

# 行政院國家科學委員會專題研究計畫 成果報告

## 基礎於相似測量並利用自我組織映射圖與支撐向量法之三 維物件辨識研究

計畫類別：個別型計畫

計畫編號：NSC94-2213-E-216-024-

執行期間：94年08月01日至95年07月31日

執行單位：中華大學資訊管理學系

計畫主持人：王偉德

計畫參與人員：劉蓉、賴政揚

報告類型：精簡報告

報告附件：出席國際會議研究心得報告及發表論文

處理方式：本計畫可公開查詢

中 華 民 國 95 年 10 月 26 日

# 基礎於相似測量並利用自我組織映射圖與支撐向量法之 三維物件辨識研究

## 3D Object Recognition Based on Similarity Measures Using Self-Organizing Map and Support Vector Machine

計畫編號：NSC 94 - 2213 - E - 216 - 024 -

執行期間：2005/8/1 ~ 2006/7/31

主持人：王偉德 中華大學助理教授

### 一、中英文摘要

從文獻中，我們得知許多學者在三維物件辨識研究有不同的研究方向。不過，學術界最近重新在相似性技術，無論在新模式發展與應用上，顯露出濃厚興趣。在2003年貴會補助下，我們發展出一套相似測量法供二維形狀影像擷取之用。這相似測量值能符合視覺感受必備的不變性條件，如平移、旋轉、放大與縮小。在已有的基礎下，又在學術界新重視此議題上，這驅使我們當初提出本計劃，並嘗試藉着我們的相對測量值供三維形狀物件辨識。但在實驗過程作比較測試時，我們發現以 Complex 及 Centroid 傅立葉描述子所得之結果更好，其中以 Complex 傅立葉描述子所得之成果最為出色，故此我們便更換特徵方法，繼續研究，最後之實驗結果為對 Sebastian 之二維影像、及 Cyr 和 Kimia 之三維形狀影像資料庫獲得平均為 84% 及 98% 之辨認率。至於 SOM 方面，目前只能成功地運用於二維影像檢索，加強檢索之效率。

So far as we know, there have already been a wide variety of approaches that propose different solutions to recognize 3D object. However, there is a renewed interest in similarity-based techniques recently, both in developing and studying new effective paradigms. Due to the refocusing on the similarity-based pattern recognition in academic, and our work of similarity measures, which can satisfy necessary requirements of visual perception such as translation, rotation, scaling and minor

shearing invariance, under the support NSC in 2001, we were attracted to propose this project to address this concerned research issue in 2004. However, when we performed comparison testing with the Complex and Centroid Fourier Descriptor, we discovered that such two methods were better than us. Therefore, we had changed our feature extraction method and continued the research. Finally, we obtain 85% and 98% accuracy for the recognition test using Cyr and Kimia 3D shape image database, respectively. On the hand, we have successfully applied the SOM approach on 2D image retrieval in improve the efficiency.

### 二、緣由與目的

#### 1. 緣由

隨著科技的進步與網際網路的發達，數位多媒體迅速的普及化，使得人們可以不受限制隨時隨地的獲得數位影像資訊，然而，相對的延伸出另一問題，要如何以有效的方式快速且正確找尋出符合需求的影像資料，影像檢索的技術也就變得越來越重要了，由於在數位時代的影像資料之多，無法用人工方式一個一個比對，為了要讓檢索的效果達到最佳，我們的研究採用以物件形狀為基礎的影像檢索方法。

#### 2. 目的

本研究期望能分別以 Cyr、Kimia[1] 所提供之影像資料庫做為測試基礎，經由特徵點擷取的演算法擷取影像圖形的特

徵點，透過 SVM 分類器[2]來解決多分類問題之結果，檢索出最適當之影像，並用最相似影像 (the best match unit, BMU) 與 SOM[3]快速有效的檢索出最相似之影像物件樣式。

### 三、研究方法及成果

#### 1. 研究方法

##### (a) 影像檢索特徵萃取方法

影像特徵值擷取的方法有許多種，在過去的研究成果也相當豐富，包含：傅立葉描述子 (Fourier Descriptor)、Moment Invariants、以及我們曾經提出的相似形狀特徵值[4]等等。在傅立葉描述子方面我們採用了 Centroid 及 Complex 兩種方式[5]，在 Moment Invariants 方面我們採用了 Chen's Moments [6]。我們首先取得圖形輪廓上的像素點座標，並同時計算出圖形重心的座標位置。取得了圖形輪廓上每一個像素點的座標位置與重心座標之後，再以三種不同的特徵點擷取方式 Centroid Fourier Descriptor、Complex Fourier Descriptor 與 Moment Invariants、我們所研發的 Similarity Measure 來擷取影像圖形之特徵點。

