行政院國家科學委員會專題研究計畫　成果報告

以六立面圖為基礎之 3D 模型搜尋引擎

計畫類別：個別型計畫
計畫編號：NSC94-2213-E-216-018-
執行期間：94 年 08 月 01 日至 95 年 07 月 31 日
執行單位：中華大學資訊工程學系

計畫主持人：石昭玲
共同主持人：李建興
計畫參與人員：楊長慎、陳弘裕

報告類型：精簡報告

處理方式：本計畫涉及專利或其他智慧財產權，1 年後可公開查詢

中　華　民　國 95 年 10 月 23 日

行政院國家科學委員會補助專題研究計畫 □ 成 果 報 告 □期中進度報告

以六立面圖為基礎之 3D 模型搜尋引擎

計畫類別：☑ 個別型計畫　　□ 整合型計畫
計畫編號：NSC 94-2213-E-216-018-
執行期間：　　2005 年 08 月 01 日 至 2006 年 07 月 31 日

計畫主持人：石昭玲
共同主持人：李建興
計畫參與人員： 楊長慎、陳弘裕

成果報告類型(依經費核定清單規定繳交)：□精簡報告　☑完整報告

本成果報告包括以下應繳交之附件：
□赴國外出差或研習心得報告一份
□赴大陸地區出差或研習心得報告一份
□出席國際學術會議心得報告及發表之論文各一份
□國際合作研究計畫國外研究報告書一份

處理方式：除產學合作研究計畫、提升產業技術及人才培育研究計畫、列管計畫及下列情形者外，得立即公開查詢
□涉及專利或其他智慧財產權，□一年□二年後可公開查詢

執行單位：中華大學資訊工程學系

中　華　民　國　95　年　10　月　31　日

摘要

隨著 3D 形狀模組化、數位化、以及視覺化等技術快速發展，導致了一個 3D 模型爆炸的時代，在網路上或是特定領域的資料庫中到處都充斥著可以利用的 3D 模型。因此如何建立一個有效的 3D 模型搜尋系統，讓使用者可以利用此一系統快速地找到在大型 3D 模型資料庫中符合使用者個人期待的相同或相似 3D 模型是本計畫的首要目標。因此在本計畫中提出一個全新以形狀為基礎的 3D 模型特徵擷取演算法，六立面圖特徵（elevation descriptor）擷取法。

一. 報告內容

1. 前言

因為 3D 模型在數位圖書館中的數量逐漸暴增，我們是很迫切需要一個搜尋系統去幫助人們找到他們所要的 3D 資料。因此在本計畫中提出一個全新的 3D 模型檢索系統的特徵方法，以六立面圖特徵（elevation descriptor）擷取法。它的關鍵就在於如何以 3D 模型本身的內容（content）搜尋出符合使用者需求的 3D 模型。

2. 研究目的與研究方法

我們提出了六立面圖特徵（elevation descriptor），主要概念為收集 3D 模型在六個不同的視角下的立面形狀分布，這六個立面圖分別為：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。首先，將這多邊形模型分割成 $2R\times 2R\times 2R$ 的網格（Voxel grid），如圖一(b)。在這網格 (m, n, h)中如果有 3D 模型中多邊形的面存在則$Voxel_{mnh} = 1$，反之 $Voxel_{mnh} = 0$。這方法可以過濾物體的外型的雜訊，也可充分表現該物體的外型。為了能夠正規化此系統，移動 3D 模型的質量中心到（R, R, R）的位置，然後縮放 3D 模型(注意！並非縮放網格）讓非零網格到達中心的平均距離必須為 $R/2$ ，如圖二(b)。在本計劃中此 R 將設定為 32。以這方法來對整個 3D 模型做取樣(sampling)的動作。

接下來擷取這網格的六個立面圖，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。我們將以灰階值代表模型在該立面圖的高度分布，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

$$f^1_{mn} = \max\{(63 - h)Voxel_{mnh}\}，$$

其中 $m=1,2,\ldots,64$ 、$n=1,2,\ldots,64$ 、$h=1,2,\ldots,64$。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 = 2, 3, …,6，k 是立面圖的編號，如圖三。

(a) (b)

圖一 3D 模型及其網格座標 (a)3D 吉普車模型，(b)網格座標。

(a) (b)

圖二 利用網格正規化 3D 模型 (a)正規化 3D 模型前，(b)正規化 3D 模型後。

在收集這些立面圖的高度資訊上，本計劃採用不同半徑的同心圓來分割立面圖，如圖四。其中正立面圖一個半徑 $r-1$ 與 r 之間的環狀圖形的灰階值總合 g_r^1 定義為

$$g_r^1=\sum_m\sum_n\left(f_{mn}^1\,\middle|\,r-1\le\sqrt{(m-R)^2+(n-R)^2}<r\right),$$

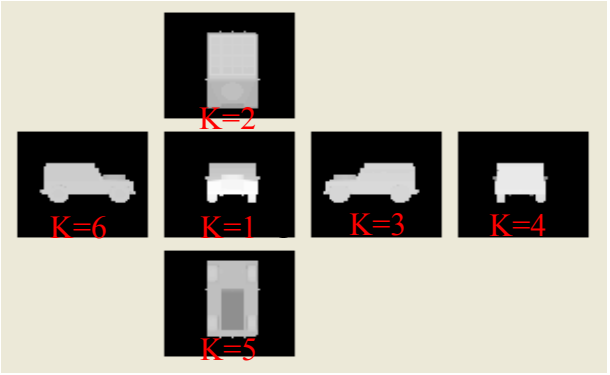
其中 $r=1,2,\ldots,32$。同理，其他五個立面圖之環狀圖形的灰階值總合則分別定義為 g_r^k ，$k=2,3,\ldots,6$ ，k 是立面圖的編號，$r=1,2,\ldots,32$。為了正規化這些數值，將定義整張立面圖的灰階值總合為 $G^k=\sum_{r=1}^{32}g_r^k$ ，其中 $k=1,2,\ldots,6$。則六立面圖特徵向量 x 定義為：

$$\mathbf{x}=\left[(\mathbf{x^1})^T,(\mathbf{x^2})^T,\ldots,(\mathbf{x^6})^T\right]^T,$$

其中 $\mathbf{x}^k=\left[x_1^k,x_2^k,\ldots,x_{32}^k\right]^T$ 且

$$x_r^k = \frac{g_r^k}{\sum_{k=1}^{6} G^k},$$

其中 $k = 1, 2, \ldots, 6$，$r = 1, 2, \ldots, 32$。



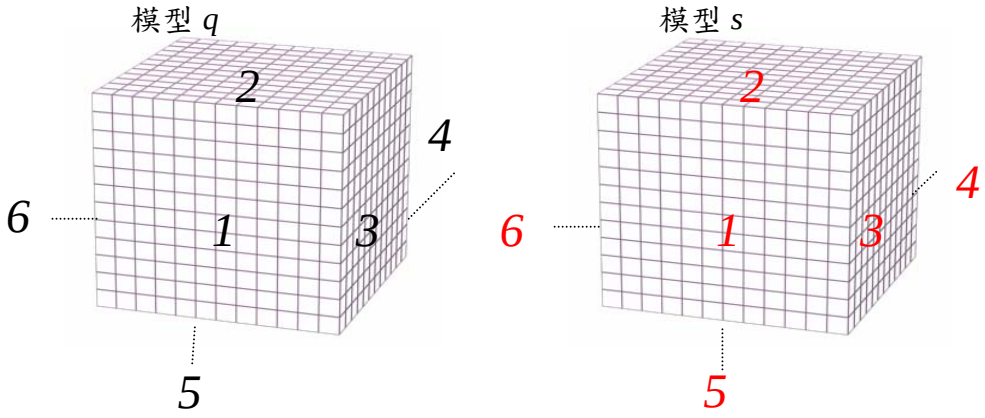
圖三 3D 吉普車_1 模型的原圖及其六立面圖，正立面圖(K=1)、背立面圖(K=4)、右側立面圖(K=3)、左側立面圖(K=6)、俯視圖(K=2)與底視圖(K=5)。

圖四 立面圖上不同半徑的同心圓示意圖。

再來就是本計畫中所提出六立面圖特徵向量的比對方法。為了比較兩個模型的相似程度，必須計算兩個模型的六立面圖特徵向量的差異值，基於六立面圖特徵向量是擷取 3D 模型的六個角度視角的關係，在計算兩模型的差異值不能僅僅只計算兩個向量的差值，則必須考慮每一個面與其他的面的差異值。如圖五，理論上，模型 *q* 與模型 *s* 的比對將有 720 種（6 階層）種排列組合，因此計算模型 *q* 與模型 *s* 的六立面圖特徵向量差異值必須取其中一種使得差異值為最小的排列方式。

為了減少計算量，我們將模型六個面彼此的關係也考慮進去，減少兩模型六個面比對時的排列組合。基於一個概念，模型的面必須跟其背對的面相對應，以圖七，的模型 *q* 為例子，模型 *q* 的面 1 必須與模型 *q* 的面 4 相對應、模型 *q* 的面 2 必須與模型 *q* 的面 5 相對應、模型 *q* 的面 3 必須與模型 *q* 的面 6 相對應。因此在計算模型 *q* 與模型 *s* 的差異值時，模型 *q* 的面 1 與模型 *s* 的面 1 計算差值的話，則模型 *q* 的面 4 就必須與模型 *s* 的面 4 計算差值；如果模型 *q* 的面 2 與模型 *s* 的面 6 計算差值的話，則模型 *q* 的面 5 與模型 *s* 的面 3 計算差值。若加

以考慮鏡射的情況，當模型 q 的面 1 與模型 s 的面 1 相比較時，在模型 q 的面 3 可能會與模型 s 的面 3 或面 6 發生對應，如此可以將其原本六個面的排列組合 720 種大幅縮減為 48 種，如表一所列，我們希望找到一組排列組合使得六立面圖特徵向量的差異值為最小。



圖五 兩個模型的六個面作比對的排列組合。

當 $d=1$ 時，模型 q 的六立面圖特徵向量為 x，模型 s 六立面圖特徵向量為 y， 則兩模型的特徵向量差異值為

$$Dis\_Temp_{q,s}^{1}=\sum_{k=1}^{6}\left\|\mathbf{x}^{k}-\mathbf{y}^{m(k)}\right\|=\sum_{k=1}^{6}\sum_{r=1}^{32}\left|x_{r}^{k}-y_{r}^{m(k)}\right|,$$

