度，其原

因就是相關比例大小的重要性，而之後的 F(1)跟 Z(1)的檢索正確度都下降許多，其原因就是低階特徵屬於較為籠統簡單比較難分辨物體，但從實驗中卻發現若就此刪除此特徵，卻會造成檢索正確度下降許多，比刪除高階但單獨重要性較高的特徵還來的不好，其原因是低高階所檢索出來的答案不同，雖然都有好的檢索效果，但會使正確的跟錯誤的答案相結合，使得最後結果錯更多，故對於大型資料庫，當要做特徵數量的最佳化時，以高階特徵慢慢減少為佳。而在小型資料庫檢索實驗中，卻發現小型資料庫對於刪除低階特徵反而能提高整體正確度，其原因是小型資料庫，所需要的特徵較少，故每項特徵的重要性更為重要，把不必要的特徵刪除，反而能提高其他特徵的特性。且在小型資料庫中並不一定還要再使用傅立葉描述子跟 Zernike 矩結合，單一特徵種類加上本論文提出的相關大小特徵就能得到最佳的檢索效果，而且少一種特徵更能減少相結合時所需的權重值最佳化此一步驟，故對於小型資料庫特徵能取越少越簡單越好。而未來如果要在針對特徵做最佳化，可以使用特徵選取的技術，使檢索的特徵數更為有效率，而且不只有可以過濾不必要的特徵，更可以增加正確率。以下是三種常用的特徵選取的技術：順序後退選擇法(sequential backward selection, SBS) [24]、順序前進選擇法(sequential forward selection, SFS) [24]和順序浮動選擇法(sequential floating selection)[24]，其中順序後退選擇法適用於有眾多特徵變數要刪減不必要的特徵的時候，而順序前進選擇法則是適合想要以低維度特徵作為代表的時候，但這兩種方法都有一缺點，那就是會有巢效應(nesting effect)：高維度往下降或低維度往上升時，其特徵不一定是同維度下最好的特徵，使得找到的特徵並不一定最佳，而順序浮動選擇法則是結合了順序後退選擇法和順序浮動選擇法，其能大大降低巢效應，找到這三種方法最佳的特徵數量。而更進一步求取最佳特徵數量，可以使用基因演算法，類神經網路等方法尋找最佳解。

第五章 結論與未來發展

本論文提出一 3D 模型檢索系統，其首先藉由主成分分析法找出三主軸，之後針對三個主軸作格子點化並投影作特徵抽取，其中傅立葉描述子跟 Zernike 矩皆為 2D 圖形檢索的特徵抽取方法，而本論文增加一新特徵：三個投影圖的大小相關性，使特徵不再只有 2D 圖形特徵，還包含了 3D 特徵且不需要經過大量計算，其早已經隱藏在原本的兩種特徵中。而且本論文的方法不像其他演算法在做完主成分分析法後，對於旋轉變量無法在作適應，能再靠著也具有旋轉不變性的傅立葉描述子跟 Zernike 矩克服偏軸跟圖像旋轉或顛倒的情形，並靠著最後旋轉投影對應找出最佳解，克服主軸順序顛倒的情形。最後在各個資料庫的實驗中的檢索效果整體效果皆比其他方法為佳，證明本論文在檢索物體時對於不同的物體有者不錯的鑒別力而對於相同的但不是很相似的物體也具備足夠的容忍度，並且在實驗中也證明其對於因部分區域不同所造成的偏軸能有所克服。最後對於所需時間跟空間也相對於 LFD 有極大的下降，且同時能保持高效能並增加對於相似物體的差異能有更高的容忍度，使得本論文在回傳相當多物體時還是能保有很好的正確率，而最後對於不同大小的資料庫，提出不同特徵數的最佳化方法，使整體效率能再加以提昇。

在未來的發展上可分為二個主要方向，尋找更佳的姿態評判標準且維持一定效率，使物體能完全轉為一致的形態，不再產生因相似物體但因為整體分佈不同而有很大的差異。另一方面則是尋找對於區域特徵能有更好的特徵抽取方法，在傅立葉描述子中只考慮了輪廓變化，而 Zernike 矩也只考慮了二值化的投影形狀變化，皆無法考慮圖形的深度變化，使其對於有遮蔽的物體檢索效能不佳。而如果未來的 3D 模型能加上色彩跟質地，則再增加此兩種特徵，將能使檢索效果更大福提昇，畢竟色彩跟質地對於人眼也都是分類物體的重要指標。