##### (b) 分類方法

在圖形識別的領域上，分類是相當重要的一個步驟，將影像進行分類之後，可以減少檢索的範圍，加快檢索的速度與有效性。現行研究的分類方法相當的多，每一個都有其不同的特性，但什麼樣類型的資料適合用哪一個分類方法，到目前並未有一定的答案。我們採用近年來受矚目的 SVM 分類技術，並選用台大林智仁博士之 LIBSVM[7]作為多類別分類器。

##### (c) 輔助檢索方法

當樣本經過 SVM 訓練後，會產生一 model 供分類用，雖然 SVM 會提供每一測試樣本一個預測之類別，但若類別中的原始影像資料過多時，需要一一與測試樣本比對亦屬耗時，故此我們預先對每一類

別使用 SOM 產生一個同類別的原始影像鄰近地圖供快速檢索用，當一個測試樣本經過分類得出預測類別後，我們對該類的每一個原始影像的傅立葉描述子數值求出與樣本之相對數值的最小差異的原始影像，爾後當使用者需要檢索更多之原始影像時，我們則從 SOM 地圖直接獲得，以加快檢索速度。

#### 2. 成果

##### (a) 實驗流程設計

2D 影像資料研究如圖一實驗流程圖來進行實驗，其各階段分別簡述如下：

##### (i) 特徵擷取階段：

將所有的 2D 影像資料經由 CeFD 的轉換取得影像特徵值。並分別將每個影像類別之資料各別分成兩個部份—20%作為測試資料和 80%作為訓練資料。

##### (ii) 學習階段：

在此階段中，我們主要先分別將 20% 測試資料和 80% 訓練資料進行正規化 (normalization) 的動作，將所有特徵值之範圍正規化於+1~-1 區間。正規化資料的主要目的希望能減少計算上的複雜度。經由正規化後，並依據交叉驗證之實驗得出之最佳參數值後，將 80%的訓練資料透過 SVM 分類器進行訓練，並獲得訓練後之學習模組 (learning model)。

##### (iii) 影像檢索階段：

SVM 檢索部份，此階段主要將測試資料依據 SVM 學習階段所產生的學習模組來進行預測的動作，並將預測所得之結果類別資訊記錄於資料庫中，以方便日後檢索。SOM 檢索部份，首先我們選擇最佳檢索之影像類別，並於檢索類別中套用最小平方差測量方式產生 SOM 地圖表，方便了解其檢索所對應之相似影像。

3D 影像資料研究如圖二驗流程圖來進行實驗，其各階段分別簡述如下：

##### (i) 前置處理：

取得以視覺為基礎的影像，我們依立體空間投影 3D 影像物件產生之 2D 影像投影，依此原理，我們使用 Cyr 和 Kimia

所提供之 3D 影像資料庫。

(ii) 特徵擷取階段：

將所有的 2D 影像投影資料經由 CeFD 與 CoFD 傅立葉轉換進行轉換，並取得影像特徵值。分別將每個影像類別之資料各別分成兩個部份—20%作為測試資料和 80%作為訓練資料。