其中 $m(k)=1,2,3,4,5,6$ 當 $k=1,2,3,4,5,6$。如果當 $d=2$ 時，模型 q 與模型 s 的特徵向量差異值變為

$$Dis\_Temp_{q,s}^{2}=\sum_{k=1}^{6}\left\|\mathbf{x}^{k}-\mathbf{y}^{m(k)}\right\|=\sum_{k=1}^{6}\sum_{r=1}^{32}\left|x_{r}^{k}-y_{r}^{m(k)}\right|,$$

其中 $m(k)=1,3,2,4,6,5$ 當 $k=1,2,3,4,5,6$。以此我們定義搜尋模型 q 與資料庫中欲比對之模型 s 的六立面圖特徵向量差異值為

$$Dis_{q,s}=\min_{d} Dis\_Temp_{q,s}^{d}$$

其中 $d=1,2,3,\ldots,48$。即兩個模型的特徵向量差異值越小，則其相似程度就越大。如圖六，上面兩個模型的特徵向量差異值比較小，則此兩模型的相似程度就較大；反之，下面兩個模型的特徵向量差異值比較大，則此兩模型的相似程度就較小。

3. 實驗結果與討論

將介紹我們所採用的資料庫以及所做的實驗。實驗裡使用的量測標準為 Recall 與 Precision [1]，定義為

$$recall = \frac{N}{T} \quad 與 \quad precision = \frac{N}{K}$$

其中 N 是檢索結果中與搜尋目標有相關的個數，T 是全部與搜尋目標相關的個數，K 是所有被檢索出的 3D 模型個數。

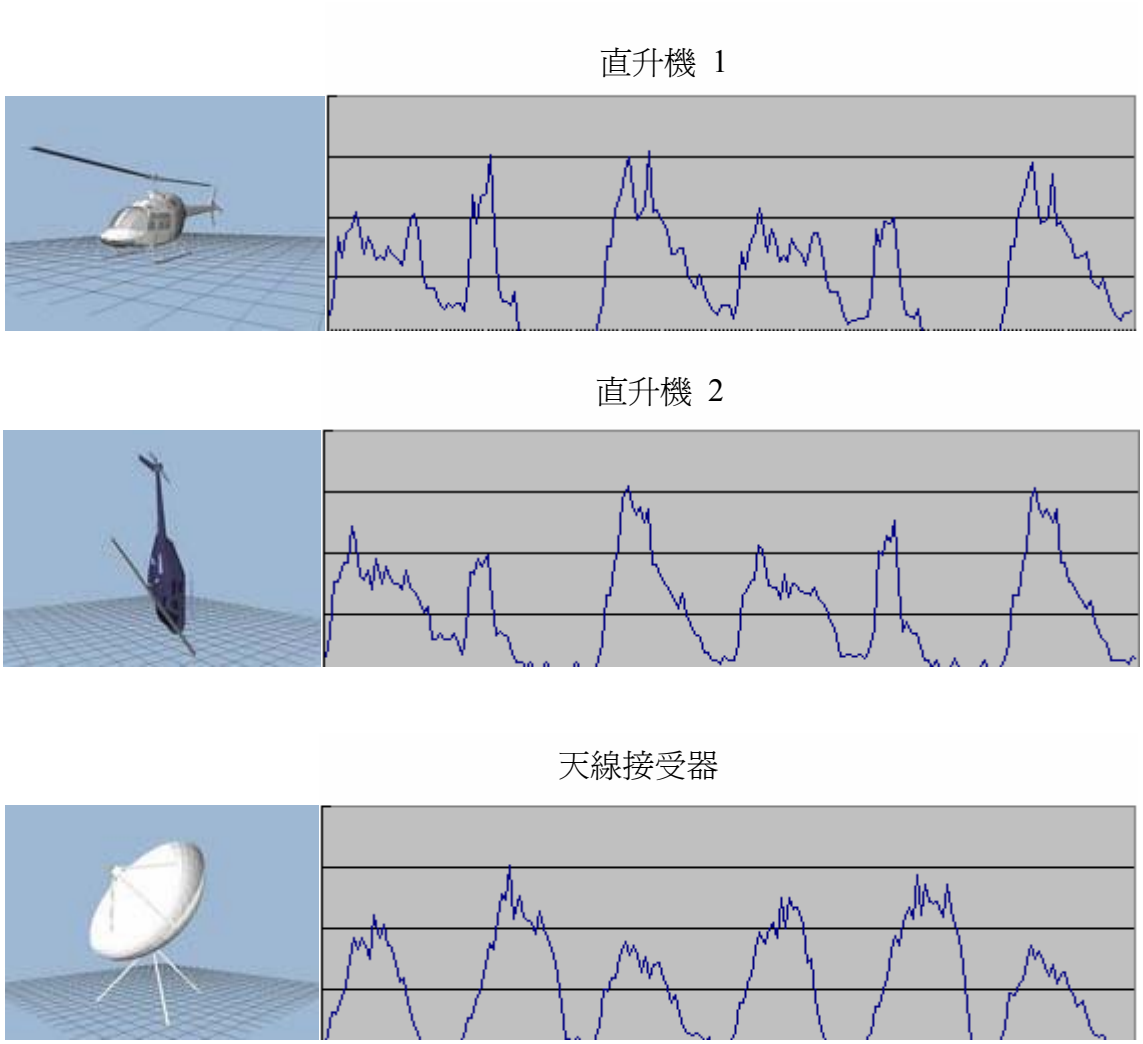
實驗所採用的資料庫是「普林斯頓形狀基準資料庫（Princeton Shape Benchmark）」[2, 3]，此資料庫是提供給對 3D 模型檢索研究有需求的使用者免費使用，大部分的 3D 模型特徵擷取[4-11]都有用此資料庫來做檢索研究。此資料庫含有 1814 個 3D 模型，包含 161 個不同的類別，如圖七。

表一 六個面比對時的 48 種排列組合。

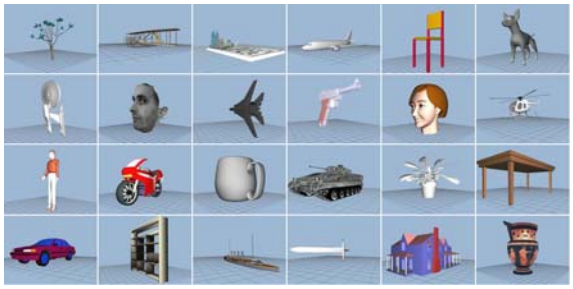
k	1	2	3	4	5	6	k	1	2	3	4	5	6	k	1	2	3	4	5	6
$p_1(k)$	1	2	3	4	5	6	$p_{17}(k)$	2	1	3	5	4	6	$p_{33}(k)$	3	2	1	6	5	4
$p_2(k)$	1	3	2	4	6	5	$p_{18}(k)$	2	3	1	5	6	4	$p_{34}(k)$	3	1	2	6	4	5
$p_3(k)$	1	5	3	4	2	6	$p_{19}(k)$	2	4	3	5	1	6	$p_{35}(k)$	3	5	1	6	2	4
$p_4(k)$	1	3	5	4	6	2	$p_{20}(k)$	2	3	4	5	6	1	$p_{36}(k)$	3	1	5	6	4	2
$p_5(k)$	1	2	6	4	5	3	$p_{21}(k)$	2	1	6	5	4	3	$p_{37}(k)$	3	2	4	6	5	1
$p_6(k)$	1	6	2	4	3	5	$p_{22}(k)$	2	6	1	5	3	4	$p_{38}(k)$	3	4	2	6	1	5
$p_7(k)$	1	5	6	4	2	3	$p_{23}(k)$	2	4	6	5	1	3	$p_{39}(k)$	3	5	4	6	2	1
$p_8(k)$	1	6	5	4	3	2	$p_{24}(k)$	2	6	4	5	3	1	$p_{40}(k)$	3	4	5	6	1	2
$p_9(k)$	4	2	3	1	5	6	$p_{25}(k)$	5	1	3	2	4	6	$p_{41}(k)$	6	2	1	3	5	4
$p_{10}(k)$	4	3	2	1	6	5	$p_{26}(k)$	5	3	1	2	6	4	$p_{42}(k)$	6	1	2	3	4	5
$p_{11}(k)$	4	5	3	1	2	6	$p_{27}(k)$	5	4	3	2	1	6	$p_{43}(k)$	6	5	1	3	2	4
$p_{12}(k)$	4	3	5	1	6	2	$p_{28}(k)$	5	3	4	2	6	1	$p_{44}(k)$	6	1	5	3	4	2
$p_{13}(k)$	4	2	6	1	5	3	$p_{29}(k)$	5	1	6	2	4	3	$p_{45}(k)$	6	2	4	3	5	1
$p_{14}(k)$	4	6	2	1	3	5	$p_{30}(k)$	5	6	1	2	3	4	$p_{46}(k)$	6	4	2	3	1	5
$p_{15}(k)$	4	5	6	1	2	3	$p_{31}(k)$	5	4	6	2	1	3	$p_{47}(k)$	6	5	4	3	2	1
$p_{16}(k)$	4	6	5	1	3	2	$p_{32}(k)$	5	6	4	2	3	1	$p_{48}(k)$	6	4	5	3	1	2

使用的特徵向量有以下五種：六立面圖特徵（ED）擷取法改良式 D2（AD2）[12]、球型諧波（SH）[13]、MPEG-7 的 3D 形狀頻譜描述（SSD）[5]、3D 幾何形狀分佈的 D2[6]。首先的實驗是五種特徵向量之間對不同類別的 3D 模型之正確率的比較，我們挑選普林斯頓形狀基準資料庫中的一些類別來做檢索，挑選的類別都是包含 15 個模型以上的類別，有 barren、biplane、city、commercial、dining_chair、enterprise_like、face、fighter_jet、handgun、head、helicopter、human、human_arms_out、military_tank、potted_plant、rectangular、sedan、shelves、ship、sword、two_story_home 與 vase 以上幾個類別，如圖七。將以上類別的每個

