

行政院國家科學委員會補助
大專學生參與專題研究計畫研究成果報告

* *****
* 計 畫
* : 3D 模型檢索
* 名 稱
* *****

執行計畫學生： 陳詠倫
學生計畫編號： NSC 99-2815-C-216-008-E
研究期間： 99年07月01日至100年02月28日止，計8個月
指導教授： 石昭玲

處理方式： 本計畫可公開查詢

執行單位： 中華大學資訊工程學系

中華民國 100年03月10日

(一) 摘要

隨著電腦技術的快術發展、電腦運算效能大幅提升，人類在電腦螢幕的世界裡不再滿足於圖片、影像為主所構成的二維平面世界，所以積極追求由 3D 模型所構成的虛擬世界。隨著製作 3D 的軟硬體操作越來越容易，使用者可以方便又快速的製作出 3D 模型。由於 3D 影片風潮，有大量的 3D 模型應用在生活中(如圖一)。而且也隨著網際網路的普及，能在網際網路上輕易的散播和取得 3D 模型，以致於要如何在龐大的 3D 模型資料庫裡面搜尋所需的 3D 模型，變成現在一個重要的研究方向。在過去，傳統的關鍵字搜尋方式不僅不直覺，也隨著個人主觀不同而分類不同，往往達不到自己所需的 3D 模型。而且只用一種特徵擷取的方法是無法對所有的模型有很好的檢索效果，希望能組合成多種特徵以達到更好的檢索效果，且試著對每個查詢模型，找到其最適合的特徵組合，進而研發出一個搜尋更完善的 3D 模型搜尋引擎。

在本計畫中，利用深度投影、曲率投影及角度投影三種不同投影方式之下進行 3D 模型的特徵擷取。接著再進行 3D 模型的相似度的比對，以搜尋出相似的 3D 模型。



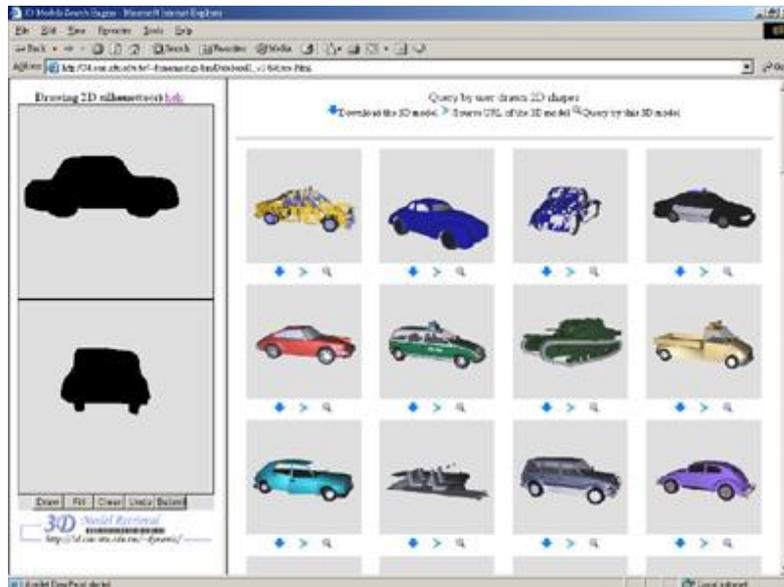
圖一 3D 電影-食破天驚[1]

(二) 研究動機與研究問題

由於電腦技術發展的日新月異，人類在電腦螢幕的世界裡不再滿足於圖片、影像為主所構成的二維平面世界，也越來越想在電腦上模擬真實世界的物體，而積極的追求由 3D 模型所構成的虛擬世界，所以漸漸的趨向使用 3D 的立體影像。而現今 3D 模型的工具和掃描技術的大幅提升，使 3D 模型的製做變得相當容易，因此使用者可以更方便的製作出一個 3D 模型。因為製作 3D 趨於簡單化和網路的快速發展，增加了資料的流通與獲取的管道，所以要取得 3D 的模型也變得更為容易。因此面對現今數量龐大的 3D 模型資料庫，在搜尋上將產生相當大的問題，要如何從許多的 3D 模型中去搜尋一個符合使用者需求的 3D 模型，成為現在一個重要的研究目標。在 3D 模型使用越來越普及的同時，3D 模型的搜尋引擎也將會扮演一個重要的角色。

一般傳統使用的搜尋方法都是藉由輸入關鍵字去找出資料庫中的 3D 模型。但是傳統的關鍵字搜尋方式不僅不直覺，且隨著每個人主觀的分類不同，往往不能滿足我們在 3D 模型萃取上的需求。倘若檔名是沒有意義的字母或數字組成來表示(例如:A1234.jpg)、或是一般的檔名(例如:GASBY.gif 或 arc.bmp)，導致這些關鍵字都不具有足以象徵表示該檔案的資訊時，那麼所擷取的關鍵字將沒有任何代表性。近來 MPEG-7 [2]提供了一種描述多媒體資料內容的標準，經由 MPEG-7 格式的定義，使得多媒體資料得以更加有效地被產生、儲存、交換和傳輸。