(iii) 學習階段：

在此階段中，將 80%的訓練資料依據交叉驗證之實驗得出之最佳參數值後，透

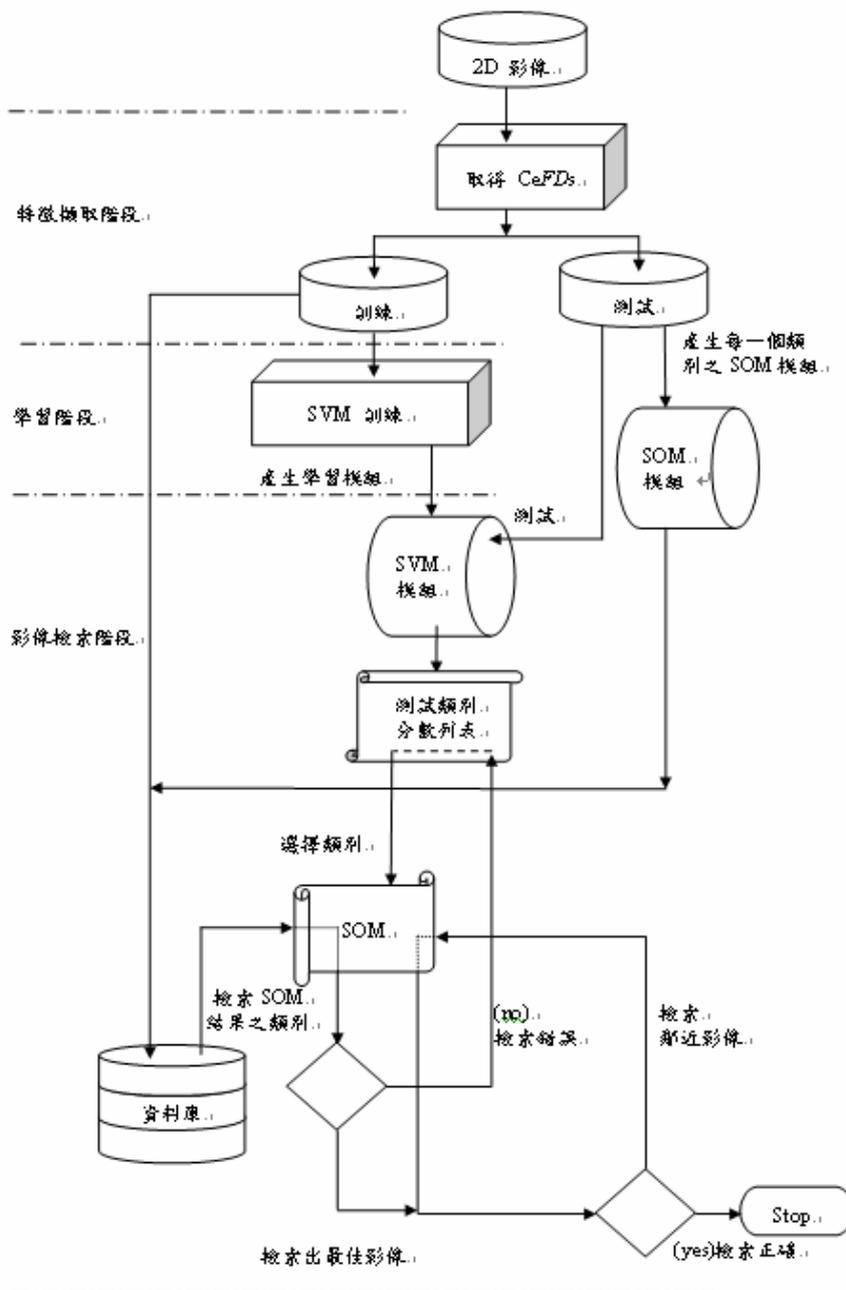
過 SVM 分類器進行訓練，並獲得訓練後之學習模組 (learning model)。

(iv) 預測階段：

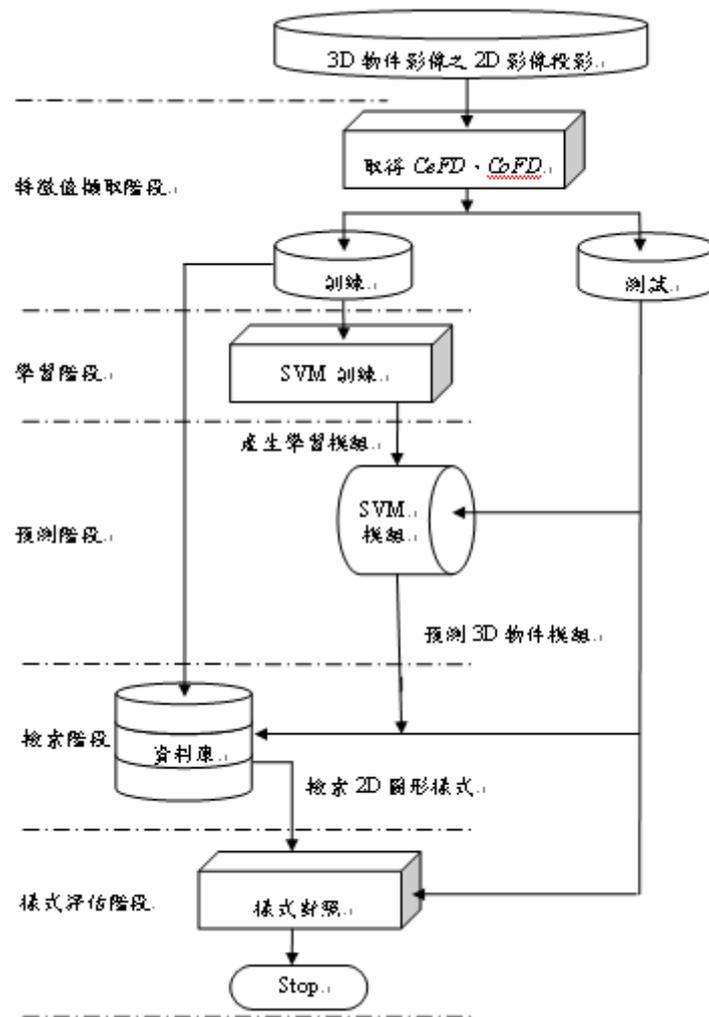
將 20%測試資料依據學習階段所產生之 SVM 模組進行預測動作，並將結果存於資料庫中。