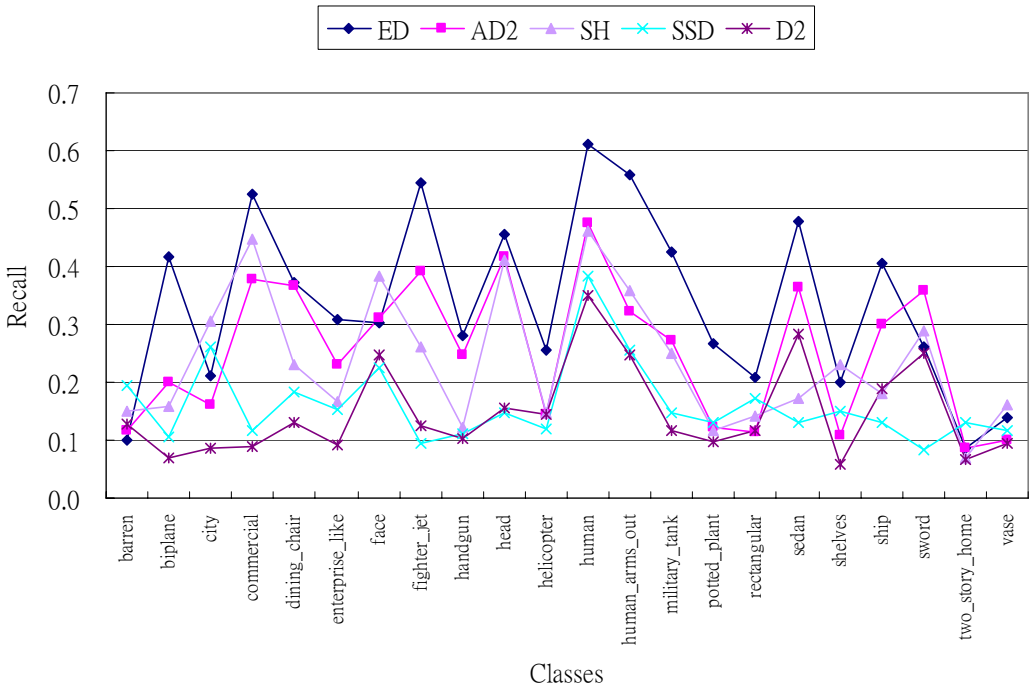
3D 模型都當成檢索目標在普林斯頓形狀基準資料庫中做檢索，並計算其正確率，圖八為實驗結果。可以很明顯的看出六立面圖特徵（elevation descriptor）擷取法特徵向量在大部分的類別擁有最好的正確率，而只有少部分的類別是別的特徵向量較好。由此驗證不同類別使用某些特徵擷取演算法將得到較好的結果。故本計畫以此論點採用六立面圖特徵（elevation descriptor）擷取法，以期達到較好的效果。接下來進行的實驗是五種不同的特徵向量之正確率的比較，將普林斯頓形狀基準資料庫中的每個 3D 模型都當成檢索目標在資料庫中做檢索，並計算其正確率，表二與圖九為實驗結果。以上，我們所提出的六立面圖特徵（elevation descriptor）擷取法在普林斯頓形狀基準資料庫中都有很不錯的檢索效果，並更加符合使用者的需求，且提高了檢索的正確率，所以由此實驗可說明本計畫所提出的方法是非常有成效的。



圖六 三個模型的六立面圖特徵向量。



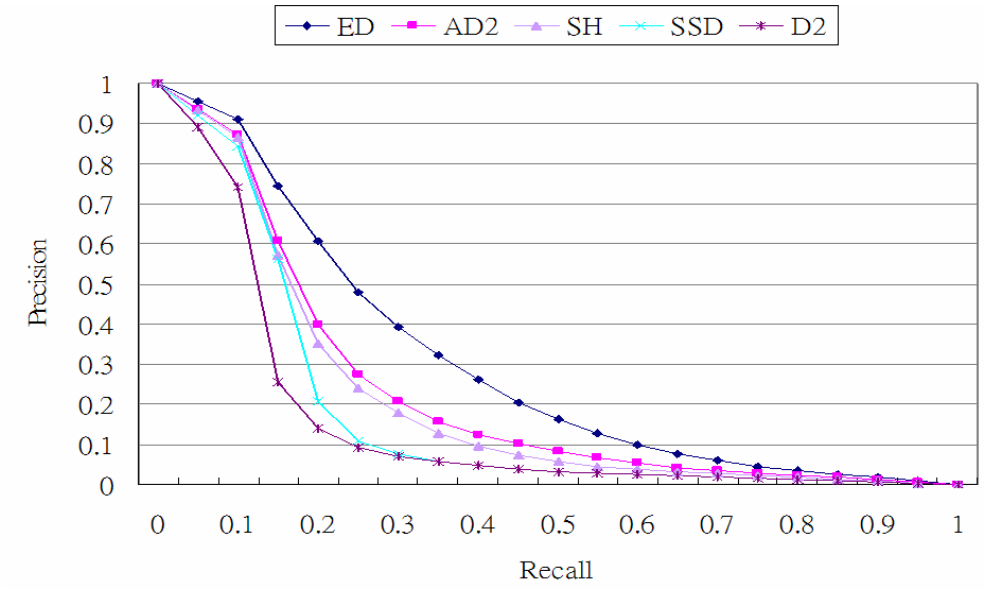
圖七 普林斯頓形狀基準資料庫中的一些類別。



圖八 不同類別的模型之正確率的實驗結果。

表二 不同特徵向量之間正確率的比較。

	Recall
六立面圖特徵（ED）	0.337000
改良式 D2（AD2）	0.2576262804
球型諧波（SH）	0.2451209359
MPEG-7 的 3D 形狀頻譜描述符（SSD）	0.2010285473
3D 幾何形狀分佈的 D2	0.1745288704



圖九 不同特徵向量之間正確率的比較。

二. 參考文獻

[1] Y. Deng and B. S. Manjunath, " An efficient low-dimensional color indexing scheme for region-based image reitrieval," in Proc. of IEEE Int. Conf. Acoustics, Speech, and Signal Processing, Vol. 6, pp. 3017-3020, 1999.

[2] Princeton Shape Benchmark, http://shape.cs.princeton.edu/benchmark/.

[3] P. Shilane, P. Min, M. Kazhdan, and T. Funkhouser, "The Princeton Shape Benchmark," in Proc. of Shape Modeling International, Genova, Italy, June 2004.

[4] W. B. Frakes and R. Baeza-Yates, "Information Retrieval: Data Structure and Algorithms," Prentice Hall, 1992.

[5] B. S. Manjunath, P. Salembier, and T. Sikora, "Introduction to MPEG-7 Multimedia Content Description Interface." John Wiley & Sons Ltd., pp. 247-260, 2002.

[6] R. Osada, T. Funkhouser, B. Chazelle, and D. Dobkin, "Matching 3D models with shape distributions," Shape Modeling International, pp. 154-166, May 2001.

[7] M. Ankerst, G. Kastenmüller, H.P. Kriegel, and T. Seidl, "3D shape histograms for similarity search and classification in spatial databases," in Proc. SSD, 1999.

[8] D. Saupe and D. V. Vrani, "3D model retrieval with spherical harmonics and moments," DAGM 2001, pp. 392-397, 2001.

[9] B. Horn. "Extended Gaussian images," in Proc. of the IEEE, Vol. 72(12), pp. 1671-1686, December 1984.

[10] S. Kang and K. Ikeuchi. "Determining 3-D object pose using the complex extended Gaussian image," in Proc. of CVPR, pp. 580-585, June 1991.

[11] D. V. Vranic. "An improvement of rotation invariant 3D shape descriptor based on functions on concentric spheres," in Proc. of IEEE International Conference on Image

Processing (ICIP 2003), vol. 3, pp. 757-760, September 2003.

[12] J. T. Wang and J. L. Shih, “Shape-based 3D model retrieval system,”NSC 92-2213-E-216-023-, 2004.

[13] T. Funkhouser, P. Min, M. Kazhdan, J. Chen, A. Halderman, D. Dobkin, and D. Jacobs “A Search Engine for 3D Models” in Proc. of ACM Transactions on Graphics, Vol. 22(1), pp. 83-105, January 2003.

三. 計畫成果自評

我們把六立面圖特徵（elevation descriptor）擷取法運用到 3D 模型檢索的領域，並且提升了 3D 模型檢索上的正確性。此六立面圖特徵（elevation descriptor）擷取法，在 3D 模型檢索系統方面有一定的研究價值。未來我們計畫要加大 3D 模型資料庫，且我們深信某些特別類別的 3D 模型，用某些特徵會有較好的效果，所以未來還可以加入其他檢索效果也不錯的特徵擷取方法，讓其檢索結果更佳。

尤其在最近幾年 3D 模型設計主要因為軟、硬體技術上的進步，導致 3D 模型設計上的簡單，使得 3D 模型被廣泛應用，也因此我們認為“3D 模型檢索系統”在學術界及業界裡有相當的研究價值，目前我們已發表五篇與 3D 模型檢索系統相關的論文，包括三篇期刊論文及一篇研討會論文。相關論文如下 ：

期刊論文 (Journal Papers) ：

[1] 王建棠 石昭玲, “具有相關性回饋演算法之智慧型 3D 模型搜尋引擎,” Chung Hua Journal of Science and Engineering, Vol. 2, No. 1. pp. 53-61 , March 2004.

[2] Jau-Ling Shih, Chang-Hsing Lee, and Jian Tang Wang, “3D Object Retrieval System Based on Grid D2,” Electronics Letters Vol. 41 No. 4 2005 pp23-24. (SCI)

[3] J. L. Shih, C. H. Lee, and J. T. Wang, “A New 3D Model Retrieval Approach Based on the Elevation Descriptor,” accepted by Pattern Recognition. (SCI) (見附件三)

研討會論文 (Conference Papers) ：

[1] Jian Tang Wang and Jau-Ling Shih, “Shape-Based 3D Model Retrieval System based on Elevation Descriptor,” Proceeding of The CVGIP, August 2004.