近年來，國內外已經有很多學術單位投入 3D 模型搜尋的研究，而且在網路上也有好幾個系統提供人使用，例如台灣大學的 3D Model Retrieval System[4](如圖二)、日本文部科學省多媒體教育研究所的 Ogden VI rxw[5]、希臘資訊電信協會的 3D Search[6]、普林斯頓大學的 3D Models Search Engine[7]。以上這些幾乎都是使用關鍵字搜尋(searching by keywords)、手繪 2D 形狀搜尋(searching by sketch)，和以範例做搜尋(searching by example)。由於關鍵字搜尋效果不佳，以及手繪 2D 形狀搜尋會因為手繪精準度有差異，導致搜尋效果也有限，所以我們利用多種不同的投影特徵，對 3D 模型做特徵擷取，在進行相似度的比對，希望使得 3D 模型的搜尋更為完善。



圖二 台灣大學的 3D Model Retrieval System[4]。

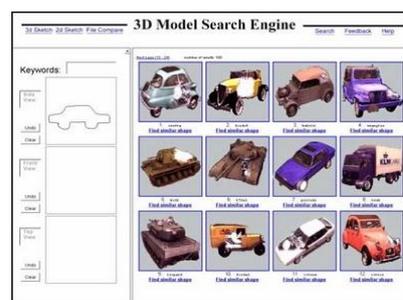
(三) 文獻回顧與探討

下面將對於國內外有關 3D 模型搜尋的相關資料作一些簡單介紹。目前現有的 3D 搜尋的方法大多是藉由擷取適合的特徵來描述 3D 模型的形狀，以便於找出類似的 3D 模型，大約分成利用剪影圖形、2D 輪廓剪影和光影描述。

以 T. Funkhouser [8]提出之 3D 模型搜尋引擎舉例來說，使用者可以畫出一個 3D 形狀去 3D 的資料庫做搜尋(如圖三(a))，或使用者可以利用前視圖、側視圖和俯視圖描繪出 3D 形狀，再利用這三張影像去做搜尋(如圖三(b))，也可以用關鍵字做搜尋，這類的方法大多是使用剪影圖形來比對，並且把比對結果最接近前幾名的 3D 模型回應給搜尋者。



(a)

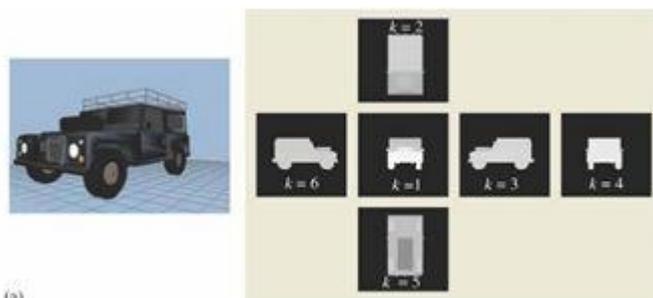


(b)

圖三 3D 模型 尋引擎[8]。(a) 使用 3D 繪圖之檢索結果。(b) 使用 2D 繪圖之檢索結果。

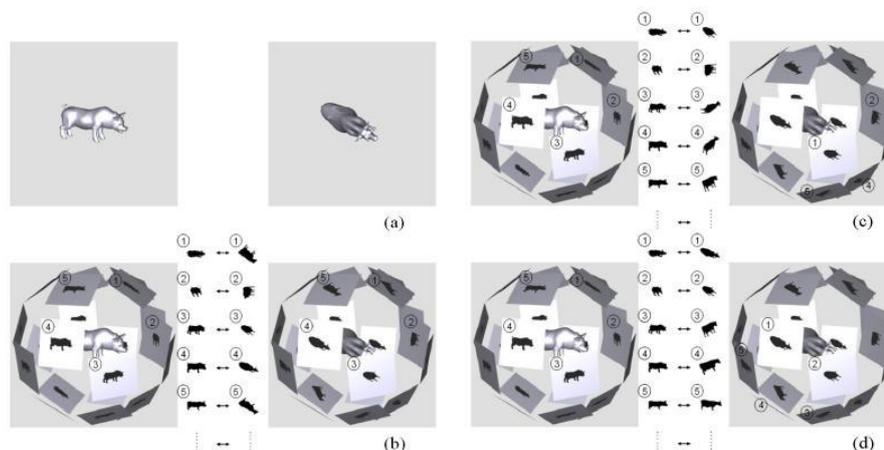
Shih et al. [9] 針對 3D 模型的檢索提出了一個六立面體的方法[9]。六立面體這個特徵對於 3D 模型的小程度的變形是不會受其影響而產生變化的。首先，一個 3D 模型從“前、後、左、右、上、下”六個不同的視角，產生六個不同的投影灰階平面(如圖四)，得到的六個灰階平面，我們稱之為立面圖。擷取每張立面圖的灰階剪影圖，並且以灰階值的深淺代表模型在該立面圖的高度分佈資

訊，其中越靠近觀察者的灰階值越高，反之則越低。接著將每一個立面利用同心圓擷取特徵。藉由每個立面上的兩兩相鄰的同心圓上，所得到高度資訊總和值的差異即為該立面的特徵。利用各個立面的特徵，就可以算出兩個 3D 模型的相似程度。