(v) 影像檢索階段：

此階段主要是依據預計所得之類別與原先資料庫中之同一類別做檢索，找出最相似之影像。



圖一 SVM 與 SOM 應用於 2D 影像檢索實驗流程圖



圖二 SVM 應用於 3D 影像物件檢索實驗流程圖

(b) 實驗結果

2D 影像資料實驗過程與結果簡述如下：

依實驗之經驗法則，我們採用 LIBSVM 工具，並使用 RBF 函數，因為它在解決多重分類的效能方面有較佳的表現。經由交叉驗證，我們訂定了相關的參數範圍  $\delta^2 \in \{5, 2, 1, 0.5, 0.1\}$  和  $C \in \{1000, 500, 100, 50, 10, 1\}$  來進行實驗，結果如表一和表二所示。

為了確保實驗的公正性，我們將資料拆成五等份並以循環方式去做測試，並和 Sebastian's 與 Chen's 的方法做比較，經由結果得知，如表三與表四，我們所使用的方法和 Sebastian's 的差不多；而與

Chen's 方法比較，則在辨識率上皆有較佳之表現。

在 SOM 方面，我們對同一類別影像物件中使用 LMSED 方法來分別找出每個影像之最相似影像 (the best match unit, BMU)，LMSED 對每個影像 Q 的計算方法如以下公式所示，m 為傅立葉之特徵值數；M 為同一類別中之另一影像物件。

$$LMSED_Q = \text{Min} \left( \sqrt{\sum_{i=1}^m (CeFD_{Q_i} - CeFD_{M_i})^2} \right)$$

對於訓練資料而言，針對我們調整參數讓每個影像之 BMU 皆於在 5x5 之 SOM 鄰近格中，並將此結果存於資料庫，方便

日後快速查詢檢索。

經由 SVM 得知影像所屬之類別後，後續動作則透過 SOM 方法再加以檢索此類別之相關影像來驗證檢索結果是否一致。如圖三所示，類別 textbox 經由 SOM 訓練分群之結果與表五得知自我類別中

的最相似影像 (the best match unit, BMU) 一樣。由各項實驗數結果可知，本研究所提出的方法在利用支持向量機進行分類時能取得較高的辨識率，並且可快速的透過 SOM 方法得知其最相似之影像。

表一 SVM 訓練與測試之結果

$C \setminus \delta^2$	5		2		1		0.5		0.1	
	Train	Test	Train	Train	Train	Test	Train	Test	Train	Test
1	70.48	79.62	66.36	74.41	62.50	68.25	54.39	60.19	29.39	33.16
10	73.80	81.04	74.07	81.99	75.27	82.46	73.01	81.52	64.10	67.77
50	73.27	79.15	74.34	80.09	75.40	81.52	76.33	82.46	73.27	79.15
100	73.40	79.15	74.20	80.09	74.34	81.04	76.86	82.46	74.34	81.52
500	73.40	79.15	73.94	79.62	74.73	80.09	76.60	80.57	76.73	80.09
1000	73.40	79.15	73.94	79.62	74.73	80.09	76.60	81.04	76.06	81.99
Avg	72.96	79.54	72.81	79.30	72.83	78.91	72.30	77.88	65.65	70.61