附件二

可供推廣之研發成果資料表

☐ 可申請專利　☑ 可技術移轉　　日期：95 年 10 月 31 日

國科會補助計畫	計畫名稱：以六立面圖為基礎之 3D 模型搜尋引擎 計畫主持人：石昭玲 計畫編號：NSC 94-2213-E-216-018-　　學門領域：影像處理
技術/創作名稱	3D 模型搜尋引擎
發明人/創作人	石昭玲副教授
技術說明	中文：我們利用六立面圖特徵（elevation descriptor）來計算每一個 3D 模型的特徵，主要概念為收集 3D 模型在六個不同的視角下的立面形狀分布，這六個立面圖分別為：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。做法為擷取每個立面圖的高度分布顯示這些立面圖為灰階圖，並且以灰階值代表模型在該立面圖的高度分布，其中越靠近觀察者的灰階值越高，反之則越低。且我們提出一個新的特徵比對方法去計算兩個 3D 模型的相似程度，在資料庫中找出與使用者想搜尋的 3D 模型相似度較高的模型回應給使用者。
	英文：We proposed a novel feature, elevation descriptor, for 3D model retrieval. First, a 3D model is represented by six gray-level images which describe the altitude of a 3D model from six viewing angles including front, left, right, rear, top and bottom. Then each gray-level image, called elevation, is partitioned into several concentric circles to extract the altitude information as the elevation descriptor. Since, there are six elevations for each 3D model, a specially designed similarity measure is then provided to find the best match of two models.
可利用之產業及可開發之產品	可應用於網際網路之線上 3D 模型搜尋系統。
技術特點	利用六面體的六個面很快的將 3D 模型的特徵向量擷取出來，以立面的想法來代表 3D 模型的高度資訊搜尋速度及效率極為優異。
推廣及運用的價值	3D 模型搜尋引擎

※　1.每項研發成果請填寫一式二份，一份隨成果報告送繳本會，一份送　貴單位研發成果推廣單位（如技術移轉中心）。

※　**2.本項研發成果若尚未申請專利，請勿揭露可申請專利之主要內容。**

3.本表若不敷使用，請自行影印使用。

附件三

A NEW 3D MODEL RETRIEVAL APPROACH BASED ON THE ELEVATION DESCRIPTOR

Jau-Ling Shih*, Chang-Hsing Lee, and Jian Tang Wang

Department of Computer Science and Information Engineering,

Chung Hua University, Hsinchu, Taiwan, R.O.C

Abstract

The advances in 3D data acquisition techniques, graphics hardware, and 3D data modeling and visualizing techniques have led to the proliferation of 3D models. This has made the searching for specific 3D models a vital issue. Techniques for effective and efficient content-based retrieval of 3D models have therefore become an essential research topic. In this paper, a novel feature, called *elevation descriptor*, is proposed for 3D model retrieval. The elevation descriptor is invariant to translation and scaling of 3D models and it is robust for rotation. First, six elevations are obtained to describe the altitude information of a 3D model from six different views. Each elevation is represented by a gray-level image which is decomposed into several concentric circles. The elevation descriptor is obtained by taking the difference between the altitude sums of two successive concentric circles. An efficient similarity matching method is used to find the best match for an input model. Experimental results show that the proposed method is superior to other descriptors, including spherical harmonics, the MPEG-7 3D shape spectrum descriptor, and D2.

Keywords: 3D model retrieval, Elevation descriptor

1. Introduction

The development of image, video, and 3D model archives has made multimedia retrieval become a popular research topic. Most of the commercial multimedia retrieval systems employ keyword search to assist users to find desired multimedia data. To facilitate search accuracy, the managers of the multimedia database must empirically annotate well-chosen keywords for all multimedia data. If the database is very large, the task is laborious and time consuming. Moreover, the appropriate keywords differ from person to person. In general, the simplest approach is to extract keywords from filenames, captions, or context (e.g., Google). However, this approach fails when the filenames are not well annotated (e.g., "c0033.jpg") or unspecified filenames are defined (e.g., "jeffrey.gif" or "circle.bmp"). Thus, the demand for an automatic and efficient content-based multimedia retrieval system has become a crucial issue.

With the proliferation of computer graphics and computer animations, 3D models are as plentiful as images and video. The primary challenge to a content-based 3D model retrieval system [1] is to extract proper features for discriminating the diverse shapes of 3D models for efficiently indexing similar ones. The 3D model retrieval methods can be roughly classified into three categories: low-level feature based methods, high-level structure based methods, and view based methods. The low-level feature based methods try to represent the shape of 3D models by their geometric and topological properties. The features can be a single vector consisting of a fixed number of feature values or distributions of a set of feature values. The high-level structure based methods try to decompose a 3D model into a set of key parts and capture the geometric relationships of the key parts. The view based methods project the shape of a 3D model on a number of 2D projections from different views.

The rest of the paper is organized as follows: The related work is described in Section 2. In Section 3, the proposed elevation descriptors are introduced. Section 4 gives the experimental results to show the effectiveness of the proposed elevation descriptor. Finally, conclusions are given in Section 5.

2. Related Work

In this section, some related work for 3D model retrieval is described. The 3D model retrieval methods are classified into three categories: low-level feature based methods, high-level structure based methods, and view based methods.

2.1 Low-Level Feature Based Methods

In 3D model retrieval systems, low-level feature descriptors are usually extracted to describe the geometric properties [2, 3], spatial properties [4-9], and shape distributions [10-16] of 3D models. The similarity between two 3D models can be measured by comparing their features.

Zhang and Chen [2] proposed methods to efficiently calculate features such as area, volume, moments, and Fourier transform coefficients from mesh representation of 3D models. Paquet et al. [3] employed moments to describe symmetries of 3D objects, cord-based descriptors to represent shape information in fine details, and wavelet transform descriptors to describe the density distribution through a volume.

Vranic et al. [4] performed Fourier transform on the sphere with spherical harmonics to get the feature vectors. This method requires pose normalization to be rotation invariant. A modified rotation invariant shape descriptor based on the spherical harmonics without pose normalization has been proposed by Funkhouser et al. [5, 6]. First, a 3D model is decomposed into a collection of spherical functions by intersecting the model with concentric spheres of different radii. Each spherical function is decomposed into a set of harmonics of different frequencies. The sum of norms of each frequency component at each radius forms the shape descriptor. The reason for the descriptor being rotation invariant is that rotating a spherical function does not change the energies in each frequency component. Novotni and Klein [7] used 3D Zernike moments for 3D shape retrieval. It is naturally an extension of spherical harmonics based descriptors. The 3D Zernike moments is a 2D histogram indexed by radius and frequency. The benefits of the 3D Zernike moments are that they are rotation invariant and less sensitive to geometric and topological

artifacts. Yu et al. [8] generated a surface penetration map in which the number of surfaces that the ray emitted from the center of the sphere penetrates is counted. Fourier transform of the map are used for retrieval or comparison purpose. Ankerst et al. [9] proposed shape histograms to characterize the area of intersection of a 3D model with a collection of concentric shells and sectors. Quadratic form distance measure is employed to compute the distance between the histogram bins.

Osada et al. [10] tried to represent each 3D model by the probability distributions of geometric properties computed from a set of randomly selected points located on the surface of the model. These geometric properties, including distance, angle, area, and volume, are employed to describe the shape distribution. Among these distributions, the most effective is D2, which measures the distribution of distances between any two randomly selected points. Ip et al. [11, 12] refined the D2 descriptor by classifying the D2 distance into three categories: IN distance if the line segment connecting the two points lies completely inside the model, OUT distance if the line segment lies completely outside the model, and MIXED distance if the line segment passes both inside and outside the model. The dissimilarity measure is a weighted distance of D2, IN, OUT, and MIXED distributions. However, it is difficult to do the classification task, if a 3D model is represented by polygon meshes. Ohbuchi et al. [13, 14] combined the absolute angle-distance histogram (AAD) with the D2 descriptor for 3D model retrieval. AAD measures the distribution of angles between the normal vectors of two surfaces on which the two randomly selected points locate. In their experimental results, AAD outperforms D2 at the expense of about 1.5 times computational cost. In typical mesh-based representation of 3D models, many polygonal meshes are required to finely represent the complex components of a 3D model. As a result, an area weighted defect will occur since the random sampling of surface points is greatly affected by the complex components. Therefore, Shih et al. [15] proposed a new descriptor called grid D2 (GD2) to alleviate this problem. In GD2, a 3D model is first decomposed into a voxel grid. Rather than on random points, the random sampling operation is performed on voxels within which some polygonal surfaces are

located. The shape spectrum descriptor (SSD) [16] is adopted in the MPEG-7 standard for 3D model retrieval. SSD represents the histogram of curvatures of all points on the 3D surface. The advantages of SSD are that it can match two 3D models without first aligning the 3D objects, and that it is robust to the tessellation of the 3D polygonal model.

2.2 High-Level Structure Based Methods

The low-level feature based methods discussed above only take the geometric or topological properties of 3D models into consideration. On the other hand, high-level structure based methods describe the relationship between model components. Hilaga et al. [17] used multi-resolution Reeb graphs (MRG) to describe the skeleton structure of a 3D model. Mathematically, the Reeb graph is defined as the quotient space of a shape and a quotient function. The Reeb graph used by Hilaga et al. is based on a quotient function defined by an integral geodesic distance. Bespalov et al. [18] applied the Reeb graph for description of solid models. One major advantage of using the Reeb graph to measure the distance between two 3D models is that it is robust to 3D shape deformation. However, computation of the Reeb graph is time consuming and very sensitive to the fine components of 3D models.

2.3 View Based Methods

The main idea of view based methods is to represent a 3D model using a number of binary images. Therefore, a set of 2D features can be used to index similar 3D models. Each binary image is obtained from the boundary contour of the 3D model from different views. Several methods provide a 2D query interface to facilitate view based retrieval of 3D models [5, 19]. Super and Lu [20] exploited 2D silhouette contours for 3D object recognition. Curvature and contour scale space are extracted to represent each silhouette. Chen et al. [21] introduced the lightfield descriptor to represent 3D models. The lightfield descriptor is computed from ten silhouettes. Each silhouette is represented by a 2D binary image. Zernike moments and Fourier descriptors are used to describe

each binary image. Since a 3D model may be rotated or deformed, the number of 2D silhouettes must be large enough to represent a 3D model. On the other hand, the retrieval time increases as the number of silhouettes increases.

In fact, 2D silhouettes represented by binary images do not describe the altitude information of the 3D model from different views well. Therefore, a new descriptor, called the elevation descriptor, is proposed for 3D model retrieval. Six elevations are obtained to represent a 3D model. Each elevation is represented by a 2D gray-level image which describes the altitude information of a 3D model from different views. In addition, an effective way for extracting features from each gray-level image is employed in order to make them less sensitive to rotations. In the following section, the proposed method is described in detail.

3. The Proposed Elevation Descriptor for 3D Model Retrieval

In this section, the proposed elevation descriptor is described. Since the features are extracted from six elevations representing 2D projections from different views, a similarity matching method is used to find the best match for an input model as efficiently as possible.