圖四 3D 模型的六個投影立面圖及其特徵[9]。

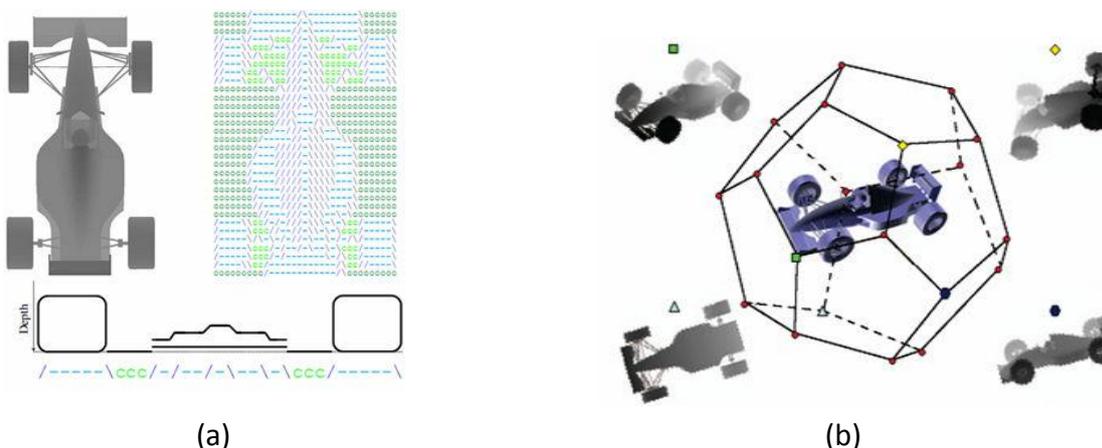
Chen et al. [10] 在三維模型的調準與檢索系統 (A 3D Model Alignment and Retrieval System) [11] 中介紹一種光影描述(lightfield descriptor)來搜尋 3D 模型。光影描述特徵的計算是用 N 個不同角度觀察獲得的 N 個剪影，每一張剪影表示一種 2D 的黑白影像，並且使用 Zernike moments 和 Fourier descriptors 描述每張黑白影像。也提到過，用不同的角度觀看 3D 模型，且擷取該角度的 2D 剪影，然後將剪影使用 2D 圖形特徵擷取的方式去計算特徵向量。以圖五來說，是用 20 個不同的角度觀看 3D 模型的 2D 剪影，再把這些 2D 剪影以 ART(Angular Radial Transform) 擷取 2D 形狀特徵來當作這個 3D 模型的特徵向量。



圖五 以 20 個不同角度去觀看 3D 豬模型的 2D 剪影[10]。

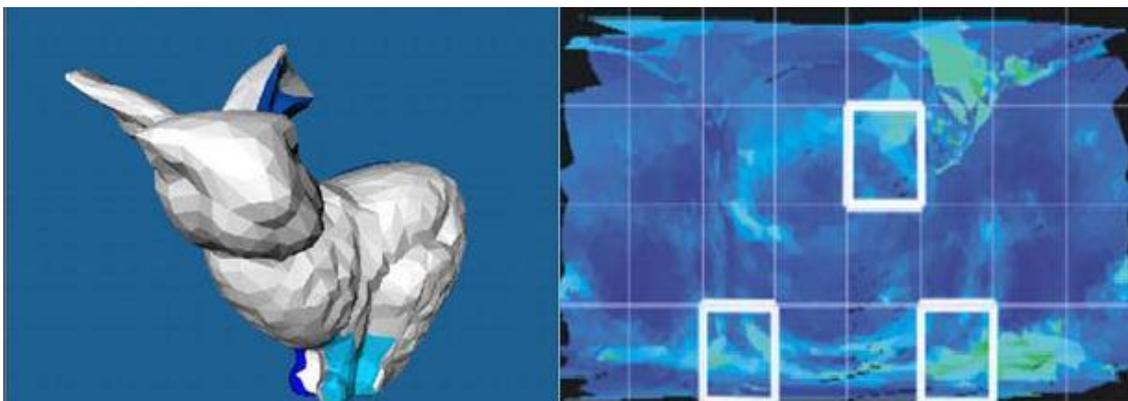
Mohamed Chaouch[12]等人提出一個 3D 模型檢索系統，這套系統主要的方法是藉由 3D 模型投影到一個正十二面體的二十個頂點方向如圖六(a)所示，所得到的二十張深度影像，再對這些深度影像作特徵擷取的動作。對於每張深度影像，根據影像大小($N*N$)，分別在水平以及垂直方向上劃出 N 條深度線，即每張影像可劃出 $2*N$ 條深度線如圖六(b)之上半部所示。對於每條深度線，使用 5 種符號分別是 o, c, /, -, \, 依序代表外部背景點、內部背景點、遞增點、水平點、遞減點，對此條線做序列編碼的動作如圖六(b)之下半部所示。利用深度線來描述 3D 形狀的邊界輪