表二 SVM 訓練與測試之結果

類別數	CeFD			Chen's 4 不變量矩		
	$\delta^2$	C	辨識率	$\delta^2$	C	辨識率
8	1	1000	100% -+ 0%	1	500	85.0% -+ 5.3%
10	1	1000	100% -+ 0%	1	500	85.4% -+ 5.3%
12	1	1000	100% -+ 0%	0.1	1000	92.9% -+ 3.9%
14	0.5	1000	100% -+ 0%	0.1	1000	92.2% -+ 4.0%
16	0.1	1000	98.7% -+ 1.2%	0.1	500	90.8% -+ 3.0%
18	1	1000	97.7% -+ 1.6%	0.5	1000	88.4% -+ 3.4%
20	1	1000	96.2% -+ 2.0%	1	250	84.0% -+ 3.9%
22	1	1000	93.1% -+ 2.7%	1	250	84.5% -+ 3.8%
24	1	1000	88.9% -+ 2.3%	1	250	84.1% -+ 2.7%
26	1	1000	88.6% -+ 2.4%	0.5	750	86.3% -+ 2.6%
28	0.5	50	88.7% -+ 2.3%	0.5	750	81.3% -+ 2.9%
30	0.5	1000	87.0% -+ 2.4%	0.5	1000	80.3% -+ 2.9%
32	0.5	10	86.5% -+ 2.5%	0.5	1000	80.6% -+ 2.3%
34	0.5	100	85.4% -+ 1.9%	0.5	1000	76.4% -+ 2.4%
36	0.5	10	84.0% -+ 1.9%	0.5	1000	72.3% -+ 2.4%
38	0.5	100	82.2% -+ 2.0%	0.5	1000	67.8% -+ 2.5%
40	0.5	100	82.5% -+ 1.9%	1	1000	66.4% -+ 2.5%