3.1 Elevation Representation

Initially, the tightest bounding box circumscribing the 3D model is constructed (see Fig. 1(a)). The bounding box is then decomposed into a $2L \times 2L \times 2L$ voxel grid (see Fig. 1(b)). A voxel located at (m, n, h) is regarded as an opaque voxel, notated as $Voxel(m, n, h) = 1$, if there is a polygonal surface located within this voxel; otherwise, the voxel is regarded as a transparent voxel, notated as $Voxel(m, n, h) = 0$. Based on the decomposition process, the area weighted defect is greatly reduced since each opaque voxel is weighted equally irrespective of the number of points located within this voxel. Secondly, the model's mass center is moved to location (L, L, L) and the average distance from all opaque voxels to the mass center is linearly scaled to be $L/2$ so that the elevation descriptor is invariant to translation and scaling, as shown in Fig. 1(c). In this paper, L is

set to be 32, which provides adequate resolution for discriminating objects and the fine-detail noise in complex components of a 3D model can be filtered out.

Next, six elevations are extracted to indicate the altitude information of 2D projections from six different views: front, top, right, rear, bottom, and left. Each elevation is represented by a gray level image in which the gray values denote the altitude information. Let the front, top, right, rear, bottom, and left elevations be notated successively as $\boldsymbol{E}_k$, $k = 1, 2, \ldots, 6$. The gray value of each pixel on these elevations is defined as

$$
\begin{aligned}
&f_1(m,n) = \mathbf{max}\{(65-h)Voxel(m,n,h) \,|\, 1 \le h \le 64\}, \text{ for } 1 \le m, n \le 64,\\
&f_2(m,h) = \mathbf{max}\{(65-n)Voxel(m,n,h) \,|\, 1 \le n \le 64\}, \text{ for } 1 \le m, h \le 64,\\
&f_3(n,h) = \mathbf{max}\{mVoxel(m,n,h) \,|\, 1 \le m \le 64\}, \text{ for } 1 \le n, h \le 64,\\
&f_4(m,n) = \mathbf{max}\{hVoxel(m,n,h) \,|\, 1 \le h \le 64\}, \text{ for } 1 \le m, n \le 64,\\
&f_5(m,h) = \mathbf{max}\{nVoxel(m,n,h) \,|\, 1 \le n \le 64\}, \text{ for } 1 \le m, h \le 64,\\
&f_6(n,h) = \mathbf{max}\{(65-m)Voxel(m,n,h) \,|\, 1 \le m \le 64\}, \text{ for } 1 \le n, h \le 64.
\end{aligned}
$$

Fig. 2 shows the six elevations of three example 3D models. From these figures, we can see that the two 3D jeep models exhibit similar elevations, whereas the jeep and ship models differ.

3.2 Feature Extraction

To extract the elevation descriptor from these six elevations, each elevation is decomposed into L concentric circles around the center point (see Fig. 3). The region within the j-th concentric circle is denoted as C_j:

$$C_j = \left\{ (r,c) \,\middle|\, \sqrt{(r-L)^2 + (c-L)^2} < j \right\},$$

for $j = 1, 2, \ldots, 32$. For the k-th elevation, the sum of gray values of all pixels located within the j-th circle, C_j, is defined as

$$g_k(j) = \sum_{(r,c) \in C_j} f_k(r,c),$$

where $j = 1, 2, \ldots, 32$. Let $g_k(0) = 0$, the difference between the sums of gray values within two successive concentric circles is then derived:

$$d_k(j) = g_k(j) - g_k(j-1),$$

for $j = 1, 2, \ldots, 32$. Furthermore, every $d_k(j)$ value is normalized by using the following equation:

$$x_k(j) = \frac{d_k(j)}{\sum_{k=1}^{6} D(k)},$$

where $D(k)$ is the sum of all $d_k(j)$ values for the k-th elevation:

$$D(k) = \sum_{j=1}^{32} d_k(j).$$

The elevation descriptor $\mathbf{x}$ is defined as

$$\mathbf{x} = [(\mathbf{x}_1)^T, (\mathbf{x}_2)^T, \ldots, (\mathbf{x}_6)^T]^T,$$

where

$$\mathbf{x}_k = [x_k(1), x_k(2), \ldots, x_k(32)]^T.$$

Fig. 4 shows the elevation descriptors for the three 3D models shown in Fig. 2. It is evident that these two jeep models exhibit similar elevation descriptors whereas the jeep and ship models have totally different ones.

In general, the elevation descriptor is less sensitive to rotation if a 3D model is rotated by a small degree. Assume a 3D model is rotated by a small degree θ (see Fig. 5), the increment/decrement Δn of the altitude value of a voxel located at radius j is:

$$\Delta n = j \tan\theta.$$

However, the altitude value of a voxel located on the other side will decrease/increase the same value Δn. On average, the sum of the gray values on the j-th concentric circle of the rotated elevation is similar to the original one (see Fig. 6).

3.3 Similarity Computations

Since each 3D model is represented by six elevations, it requires 720 (6!) elevation matching operations to compute the similarity between two models. To reduce the matching time, an

efficient similarity computation is provided to find the best match for a given query model.

The matching operations can be greatly reduced if the relative positions of the elevations are taken into account. In practice, the front elevation $\boldsymbol{E}_1$ and the rear elevation $\boldsymbol{E}_4$ locate on opposite sides. Similarly, the right elevation $\boldsymbol{E}_3$ and the left elevation $\boldsymbol{E}_6$ as well as the top elevation $\boldsymbol{E}_2$ and the bottom elevation $\boldsymbol{E}_5$ also locate on opposite sides. Let the six elevations of the query model q and the matching model s be respectively defined as $\boldsymbol{E}_k^q$ and $\boldsymbol{E}_k^s$, for $k = 1, 2, \ldots, 6$. The six elevations of a query model q can be divided into three pairs: $(\boldsymbol{E}_1^q, \boldsymbol{E}_4^q)$, $(\boldsymbol{E}_2^q, \boldsymbol{E}_5^q)$, and $(\boldsymbol{E}_3^q, \boldsymbol{E}_6^q)$. Similarly, the six elevations of a matching model s can be divided into three pairs: $(\boldsymbol{E}_1^s, \boldsymbol{E}_4^s)$, $(\boldsymbol{E}_2^s, \boldsymbol{E}_5^s)$, and $(\boldsymbol{E}_3^s, \boldsymbol{E}_6^s)$. To calculate the difference between q and s, if $\boldsymbol{E}_1^q$ matches $\boldsymbol{E}_i^s$, $\boldsymbol{E}_4^q$ must match $\boldsymbol{E}_{[(i+2) \bmod 6]+1}^s$ according to the topological relationship between $\boldsymbol{E}_1^q$ and $\boldsymbol{E}_4^q$. Similarity, if $\boldsymbol{E}_2^q$ matches $\boldsymbol{E}_i^s$, $\boldsymbol{E}_5^q$ must match $\boldsymbol{E}_{[(i+2) \bmod 6]+1}^s$, and if $\boldsymbol{E}_3^q$ matches $\boldsymbol{E}_i^s$, $\boldsymbol{E}_6^q$ must match $\boldsymbol{E}_{[(i+2) \bmod 6]+1}^s$. In summary, the number of elevation matching operations that need to be performed is $3! \times 2^3 = 48$, instead of 720 matching operations. Table 1 lists these 48 matching operations. In this table, for i-th permutation p_i, $\boldsymbol{E}_k^q$ will match $\boldsymbol{E}_{p_i(k)}^s, 1 \le k \le 6$.

Let $\mathbf{x} = [(\mathbf{x}_1)^T, (\mathbf{x}_2)^T, \ldots, (\mathbf{x}_6)^T]^T$ and $\mathbf{y} = [(\mathbf{y}_1)^T, (\mathbf{y}_2)^T, \ldots, (\mathbf{y}_6)^T]^T$ denote the elevation descriptors of q and s, respectively. For the matching operation corresponding to the i-th permutation p_i, $1 \le i \le 48$, the distance between $\mathbf{x}$ and $\mathbf{y}$ is defined as:

$$Dis_{q,s}^i = \sum_{k=1}^{6} \left\| \mathbf{x}_k - \mathbf{y}_{p_i(k)} \right\|_1 = \sum_{k=1}^{6} \sum_{r=1}^{32} \left| x_k(r) - y_{p_i(k)}(r) \right|,$$

where $p_i(k)$ denotes the k-th value for the i-th permutation, $1 \le k \le 6$. The distance between the query model q and the matching model s is defined as

$$Dis_{q,s} = \min_{1 \le i \le 48} Dis_{q,s}^i.$$

Then, the similarity measure between q and s is defined as the inverse of the distance:

$$Sim_{q,s} = \frac{1}{Dis_{q,s}}.$$

Note that the larger the similarity value, the more similar a matching model is to the query. Therefore, the retrieved models similar to a query can be determined by taking those with larger similarity values.

4. Experimental Results

To demonstrate the effectiveness of the proposed elevation descriptor for different 3D models, experiments have been conducted on two test databases. Three other features, including spherical harmonics (SH) [5], the MPEG-7 3D shape spectrum descriptor (SSD) [16], and D2 [10], are implemented to compare the retrieval results. The performance is measured by **recall** and **precision**. The recall value, *Re*, and the precision value, *Pr*, are defined by the following equations:

$$Re = N / T ,$$

and

$$Pr = N / K ,$$

where N is the number of relevant models retrieved, T is the total number of relevant models in the database, and K is the total number of retrieved models.

4.1 Experiment on Database 1

Database 1 is established to test the performance of invariance to deformations. To derive Database 1, 20 models are selected as the seed models. Then, each seed model is deformed by 14 kinds of transformations, including 4 geometric deformations, 2 scalings, 3 rotations, and 5 various resolutions (see Fig. 7). Thus, there are in total 300 models in Database 1.

Fig. 8 shows some 3D models and their deformed models representing geometric deformation, rotation, scaling, and various resolution as well as the corresponding elevation descriptors. The similarities between these pairs of models are 0.9702, 0.9304, 0.9999, and 0.9691, respectively. We

can see that even though the model is deformed, the similarity value is still large enough.

In the previous section, it is shown that the elevation descriptor of a rotated 3D model is similar to the original if the rotation degree is small. In our simulation results, the elevation descriptor is also robust if a 3D model is rotated by a large degree. Table 2 shows the distance between a rotated query model and all seed models in Database 1. The query model q is a dragon, the seed model of class 3 on Database 1 (see Fig. 7). The query dragon model is rotated about the m-axis and n-axis by different degrees θ_m and θ_n : 20°, 40°, 60°, and 80°. We can see that even though the dragon model is rotated by various degrees, the one with the smallest distance is still the original dragon model. That is, the proposed elevation descriptor is robust to rotations.