廓比起其他使用 2D 形狀表示法來的更為精確。最後再將編碼完的序列，使用動態規畫演算法計算出兩個序列的相似度。



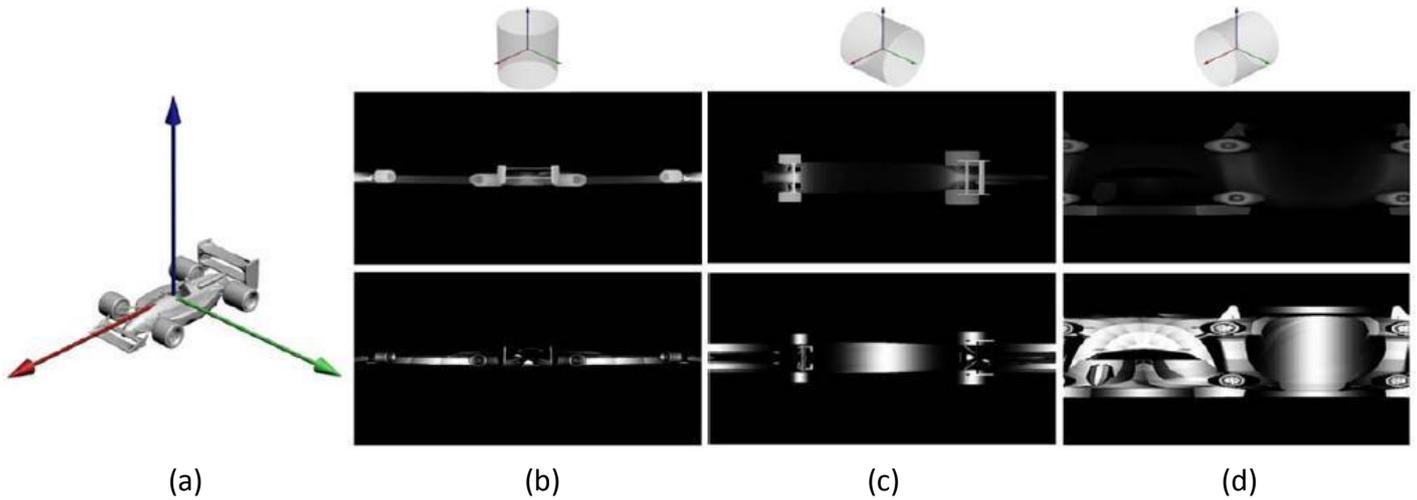
圖六 (a) 從不同的角度所投影出來的深度影像, (b)賽車影像的深度線[12]

Assfalg[14]等人提出利用曲率投影進行 3D 模型檢索的方法，主要的做法為先對每個 3D 模型上的三角片做平滑處理如圖七所示。對於 3D 模型(如圖七(a))表面上的每個點，根據該點周圍的三角片計算出該點的平均法向量，再利用每個點的法向量即可求出該模型表面上的曲率。最後，再將 3D 模型做變形處理以及曲率投影的動作。曲率投影所得到的影像是由相同大小的區塊所構成的一張棋盤式影像如圖七(b)所示，利用直方圖即可統計出整張影像的曲率分布資訊。



圖七 (a)兔子的 3D 模型 (b)由圖七(a)曲率投影所得到的影像以及三個最相近的區塊[13]

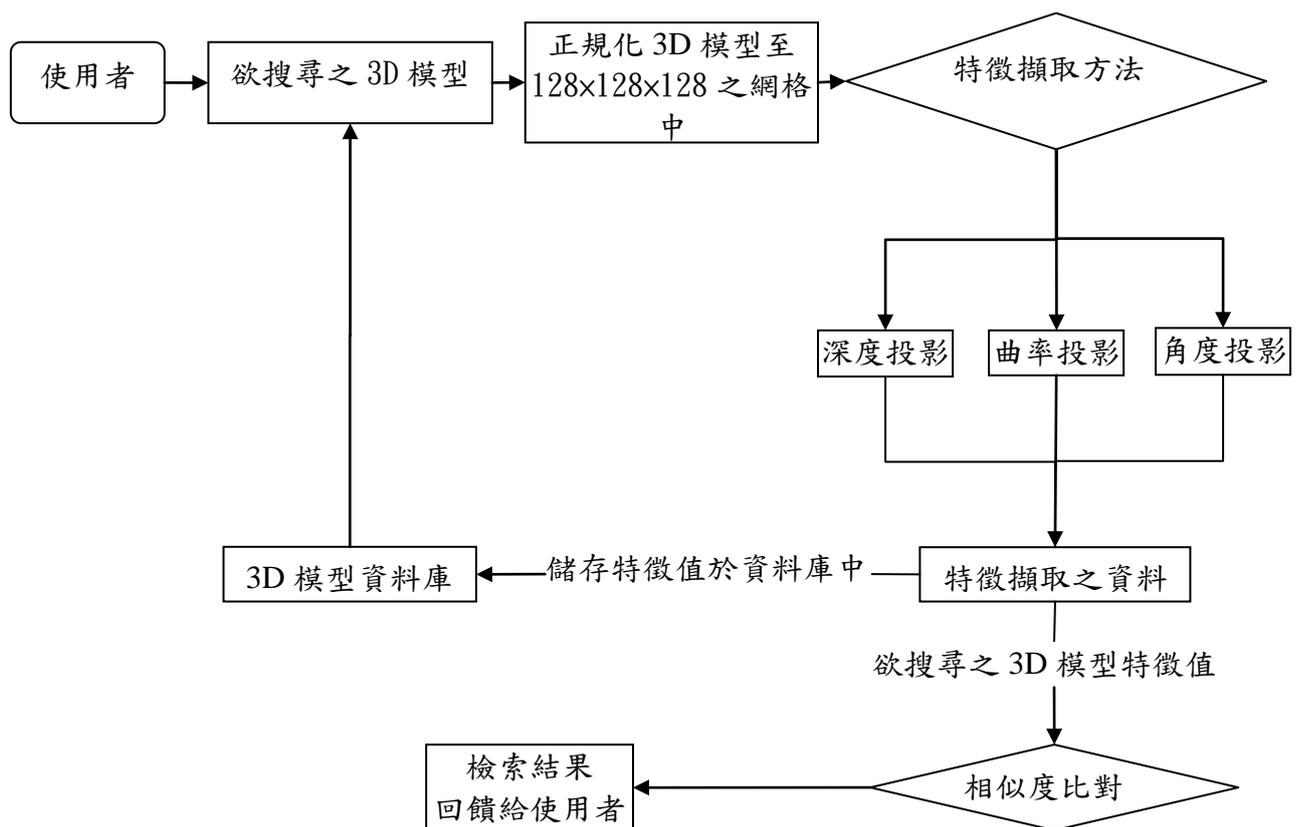
Panagiotis Papadakis et al.[14] 提出一個基於圓柱投影進行 3D 模型檢索的方法。主要作法是以 CPCA 和 NPCA 對模型做擺正，並得到三個主軸，之後利用三個主軸來做圓柱投影，獲得兩種灰階值影像，分別是計算高度分佈資訊以及模型表面三角片法向量與投影線之間的夾角來當做灰階值，最後經由傅立葉轉換和小波轉換來擷取特徵來做相似度測量。