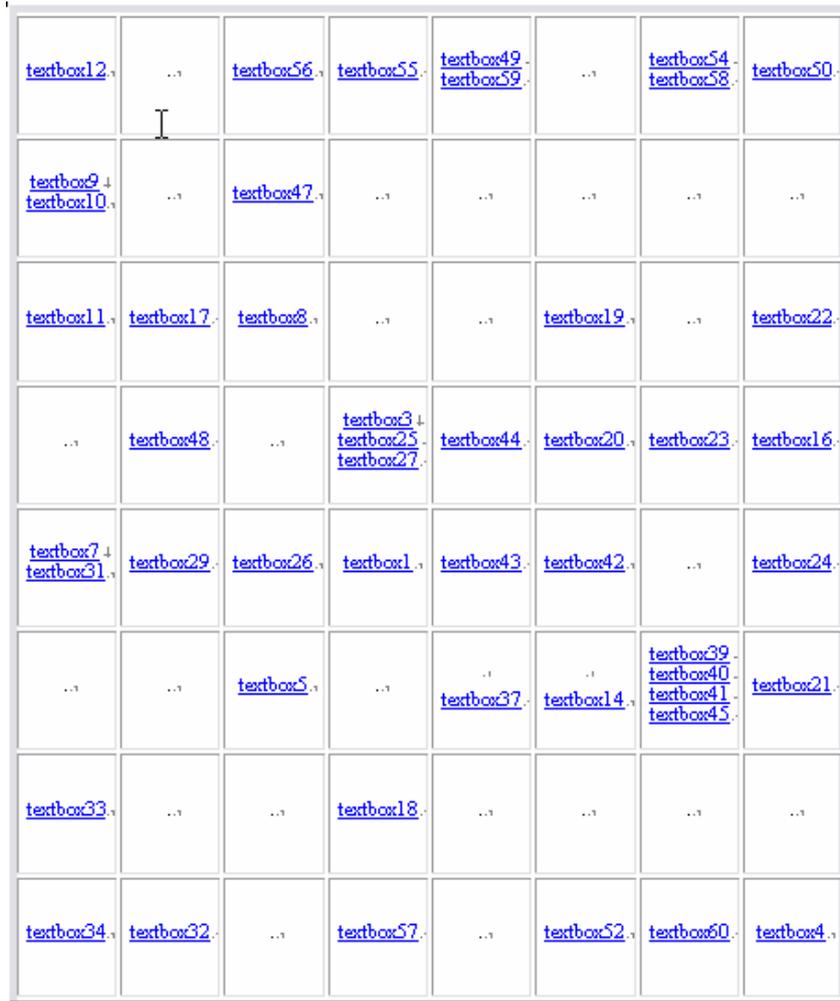
表三 不同方法之預測結果

Database		99 shapes	216 shapes	1045 shapes
第一次比對 (%)	Ours	90.0	96.7	84.1
	*Sebastian's	97.0	96.0	--
	Chen's	77.8	77.8	67.8
第二次比對 (%)	Ours	96.7	98.9	94.2
	*Sebastian's	--	--	--
	Chen's	77.8	83.3	81.0
第三次比對 (%)	Ours	96.7	100	96.3
	*Sebastian's	100	100	--
	Chen's	77.8	91.7	83.9

註：Sebastian's 的方法於第二次檢索比較與 1045 影像並無相關資料紀錄

表四 五次循環之預測與測試結果

Database	循環	第一次比對 (%)		第二次比對 (%)		第三次比對 (%)	
		Train	Test	Train	Test	Train	Test
99 shapes	Round1	100	94.44	100	94.44	100	94.44
	Round2	100	83.33	100	88.89	100	88.89
	Round3	100	94.44	100	100	100	100
	Round4	100	83.33	100	100	100	100
	Round5	100	94.44	100	100	100	100
	Average	100	90.00	100	96.67	100	96.67
216 shapes	Round1	100	100	100	100	100	100
	Round2	100	97.22	100	100	100	100
	Round3	100	97.22	100	100	100	100
	Round4	100	94.44	100	97.22	100	100
	Round5	100	94.44	100	97.22	100	100
	Average	100	96.66	100	98.89	100	100.00
1045 shapes	Round1	96.28	84.91	99.76	93.87	100	95.75
	Round2	96.88	83.49	99.88	94.34	100	96.70
	Round3	96.88	83.49	99.88	94.34	100	96.70
	Round4	96.28	84.91	99.76	93.87	100	95.75
	Round5	96.88	83.49	99.88	94.34	100	96.70
	Average	96.64	84.06	99.83	94.15	100	96.32



圖三 textbox 類別經由 SOM 訓練之結果

表五 textbox 類別之 BMU 結果

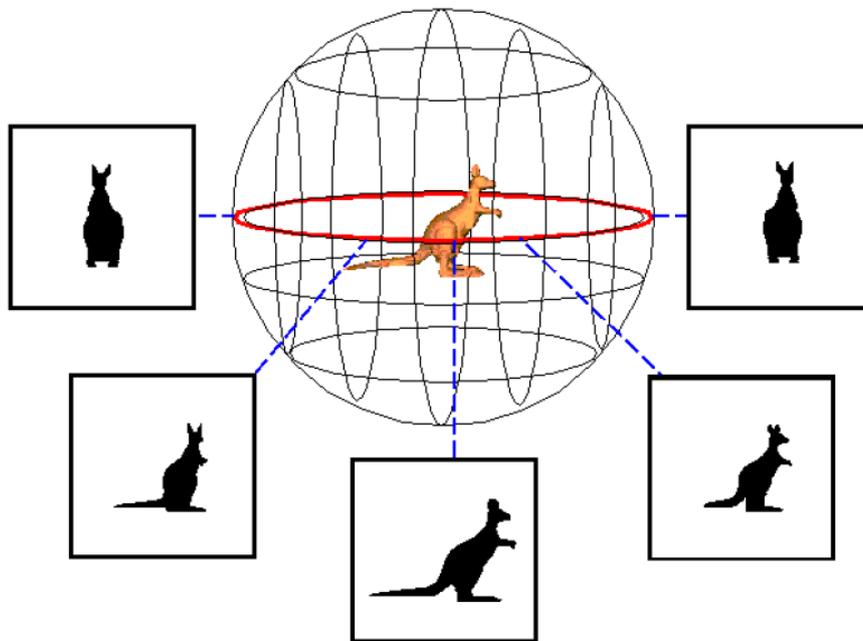
textbox					
Image	BMU	Image	BMU	Image	BMU
textbox1	textbox25	textbox21	textbox45	textbox42	textbox43
textbox3	textbox27	textbox22	textbox16	textbox43	textbox42
textbox4	textbox41	textbox23	textbox20	textbox44	textbox43
textbox5	textbox29	textbox24	textbox41	textbox45	textbox41
textbox7	textbox31	textbox25	textbox27	textbox47	textbox48
textbox8	textbox29	textbox26	textbox5	textbox48	textbox8
textbox9	textbox10	textbox27	textbox3	textbox49	textbox59
textbox10	textbox9	textbox29	textbox5	textbox50	textbox54
textbox11	textbox9	textbox31	textbox7	textbox52	textbox57
textbox12	textbox10	textbox32	textbox34	textbox54	textbox58
textbox14	textbox13	textbox33	textbox34	textbox55	textbox49
textbox16	textbox23	textbox34	textbox33	textbox56	textbox55
textbox17	textbox11	textbox37	textbox43	textbox57	textbox52
textbox18	textbox37	textbox39	textbox41	textbox58	textbox54
textbox19	textbox44	textbox40	textbox45	textbox59	textbox49
textbox20	textbox23	textbox41	textbox45	textbox60	textbox52