In our experiments, each model in Database 1 is presented as a query. Table 3 shows the average recall values for all query models using the proposed elevation descriptor (ED), spherical harmonics (SH), 3D shape spectrum descriptor (SSD), and D2. From Table 3, we can see that the elevation descriptor outperforms other descriptors. The detailed comparison of the average recall value for each class is shown in Fig. 9. The elevation descriptor has the best performance for most classes. To see what kind of deformations will dramatically affect the retrieval result, a detailed performance comparison for each kind of deformation is shown in Fig. 10. From this figure, we can see that the elevation descriptor always get the best performance.

4.2 Experiment on Database 2

The second database, Database 2, is derived from the Princeton Shape Benchmark database [22] which contains 1814 models (161 classes) and is used for evaluating shape based retrieval and analysis algorithms. Note that in this database each class contains a different number of models. The 22 classes that have the largest number of models are selected as queries. Each of them contains at least 15 models. These 22 classes are shown in Fig. 11.

The performance is also measured by **recall** and **precision**. Since the number of models in each class is different, the recall value and the precision value for the j-th query model within the i-th

class are defined as:

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

and

$$Pr_i^j = N_i^j / K,$$

where N_i^j is the number of relevant models retrieved, T_i is the total number of relevant models in the database, and K is the total number of retrieved models. The average recall and precision are defined by the following equations:

$$Re = \frac{1}{T_S} \sum_{i=1}^{22} \sum_{j=1}^{T_i} Re_i^j ,$$

and

$$Pr = \frac{1}{T_S} \sum_{i=1}^{22} \sum_{j=1}^{T_i} Pr_i^j ,$$

where $T_S = T_1 + T_2 + \cdots + T_{22}$. The overall performance of the proposed method is still better than others (see Table 4 and Fig. 12). Table 5 compares the average query time. We can see that the query time when using ED is slightly larger than when using SH, but the retrieval accuracy is much better than SH.

5. Conclusions

In this paper, a novel descriptor, called elevation descriptor (ED), for 3D model retrieval is proposed. First, a 3D model is represented with six gray-level images which describe the altitude information of 2D projections from six different views including front, left, right, rear, top and bottom. Each gray-level image, called an elevation, is then decomposed into a set of concentric circles. The sum of the altitude information within each concentric circle is calculated. The elevation descriptor is obtained from the difference of the altitude sums between two successive concentric circles. Since there are six elevations, an efficient similarity matching method is provided to find the best match for a given query model without exhaustively matching all possible 720(6!) elevation permutations. The experimental results show that for most types of 3D models,

the proposed ED outperforms other descriptors including spherical harmonics (SH), the MPEG-7 3D shape spectrum descriptor (SSD), and D2.

Acknowledgments

This research was supported in part by the National Science Council, Taiwan under Contract NSC 94-2213-E-216-018 and Chung Hua University, Taiwan under Contract CHU-94-TR-002.

References

[1] J.W.H. Tangelder, R.C. Veltkamp, A survey of content based 3D shape retrieval methods, Proceedings of the Shape Modeling Applications, 2004, pp. 145-156.

[2] C. Zhang, T. Chen, Efficient feature extraction for 2D/3D objects in mesh representation, Proceedings of IEEE International Conference on Image Processing (ICIP), Thessaloniki, Greece, 2001, pp. 935-938.

[3] E. Paquet, A. Murching, T. Naveen, A. Tabatabai, M. Rioux, Description of shape information for 2-D and 3-D objects, Signal Processing: Image Communication 16 (2000) 103–122.

[4] D.V. Vranic, D. Saupe, J. RICHTER, Tools for 3D-object retrieval: Karhunen-Loeve transform and spherical harmonics, Proceedings of the IEEE Workshop on Multimedia Signal Processing, 2001, pp. 293-298.

[5] T. Funkhouser, P. Min, M. Kazhdan, J. Chen, A. Halderman, D. Dobkin, D. Jacobs, A search engine for 3D models, ACM Trans. Graphics 22 (1) (2003) 83-105.

[6] M. Kazhdan, T. Funkhouser, S. Rusinkiewicz, Rotation invariant spherical harmonic representation of 3D shape descriptors, Symposium on Geometry Processing, 2003.

[7] M. Novotni, R. Klein, Shape retrieval using 3D Zernike descriptors, Computer-Aided Design 36 (2004) 1047-1062.

[8] M. Yu, I. Atmosukarto, W.K. Leow, Z. Huang, R. Xu, 3D model retrieval with morphing-based geometric and topological feature maps, Proceedings of Computer Vision

and Pattern Recognition, 2003, pp. 656-661.

[9] M. Ankerst, G. Kastenmuller, H.P. Kriegel, T. Seidl, 3D shape histograms for similarity search and classification in spatial databases, Symposium on Large Spatial Databases, 1999, pp. 207–226.

[10] R. Osada, T. Funkhouser, B. Chazelle, D. Dobkin, Shape distributions, ACM Trans. on Graphics 21 (4) (2002) 807-832.

[11] C.Y. Ip, D. Lapadat, L. Sieger, W.C. Regli, Using shape distributions to compare solid models, Proceedings of Solid Modeling, 2002, pp. 273-280.

[12] C.Y. Ip, L.Sieger, W.C. Regli, A. Shokoufandeh, Automated learning of model classifications, Proceedings of Solid Modeling, 2003, pp. 322-327.

[13] R. Ohbuchi, T. Minamitani, T. Takei, Shape-similarity search of 3D models by using enhanced shape functions, Proceedings of Theory and Practice of Computer Graphics, 2003, pp. 97-104.

[14] R. Ohbuchi, T. Takei, Shape-similarity comparison of 3D models using alpha shapes, Proceedings of Pacific Conf. on Computer Graphics and Applications, 2003.

[15] J.-L. Shih, C.-H. Lee, J.T. Wang, 3D object retrieval system based on grid D2, Electronics Letters 41 (4) (2005) 23-24.

[16] MPEG Video Group, MPEG-7 Visual part of eXperimetation Model Version 9.0, Doc. ISO/IEC JTC1/SC29/WG11/N3914, Pisa, January 2001.

[17] M. Hilaga, Y. Shinagawa, T. Kohmura, T.L. Kunii, Topology matching for fully automatic similarity estimation of 3D shapes, Proceedings of SIGGRAPH, 2001, pp. 203-212.

[18] D. Bespalov, A. Shokoufandeh, W.C. Regli, Reeb graph based shape retrieval for CAD, Proceedings of ASME Design Engineering Technical Conferences, 2003.

[19] J. Loffler, Content-based retrieval of 3D models in distributed web databases by visual shape information, Proceedings of Information Visualization, 2000, pp. 82-87.

[20] B.J. Super, H. Lu, Evaluation of a hypothesizer for silhouette-based 3-D object recognition,

Pattern Recognition 36 (2003) 69-78.

[21] D.Y. Chen, X.P. Tian, Y.T. Shen, M. Ouhyoung, On visual similarity based 3D model retrieval, Computer Graphics Forum 22 (3) (2003) 223-232.

[22] P. Shilane, P. Min, M. Kazhdan, T. Funkhouser, The Princeton shape benchmark, Proceedings of the Shape Modeling Applications, 2004, pp. 167-178.

Table 1 48 permutations for elevation matching between a query model and a matching model.

k	1	2	3	4	5	6	k	1	2	3	4	5	6	k	1	2	3	4	5	6
$p_1(k)$	1	2	3	4	5	6	$p_{17}(k)$	2	1	3	5	4	6	$p_{33}(k)$	3	2	1	6	5	4
$p_2(k)$	1	3	2	4	6	5	$p_{18}(k)$	2	3	1	5	6	4	$p_{34}(k)$	3	1	2	6	4	5
$p_3(k)$	1	5	3	4	2	6	$p_{19}(k)$	2	4	3	5	1	6	$p_{35}(k)$	3	5	1	6	2	4
$p_4(k)$	1	3	5	4	6	2	$p_{20}(k)$	2	3	4	5	6	1	$p_{36}(k)$	3	1	5	6	4	2
$p_5(k)$	1	2	6	4	5	3	$p_{21}(k)$	2	1	6	5	4	3	$p_{37}(k)$	3	2	4	6	5	1
$p_6(k)$	1	6	2	4	3	5	$p_{22}(k)$	2	6	1	5	3	4	$p_{38}(k)$	3	4	2	6	1	5
$p_7(k)$	1	5	6	4	2	3	$p_{23}(k)$	2	4	6	5	1	3	$p_{39}(k)$	3	5	4	6	2	1
$p_8(k)$	1	6	5	4	3	2	$p_{24}(k)$	2	6	4	5	3	1	$p_{40}(k)$	3	4	5	6	1	2
$p_9(k)$	4	2	3	1	5	6	$p_{25}(k)$	5	1	3	2	4	6	$p_{41}(k)$	6	2	1	3	5	4
$p_{10}(k)$	4	3	2	1	6	5	$p_{26}(k)$	5	3	1	2	6	4	$p_{42}(k)$	6	1	2	3	4	5
$p_{11}(k)$	4	5	3	1	2	6	$p_{27}(k)$	5	4	3	2	1	6	$p_{43}(k)$	6	5	1	3	2	4
$p_{12}(k)$	4	3	5	1	6	2	$p_{28}(k)$	5	3	4	2	6	1	$p_{44}(k)$	6	1	5	3	4	2
$p_{13}(k)$	4	2	6	1	5	3	$p_{29}(k)$	5	1	6	2	4	3	$p_{45}(k)$	6	2	4	3	5	1
$p_{14}(k)$	4	6	2	1	3	5	$p_{30}(k)$	5	6	1	2	3	4	$p_{46}(k)$	6	4	2	3	1	5
$p_{15}(k)$	4	5	6	1	2	3	$p_{31}(k)$	5	4	6	2	1	3	$p_{47}(k)$	6	5	4	3	2	1
$p_{16}(k)$	4	6	5	1	3	2	$p_{32}(k)$	5	6	4	2	3	1	$p_{48}(k)$	6	4	5	3	1	2

Table 2 Distance between a rotated query model and all 3D seed models in Database 1. The rotated query model q is derived from the seed model of class 3 in Database 1 with rotation about the m-axis and n-axis by degrees θ_m and θ_n to be 20°, 40°, 60°, and 80°. The values shown in this table are multiplied by 10000.