圖八 賽車 3D 模型利用三個圓柱投影分別產生兩種灰階值影像。

(四) 研究方法及步驟

由於目前現有的 3D 模型搜尋的方法，對於一個複雜的 3D 模型，只用一種特徵擷取的方法是稍嫌不足的。因此希望能藉由擷取 3D 模型的多種投影特徵來加強對 3D 模型的搜尋。圖九為 3D 模型搜尋系統流程圖。



圖九 3D 模型搜尋系統流程圖

首先會說明六立面投影如何描述一個 3D 模型的外部資訊以及各投影特徵的相似度比對演算法：首先，用一個最小正方體將 3D 模型包住，再利用網格方式將 3D 模型切割成 $2R \times 2R \times 2R$ 的網格，然而切割網格的主要好處是可過濾模型的細微變

化，也充分的表現出 3D 模型的外型。為了正規化此系統，因此移動 3D 模型的質量中心到 (R, R, R) 的位置，並且縮放 3D 模型(注意！並非縮放網格)讓所有有效網格到達中心的平均距離為 $R/2$ 。在此 R 設定為 128。移動 3D 模型的質量重心到此正立方體的中心。為了減少 3D 模型旋轉所產生的誤差，因此用 PCA[14] 的方法來對所有網格進行擺正的動作。在任意網格中若有包含 3D 模型的一部分時，此網格就會被標示為 1，(我們稱此網格為有效網格)；反之，網格中沒有含括到 3D 模型則標示為 0，(我們稱此網格則為無效網格)。

針對六立面體投影，我們希望結合三種投影方法：

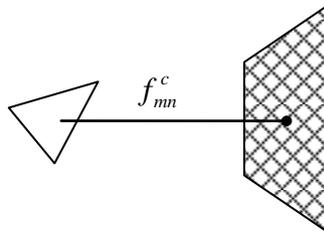
1. 深度投影。
2. 曲率投影。
3. 角度投影。

1. 深度投影

首先，一個 3D 模型從六個不同的視角，產生六個不同的投影灰階平面，分別為：正立面圖、背立面圖、右側立面圖、左側立面圖、俯視圖與底視圖的高度分部的資訊，並顯示這些立面圖為灰階[9]。灰階值的深淺代表模型在該立面圖的高度分佈資訊，其中越靠近觀察者的灰階值越高，反之則越低。其中正立面圖二維灰階影像 F^1 中的每個像素的灰階值定義為：

$$f_{mn}^c = \max\{(127 - h)Voxel_{mnh}\}$$

其中 $m = 1, 2, \dots, 128$ 、 $n = 1, 2, \dots, 128$ 、 $h = 1, 2, \dots, 128$ 。同理，其他五個立面圖之二維灰階影像則分別定義為 F^k ， $k = 1, 2, \dots, 6$ ， k 是立面圖的編號。



圖十 深度投影示意圖

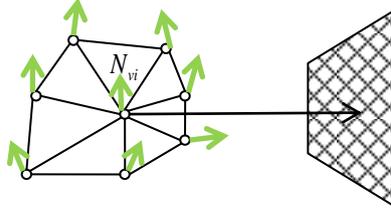
在收集這些立面圖的高度資訊上，採用不同半徑 r 的同心圓來分割立面圖，其中 $r = 1, 2, \dots, 64$ ， k 是立面圖的編號，其中 $k = 1, 2, \dots, 6$ 。則六立面圖特徵向量 x^D 定義為：

$$x^D = \left[(x^1)^T, (x^2)^T, \dots, (x^6)^T \right]^T, \text{ 其中 } x^k = [x_1^k, x_2^k, \dots, x_{64}^k]^T$$