3D 影像資料實驗過程與結果簡述如下：

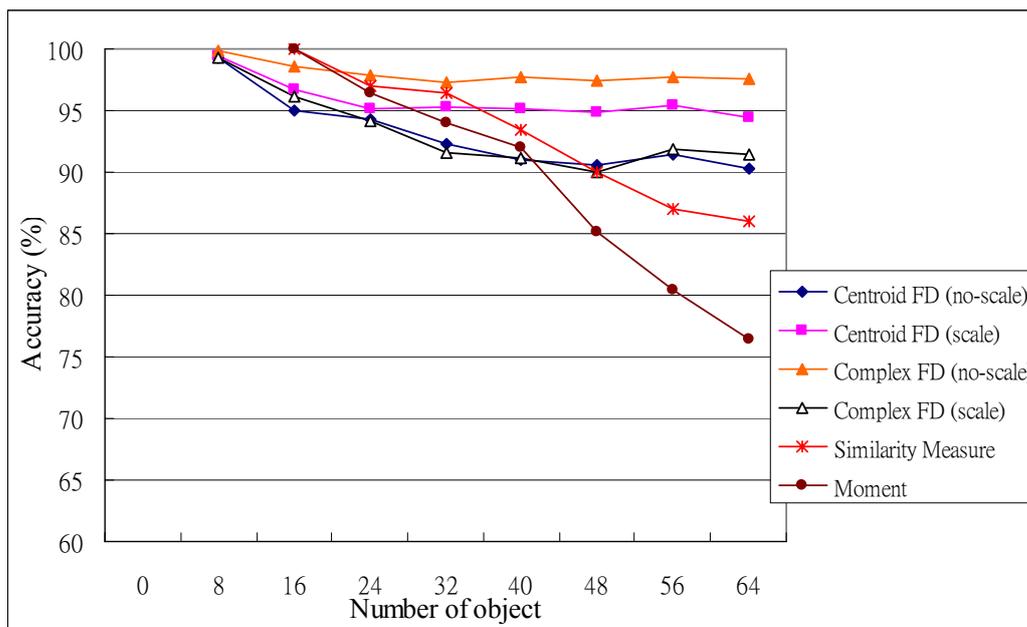
本實驗 3D 影像檢索主要是使用 Cyr 和 Kimia 所提供之 3D 影像資料庫，此 3D 資料集主要是以視覺為基礎依立體空間投影 3D 影像物件產生之 2D 影像，以水平角度每隔五度做為一空間區段進行投影而成，如圖四，影像立體物件 kangaroo 以每 5 度為一單位，進而投影產生 2D 影像。故一個完整的 3D 影像物件可拍攝出 72 張 2D 影像。此影像資料庫總共包含了六十四個影像物件。實驗使用 LIBSVM，主要核心函數為 RBF，因為此函數能針對多重類別資料分類獲得較好的結果。我們

透過 ten-fold 的交叉驗證，依所訂定相關的參數範圍和來針對不同的影像特徵擷取方法進行實驗，以取得其最佳參數與辨識率，如圖五所示。在數種影像特徵方法辨識率之比較圖，可輕易看出 Complex FD (no-scale) 於 64 個物件類別獲得較高之辨識率，因此，本研究採用 Complex FD 特徵擷取方法來做為實驗依據。