Rotated degree		The class number of matching model s																			
θ_m	θ_n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
0	0	55	32	0	42	36	44	26	41	42	40	44	61	30	31	32	29	28	30	35	33
0	20	57	33	6	43	34	45	25	45	47	44	43	63	33	31	29	30	28	30	33	35
0	40	58	35	9	44	35	48	24	44	48	45	45	66	34	34	29	29	30	29	34	36
0	60	56	35	12	45	39	50	23	41	46	43	49	67	33	38	33	29	33	31	36	35
0	80	55	32	12	45	40	47	23	41	44	41	51	65	31	38	34	28	32	32	37	33
20	0	58	34	8	41	36	46	29	44	47	44	45	64	33	33	32	31	28	31	36	37
20	20	59	34	10	41	36	46	29	46	49	45	45	64	36	33	33	31	27	31	35	38
20	40	60	35	12	43	35	48	27	47	50	46	48	66	37	35	33	30	27	30	35	40
20	60	59	37	13	44	38	51	25	46	50	47	49	68	36	38	33	32	30	31	36	38
20	80	56	34	14	43	40	49	26	43	47	44	52	65	34	39	35	30	32	32	38	34
40	0	63	36	14	40	36	47	32	46	51	47	49	64	36	38	36	34	28	35	37	42
40	20	63	35	14	40	35	46	34	48	53	49	48	63	38	35	35	36	27	35	37	43
40	40	66	37	17	41	34	47	34	51	55	51	49	64	41	37	35	35	26	35	36	46
40	60	65	38	17	42	35	49	31	50	55	52	50	66	40	37	35	37	27	35	36	44
40	80	62	36	17	41	38	48	30	47	52	48	53	66	37	40	37	34	30	36	39	41
60	0	61	34	11	40	35	46	31	45	50	47	47	64	35	36	33	33	27	32	37	40
60	20	66	38	17	39	38	49	36	49	56	54	51	65	40	39	37	38	29	37	39	47
60	40	69	40	21	39	34	49	39	51	56	55	50	64	41	42	36	39	27	37	37	50
60	60	69	41	22	40	33	49	40	53	58	56	51	64	42	42	37	40	27	37	38	51
60	80	64	38	16	39	34	47	34	49	53	50	48	64	38	36	36	36	28	34	37	44
80	0	57	33	5	41	35	45	28	41	44	43	44	63	30	32	31	31	29	31	35	36
80	20	64	39	15	39	38	49	37	45	52	51	49	65	37	36	37	36	30	37	38	44
80	40	69	41	20	39	34	49	41	51	56	54	50	63	41	42	37	40	26	38	37	49
80	60	67	41	22	37	32	46	43	52	57	55	48	60	43	42	38	39	26	34	36	50
80	80	64	39	18	37	32	45	39	49	53	51	45	61	38	37	37	36	26	33	36	46

Table 3 Comparison of the retrieval results of the proposed ED with other descriptors on Database 1.

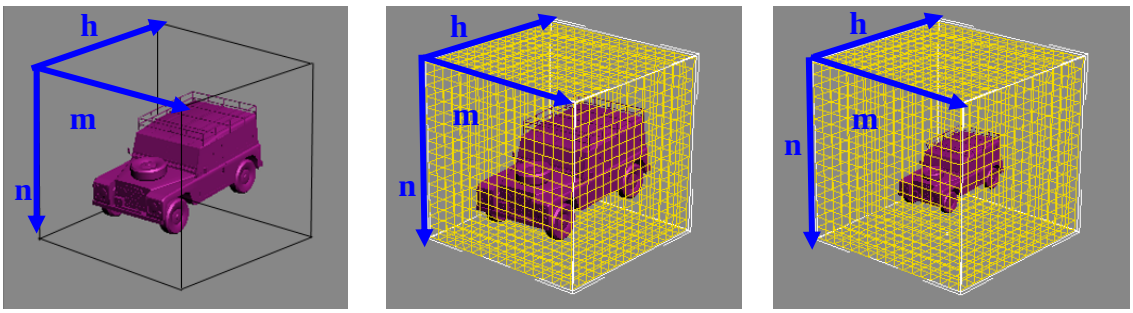
	Re (*K=15*)	*Re* (*K=30*)
ED	0.9722	0.9869
SH	0.9391	0.9700
SSD	0.8840	0.9358
D2	0.8733	0.9222

Table 4 Overall performance of the proposed ED method and other methods on Database 2.

	$Re\ (K = T_i)$	$Re\ (K = 2T_i)$	$Re\ (K = 3T_i)$	$Re\ (K = 4T_i)$
ED	0.3370	0.4364	0.4945	0.5371
SH	0.2451	0.3159	0.3607	0.3931
SSD	0.2010	0.2395	0.2705	0.2977
D2	0.1745	0.2216	0.2596	0.2902

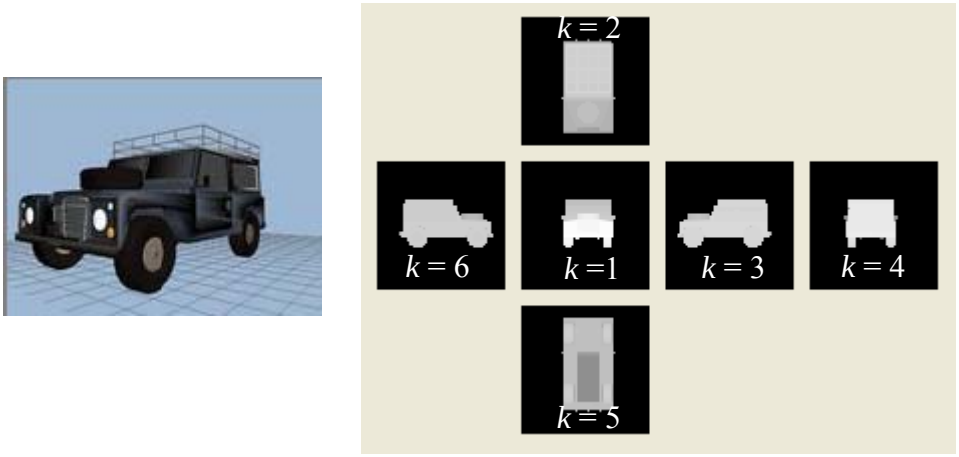
Table 5 Comparison of average query time of the proposed ED method and other methods in Database 2.

	query time(sec)
ED	3.148
SH	2.946
SSD	1.656
D2	1.661

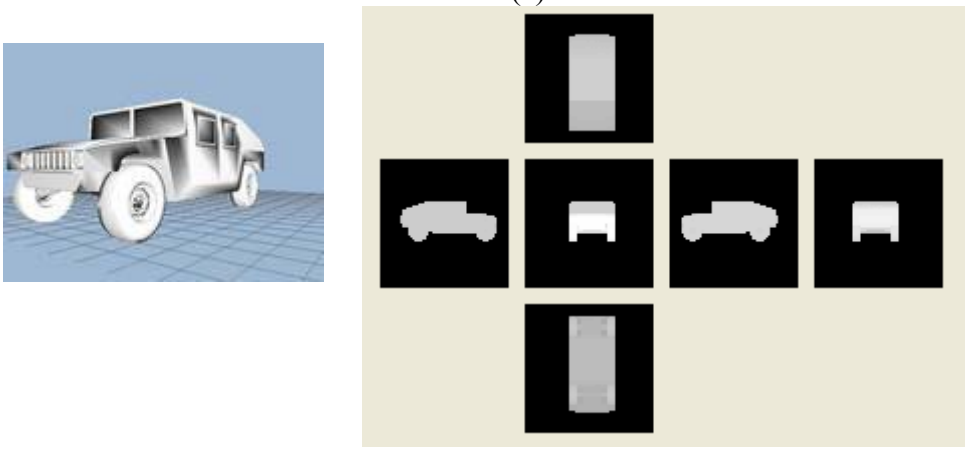


(a) (b) (c)

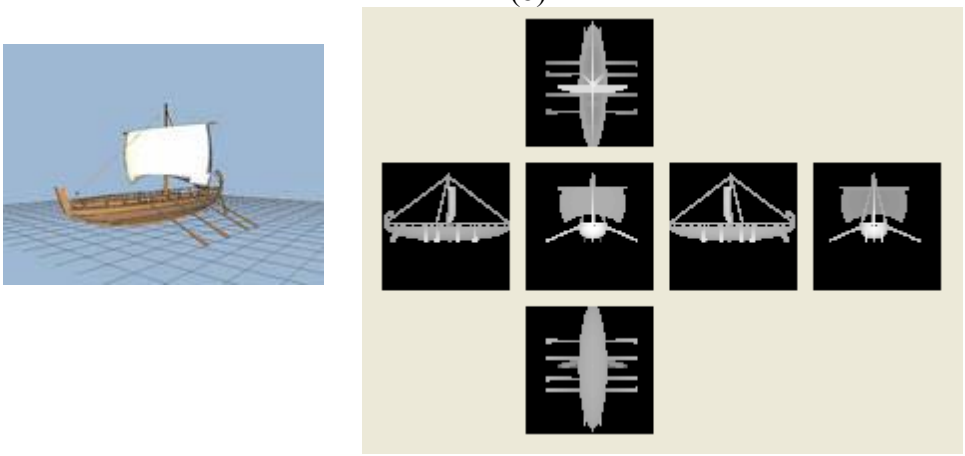
Fig. 1 Original and decomposed 3D jeep model. (a) The 3D jeep model circumscribed by a bounding box. (b) The bounding box of the 3D jeep model is decomposed into a $2L \times 2L \times 2L$ voxel grid. (c) The normalized 3D jeep model.



(a)



(b)



(c)

Fig. 2 3D models and their six elevations including front ($k=1$), top ($k=2$), right ($k=3$), rear ($k=4$), bottom ($k=5$), and left ($k=6$) elevations. (a) 3D jeep model and its six elevations (b) Another 3D jeep model and its six elevations. (c) 3D ship model and its six elevations.

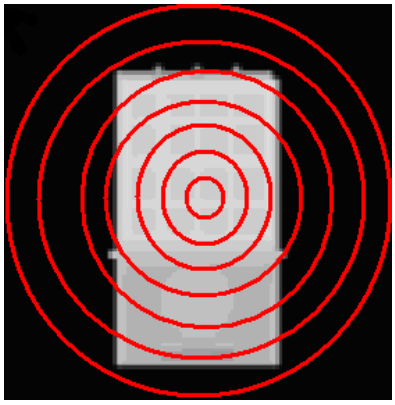


Fig. 3 Top elevation of the 3D jeep model shown in Fig. 2(a) segmented by several concentric circles.