2. 曲率投影

曲率投影是由任一端點 p_i 周圍鄰近的三角片計算法向量而求得該端點的曲率 Γ_{p_i} ，3D 模型在經過曲率計算後，每個網格都以最靠近最小立面體的端點的曲率，當做網格的曲率值，表示成 $CurVoxel(x, y, z) = \Gamma_{v_i}$ [13]。從 3D 模型“前、後、

左、右、上、下”六個不同的視角進行曲率值投影，產生六個不同的灰階影像平面。



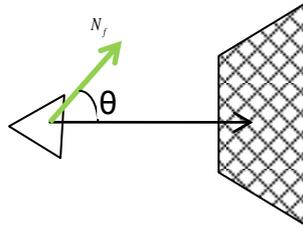
圖十一 曲率投影示意圖

在收集這些立面圖的高度資訊上，採用不同半徑 r 的同心圓來分割立面圖，其中 $r = 1, 2, \dots, 64$ ， k 是立面圖的編號，其中 $k = 1, 2, \dots, 6$ 。則六立面圖特徵向量 x^C 定義為：

$$x^C = \left[(x^1)^T, (x^2)^T, \dots, (x^6)^T \right]^T, \text{ 其中 } x^k = [x_1^k, x_2^k, \dots, x_{64}^k]^T$$

3. 角度投影

角度投影的計算方式是對模型表面三角片去計算法向量與投影線之間的夾角，利用此法向量與投影線的夾角做餘弦計算當作六立面投影的灰階值，灰階值的深淺代表模型表面三角片的角度變化[14]。



圖十二 角度投影示意圖

在收集這些立面圖的高度資訊上，採用不同半徑 r 的同心圓來分割立面圖，其中 $r = 1, 2, \dots, 64$ ， k 是立面圖的編號，其中 $k = 1, 2, \dots, 6$ 。則六立面圖特徵向量 x^A 定義為：

$$x^A = \left[(x^1)^T, (x^2)^T, \dots, (x^6)^T \right]^T, \text{ 其中 } x^k = [x_1^k, x_2^k, \dots, x_{64}^k]^T$$

網格特徵的相似度比對。假設 \mathbf{x} 和 \mathbf{y} 分別為查詢的 3D 模型 q 以及比對的 3D 模型 s 的六立面投影特徵，那麼在 \mathbf{x} 和 \mathbf{y} 之間的差異程度分別定義為：

$$Dis_{q,s}^D = \sum_{i=1}^N |x^D(i) - y^D(i)|$$

$$Dis_{q,s}^C = \sum_{i=1}^N |x^C(i) - y^C(i)|$$

$$Dis_{q,s}^A = \sum_{i=1}^N |x^A(i) - y^A(i)|$$

那麼兩個 3D 模型的相似度即： $Sim = 1/(\omega_1 Dis_{q,s}^C + \omega_2 Dis_{q,s}^C + \omega_3 Dis_{q,s}^C)$ ，其中 ω 為權重值初始設為 1，之後會在實驗過程中調整。

因此當 \mathbf{x} 和 \mathbf{y} 之間的差異越小，即表示說此兩個 3D 模型的相似度越大。假如查詢的 3D 模型和比對的 3D 模型很相似的話，那麼他們的網格的向量特徵也會非常相似。

針對合併方法，我們利用了三種特徵與四種合併方法。

三種特徵分別為：

- (1) AED (ART Elevation Descriptor)
- (2) CMD (Curvature Maps Descriptor)
- (3) AngleED (Angle Elevation Descriptor)

四種合併方式分別為：

- (1) 特徵值相加
- (2) 特徵值相乘
- (3) 特徵值除上特徵長度
- (4) 正規化

(1) 特徵值相加

特徵值相加的比對方法是將三個特徵值分別取出對減距離，在與分別對應的投影特徵取出對減距離。將兩種對減距離相加，在去執行相似度比對。假設 \mathbf{x} 和 \mathbf{y} 分別為查詢的 3D 模型 q 以及比對的 3D 模型 s 的六立面投影特徵，其對減距離利用公式來定義如下：

$$Dis_{q,s}^D = \sum_{i=1}^N |x^D(i) - y^D(i)|$$

$$Dis_{q,s}^C = \sum_{i=1}^N |x^C(i) - y^C(i)|$$

$$Dis_{q,s}^A = \sum_{i=1}^N |x^A(i) - y^A(i)|$$

(2) 特徵值相乘

特徵值相乘的方式與特徵值相加有點相似，分別取出三種特徵值得對減距離，在取出投影特徵的對減距離。將兩種對減距離相乘，在去執行相似度對比。

(3) 特徵值除以特徵長度

特徵值除以特徵長度是先將三種特徵值對減距離和投影特徵對減距離取出，在分別除以個別特徵值的特徵長度。

(4) 正規化

將全部的特徵值做出統計，其中特徵值以 f 表示，將特徵值依序從小排序到大，取全部特徵值中的 3% 為最小值，97% 為最大值，其中最小值為 Min ，最大值為 Max 。使最小值以下為 0，最大值以上為 1，其好處為不會讓極小值或極大值影響到相似度的比對。在利用公式： $F = (f - Min) / (Max - Min)$ ，來定義正規化後的特徵值，其中 F 為新的特徵值。將要合併的特徵正規劃以後，再利用特徵值相加的方式來合併。