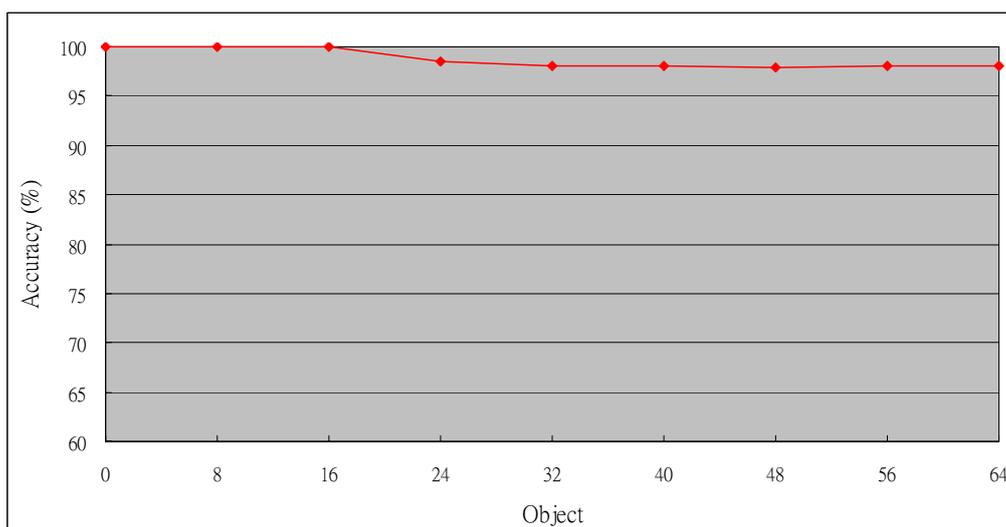
為確保實驗的公正性，我們分別以五次的循環方式，分別由 0 個物件類別增加至 64 個物件類別進行檢索預測與測試，並獲得平均高達 98% 檢索辨識之辨識率。其結果如圖六所示。



圖四 3D 影像物件之 2D 影像投影



圖五 3D 影像類別使用不同影像特徵擷取方法之預測結果



圖六 3D 影像由 0 個物件類別增加至 64 個物件類別之預測結果

#### 四、討論

在本研究中，我們比較了不同的特徵擷取方法來擷取影像圖形的特徵點，經由前置資料處理後，利用分類器 SVM 來加以訓練並快速且正確的檢索影像。至於執行效率上之比較，2D 部份約快比前人 Sebastian 快上 150 倍左右；3D 部份效率約比 Cyr 和 Kimia 快上 30 倍左右。由此可知，不管是在影像物件識別或是特徵值擷取上，我們的方法在正確性與效率上皆有較好的表現。

#### 五、計畫自評

在研究的過程中，實驗結果大致符合我們預期的目標，藉由不同特徵擷取方法並運用分類器，能夠有效且快速的檢索出正確影像資料，並預測其樣式。但在研究的過程中，仍有一些可以改進的空間，可供後續進一步研究的改善與擴展之空間。在 SOM 方面，由於執行時間之限制，未能在 3D 方面進行實驗，本人將會繼續進行研究。在成果方面，目前就夠完整之 2D 方面已寫成文章，題目為 — Shape-based image retrieval using support

vector machines, Fourier descriptors and self-organizing maps, 並載明國科會計劃編號, 投稿至 Information Sciences (SCI) 期刊, 目前正進行第三次 revised, 至於 3D 部份尚需實驗補齊並整理完整後再完稿投至 SCI 國際期刊。

#### 六、參考文獻

- [1] C.M. Cyr and B.B. Kimia, "A similarity-based aspect graph approach to 3D object recognition," International Journal of Computer Vision, vol.57, no.1, pp.5-22, 2004.
- [2] V. N. Vapnik, "The Nature of Statistical Learning Theory," Springer-Verlag, New York, 1995.
- [3] T. Kohonen, "Self-organization and Associative Memory," 2nd ed. Berlin: Springer-Verlag, 1988.
- [4] W.-T. Wong and S.-H. Hsu, "Shape-based image retrieval using support vector classification," Proceedings of the ninth IAPR conference on machine vision applications (MVA2005), Tsukuba Science City, Japan, pp. 64-67, 2005.
- [5] 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.
- [6] C.-C. Chen, "Improved moment invariants for shape discrimination," Pattern Recognition vol. 26, pp. 683-686, 1993.
- [7] C. Chang, J. Lin, "LIBSVM: a library for support vector machines," <http://www.csie.ntu.edu.tw/~cjlin/libsvm>, 2003.