參考資料

- [1] J. H. Tangelder and R. C. Veltkamp, “A survey of content based 3D shape retrieval methods,” *Multimedia Tools and Applications*, vol. 39, pp. 441-471, 2008.
- [2] B. Bustos, D. A. Keim, D. Saupe, T. Schreck, and D. V. Vranić, “Feature-based similarity search in 3D object databases,” *ACM Computing Surveys*, vol. 37, no. 4, pp. 345-387, 2005.
- [3] P. Shilane, P. Min, M. Kazhdan, and T. Funkhouser, “The Princeton shape benchmark,” in *Proc. of 6th International Conference on Shape Modeling and Applications*, pp. 167-178, 2004.
- [4] M. Ankerst, G. Kastenmüller, H. P. Kriegel, and T. Seidl, “3D shape histogram for similarity search and classification in spatial databases,” in *Proc. of 6th International Symposium on Spatial Databases*, pp. 207-228, 1999.
- [5] M. Ankerst, G. Kastenmüller, H. P. Kriegel, and T. Seidl, “Nearest neighbor classification in 3D protein databases,” in *Proc. of 7th International Conference on Intelligent Systems for Molecular Biology*, pp. 34-43, 1999.
- [6] R. Osada, T. Funkhouser, B. Chazelle, and D. Dobkin, “Shape distributions,” *ACM Transactions on Graphics*, vol. 21, no. 4, pp. 807-832, 2002.
- [7] D. Y. Chen, M. Ouhyoung, X. P. Tian, and Y. T. Shen, “On visual similarity based 3D model retrieval,” *Computer Graphics Forum*, pp. 223-232, 2003.
- [8] T. Funkhouser, P. Min, M. Kazhdan, J. Chen, A. Halderman, D. Dobkin, and D. Jacobs, “A search engine for 3D models,” *ACM Transactions on Graphics*, vol. 22, no. 1, pp. 83-105, 2003.
- [9] M. Kazhdan, T. Funkhouser, and S. Rusinkiewicz, “Rotation invariant spherical harmonic representation of 3D shape descriptors,” in *Proc. of the Symposium on Geometry Processing*, 2003.
- [10] D. Saupe and D. V. Vranić, “3D model retrieval with spherical harmonics and moments,” in *Proc. of the Symposium on Pattern Recognition*, pp. 392-397, 2001.
- [11] B. Horn, “Extended Gaussian images,” in *Proc. of the IEEE*, vol. 72, no. 12, pp. 1671-1686, 1984.
- [12] I. Kunttu and L. Lepistö, “Shape-based retrieval of industrial surface defects using angular radius Fourier descriptor,” *Image Processing, IET*, vol. 1, no. 2, pp. 231-236, 2007.

- [13] D. S. Zhang and G. Lu, "Study and evaluation of different Fourier methods for image retrieval," *Image and Vision Computing*, vol. 23, pp. 33-49, 2005
- [14] A. El-ghazal, O. Basir, and S. Belkasim, "A new shape signature for Fourier descriptors," in *Proc. of the IEEE International Conference on Image Processing*, pp. 161-164, 2007.
- [15] H. Kauppinen, T. Seppänen, and M. Pietikäinen, "An experimental comparison of autoregressive and Fourier-based descriptors in 2D shape classification," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, no. 2, pp. 201-207, 1995.
- [16] W. Y. Kim and Y. S. Kim, "A region-based shape descriptor using Zernike moments," *Signal Processing Image Communication*, vol. 16, pp. 95-102, 2000.
- [17] D. Zhang and G. Lu, "An integrated approach to shape based image retrieval," in *Proc. of the 5th Asian Conference on Computer Vision*, pp. 652-657, 2002.
- [18] D. V. Vranić and D. Saupe, "3D model retrieval," in *Proc. of Spring Conference on Computer Graphics and its Applications*, pp. 89-93, 2000.
- [19] D. V. Vranić and D. Saupe, "Tools for 3D-object retrieval: Karhunen-Loeve transform and spherical harmonics," in *Proc. of IEEE 4th Workshop on Multimedia Signal Processing*, pp. 293-298, 2001.
- [20] S. T. Wang and Z. L. Shi, "Content-based 3D model retrieval based on the relevance feedback algorithm," *Chung Hua Journal of Science and Engineering*, vol. 2, no. 1, pp. 53-61, 2004.
- [21] R. Jang, "Data clustering and pattern recognition," [Online]. Available: <http://neural.cs.nthu.edu.tw/jang/books/dcpr/> [Accessed: June 5, 2009].
- [22] K. Teknomo, "Similarity measurement," [Online]. Available: <http://people.revoledu.com/kardi/tutorial/Similarity/> [Accessed: June 5, 2009].
- [23] J. Shutler "Statistical moments," [Online]. Available: http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/SHUTLER3/CVonline_moments.html [Accessed: June 5, 2009].
- [24] S. Theodoridis and K. Koutroumbas, *Pattern recognition*. San Diego: Academic press, 1999.