實驗結果

以下是 AED、CMD、AngleED 的正確率與各別的投影特徵正確率：

	ART		Histogram(8 Bin)	
AED	0.395		0.377	
Curvature	0.437		0.293	
AngleED	0.417		0.328	

	32 Bin	16 Bin	8 Bin	4 Bin
AED	0.342	0.389	0.377	0.364
Curvature	0.260	0.279	0.293	0.281
AngleED	0.313	0.321	0.328	0.314

以下是利用四種不同的合併方式取得的正確率：

	Art + His	Art + His (除特徵長度)	Art * His	Normalize (Art + His)
AED	0.424	0.417	0.414	0.417
Curvature	0.441	0.416	0.392	0.437
AngleED	0.427	0.420	0.396	0.421

我們利用查全率(recall)的方式來比對正確率。Recall 是將相關模型取得個數除上相關模型個數(class)，其中相關模型以 N 表示，相關模型個數以 T 表示。i 是 class 編號，j 代表 3D 模型編號，其定義如下：

$$Re_i^j = N_i^j / T_i$$

那麼 recall 平均值將定義如下：

$$Re = \frac{1}{T_s} \sum_{i=1}^{92} \sum_{j=1}^{T_i} Re_i^j \quad \text{其中 } T_s = T_1 + T_2 + \dots + T_{92}.$$

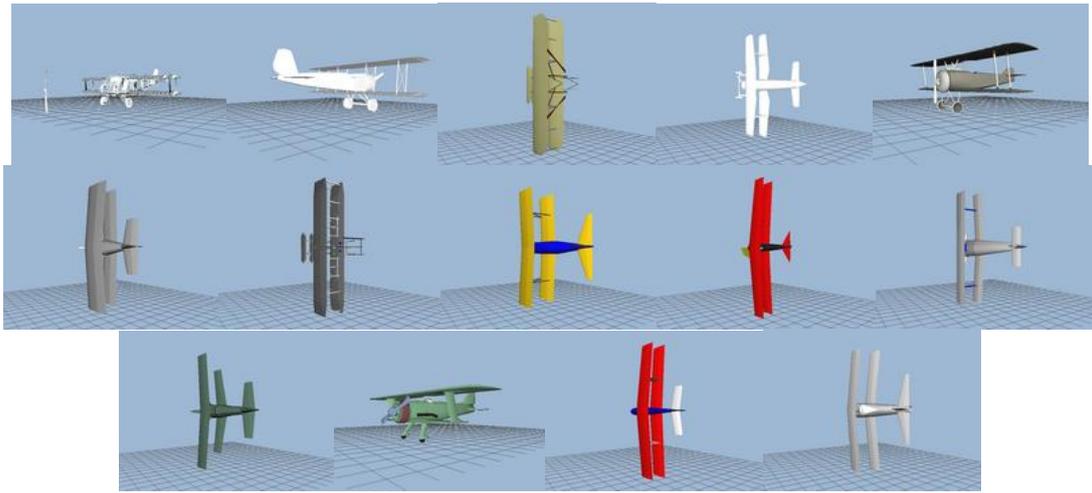
對於一個複雜的 3D 模型而言，只用單一的外部特徵是無法很正確的搜尋到所需的 3D 模型。因此利用三種特徵的方法來擷取。所以在本計劃中，我們藉由深度投影、曲率投影、角度投影的方法來進行擷取 3D 模型的外部特徵，來取得較高的相似度，增加搜尋的正確率。本計劃是利用三種特徵得整合來加強擷取 3D 模型的外部特徵，並且對三種外部特徵來進行比對。

我們使用普林斯頓的資料庫(PSB testing)，此資料庫有 92 個資料夾，共有 907 張 3D 網狀物物件檔格式(.off)。

下表為普林斯頓的資料庫內容：

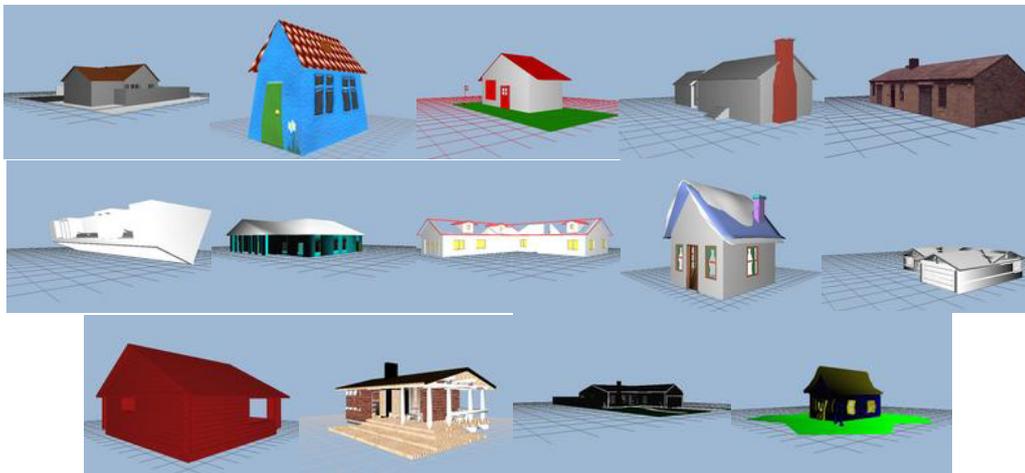
Category name	[Nc]	Category name	[Nc]	Category name	[Nc]	Category name	[Nc]
Biplane	14	Book	4	Vase	11	Flying_bird	14
Commercial	11	Barn	5	Mailbox	7	School_desk	4
Fighter_jet	50	Church	4	Electrical_guitar	13	Staircase	7
Glider	19	Gazebo	5	Newtonian_toy	4	Standing_bird	7
Stealth_bomber	5	One_story_home	14	Bush	9	Bench	11
Hot_air_balloon	9	Skyscraper	5	Flowers	4	Hammer	4
Helicopter	18	One_peak_tent	4	Potted_plant	26	Dog	7
Enterprise_like	11	Two_story_home	10	Barren	11	Dining_chair	11
Flying_saucer	13	Chess_set	9	Conical	10	Shovel	6
Satellite	7	City	10	Satellite_dish	4	Horse	6
Tie_fighter	5	Desktop	11	Large_sail_boat	6	Desk_chair	15
Ant	5	Computer_monitor	13	Ship	11	Umbrella	6
Butterfly	7	Door	18	Submarine	9	Rabbit	4
Human	50	Eyeglasses	7	Billboard	4	Shelves	13
Human_arms_out	20	Fireplace	6	Sink	4	Race_car	14
Walking	8	Cabinet	9	Slot_machine	4	Snake	4
Rectangular	25	Sedan	10	Fish	17	Single_leg	6
Covered_wagon	5	Sea_turtle	6	Geographic_map	12	Motorcycle	6
Axe	4	Handgun	10	Monster_tuck	5	Knife	7
Hat	6	Semi	7	Sword	16	Hourglass	6
Jeep	5	Face	16	Ladder	4	Train_car	5
Hand	17	Streetlight	8	Wheel	4	Head	16
Glass_with_stem	9	Gear	9	Skull	6	Pail	4

資料庫中較相近的 3D 模型：



圖十四 普林斯頓資料庫中的 biplane class

資料庫中較不相近的 3D 模型：



圖十五 普林斯頓資料庫中的 one_story_home class

(五) 結論

我們利用投影特徵求得正確率，正確率雖然沒比原本特徵的高，但是好處是能辨識較細微處的變化，例如：3D 人體模型的手或腳，假使有擺動較容易辨識。但是將投影特徵與其他特徵合併再去求得正確率，會使正確率提高，有助於對相似度比對，使更容易搜尋到正確模型。

(六)參考文獻:

- [1] <http://www.sonypictures.com.tw/movie/CloudywithaChanceofMeatballs/index.htm>
- [2] Demo:<http://disney.go.com/disneyvideos/animatedfilms/chickenlittle/>
- [3] MPEG Video Group, MPEG-7 Visual part of eXperimentation Model Version 9.0, Doc. ISO/IEC JTC1/SC29/WG11/N3914, Pisa, January 2001.
- [4] National Taiwan University, CM Lab Graphics Group, 3D Model Retrieval System, Demo:http://3d.csie.ntu.edu.tw/%7Edynamic/cgi-bin/DatabaseII_v1.8/index.html
- [5] National Institute of Multimedia Education, Motomi T. Suzuki, Demo:<http://www.nime.ac.jp/~motofumi/Ogden>
- [6] Informatics and Telematics Institute, 3D Search, Demo:<http://3d-search.iti.gr/3DSearch>
- [7] Princeton University, Princeton Shape Retrieval and Analysis Group, 3D Models Search Engine, Demo:<http://shape.cs.princeton.edu/search.html>
- [8] T. Funkhouser, P. Min, M. Kazhdan, J. Chan, A. Halderman, D. Dobkin, D. Jacobs, "A search engine for 3D models", ACM Trans. Graphics 22 (1), pp. 83-105, 2003.
- [9] J. L. Shih, C. H. Lee, and J. T. Wang, "A New 3D Model Retrieval Approach Based on the Elevation Descriptor", Pattern Recognition , Vol. 40, No.1, pp. 283-295, Jan 2007.
- [10] D. Y. Chen, X. P. Tian, Y. T. Shen, and M. Ouhyoung, "On visual similarity based 3D model retrieval", Computer Graphics Forum 22 (3), pp. 223-232, 2003.
- [11] D. Y. Chen and M. Ouhyoung, "A 3D Model Alignment and Retrieval System", Proc. of International Computer Symposium, Workshop on Multimedia Technologies, Vol. 2, pp. 1436-1443, Hualien, Taiwan, Dec. 2002.
- [12] M. Chaouch and A. Verroust-Blondet, "A New Descriptor for 2D Depth Image Indexing and 3D Model Retrieval", in Proc. of IEEE International Conference on Image Processing (ICIP 2007), Vol. 6, pp. 373-376, 2007.
- [13] J. Assfalg, A.D. Bimbo, and P. Pala, "Content-based retrieval of 3D models through curvature maps: a CBR approach exploiting media conversion", Proceedings of Multimedia Tools and Applications, Vol. 31, No. 1, pp. 29-50, 2006.
- [14] P. Papadakis, I. Pratikakis, T. Theoharis and S. Perantonis, "PANORAMA: A 3D Shape Descriptor based on Panoramic Views for Unsupervised 3D Object Retrieval", International Journal of Computer Vision, Special Issue on: 3D Object Retrieval.