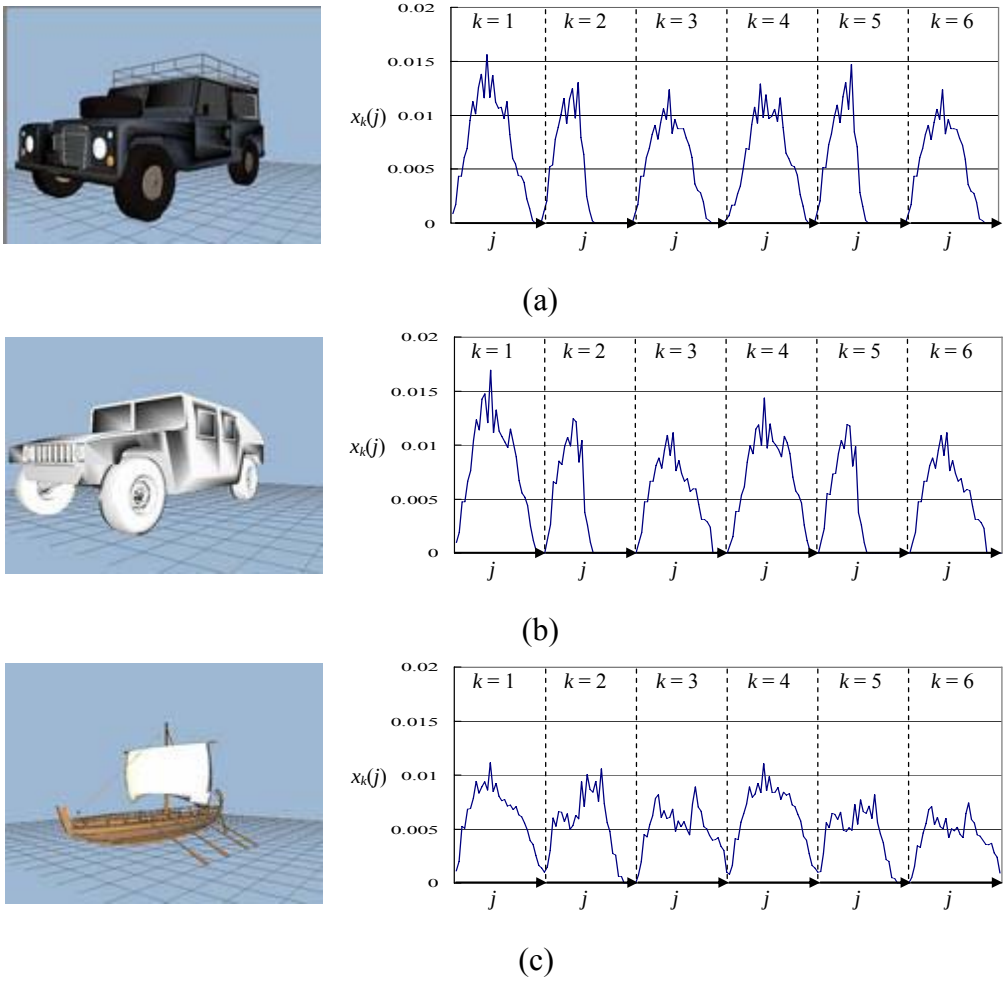


Fig. 4 Elevation descriptors for the three models shown in Fig. 2. The vertical axis represents $x_k(j)$. The horizontal axis represents j ($1 \le j \le 32$) for each $k = 1, 2, \ldots, 6$ successively. (a) The elevation descriptors for the jeep model shown in Fig. 2(a). (b) The elevation descriptors for another jeep model shown in Fig. 2(b). (c) The elevation descriptors for the ship model shown in Fig. 2(c).

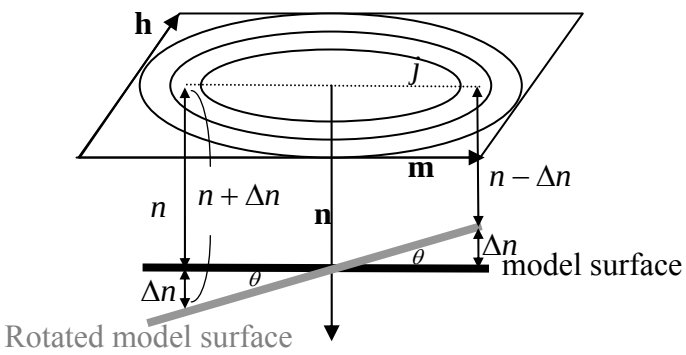
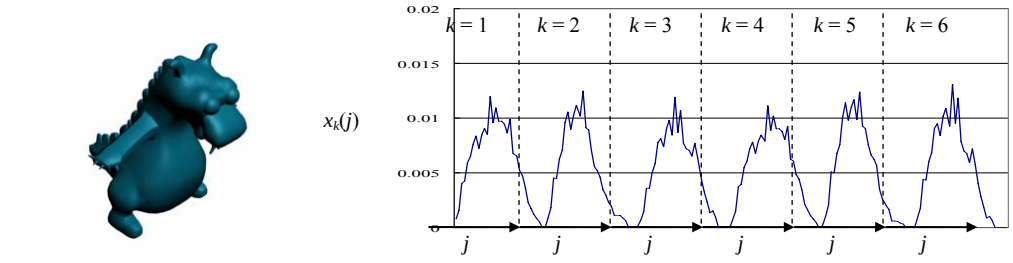
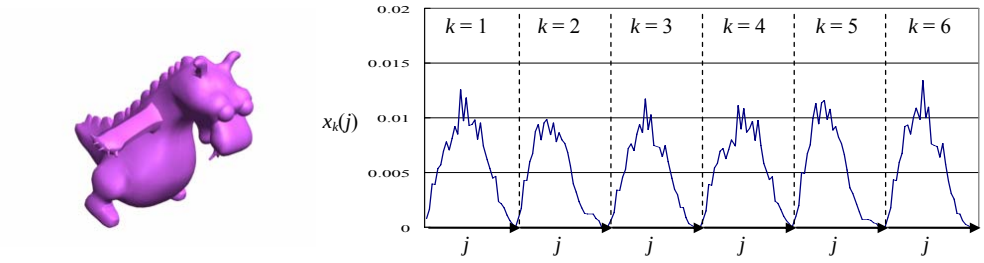


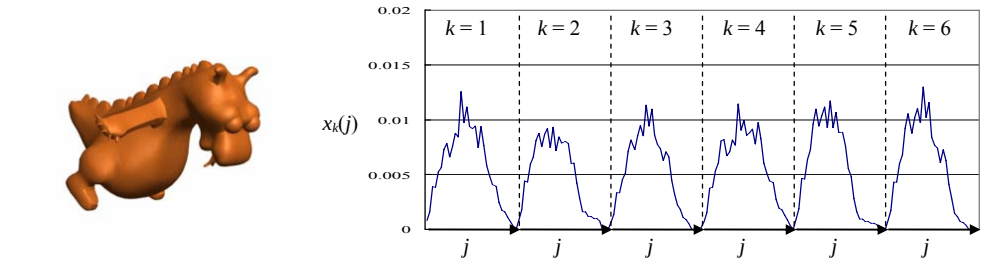
Fig. 5 3D model rotated by degree θ .



(a)



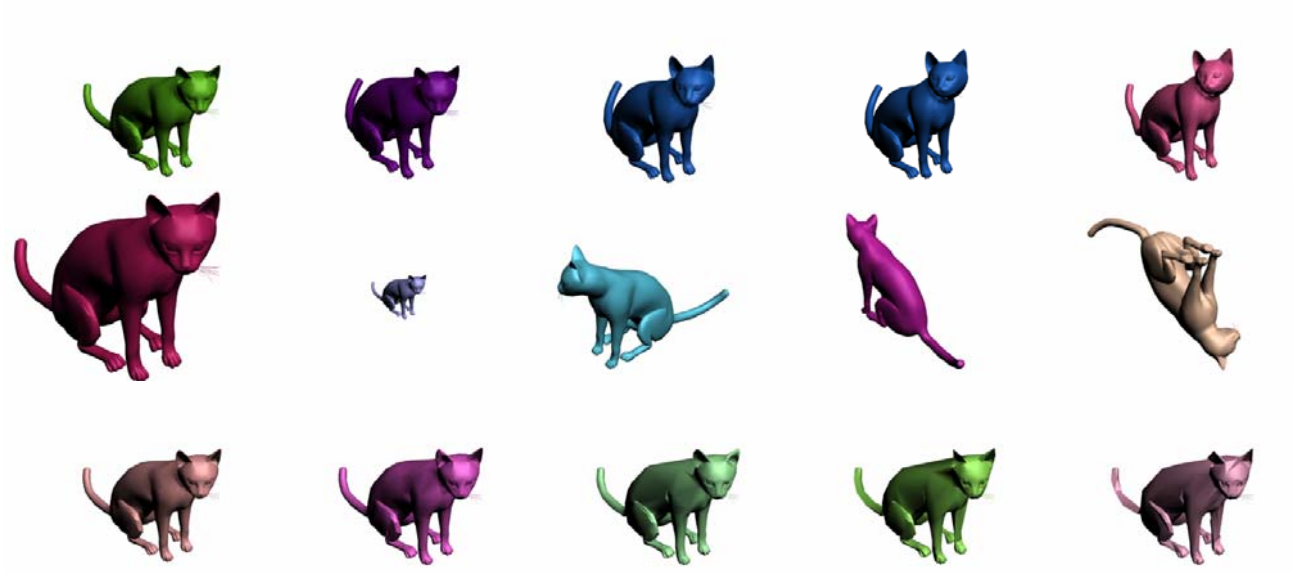
(b)



(c)

Fig. 6 Elevation descriptors for the rotated models. (a) The elevation descriptors for the dragon model. (b) The elevation descriptors for the rotated dragon model with rotation degrees $\theta_m = 10°$ and $\theta_n = 10°$. (c) The elevation descriptors for another rotated dragon with rotation degrees with $\theta_m = 20°$ and $\theta_n = 20°$.

(a)



(b)

Fig. 7 Database 1. (a) 20 seed models. (b) 15 deformed models for the cat class.

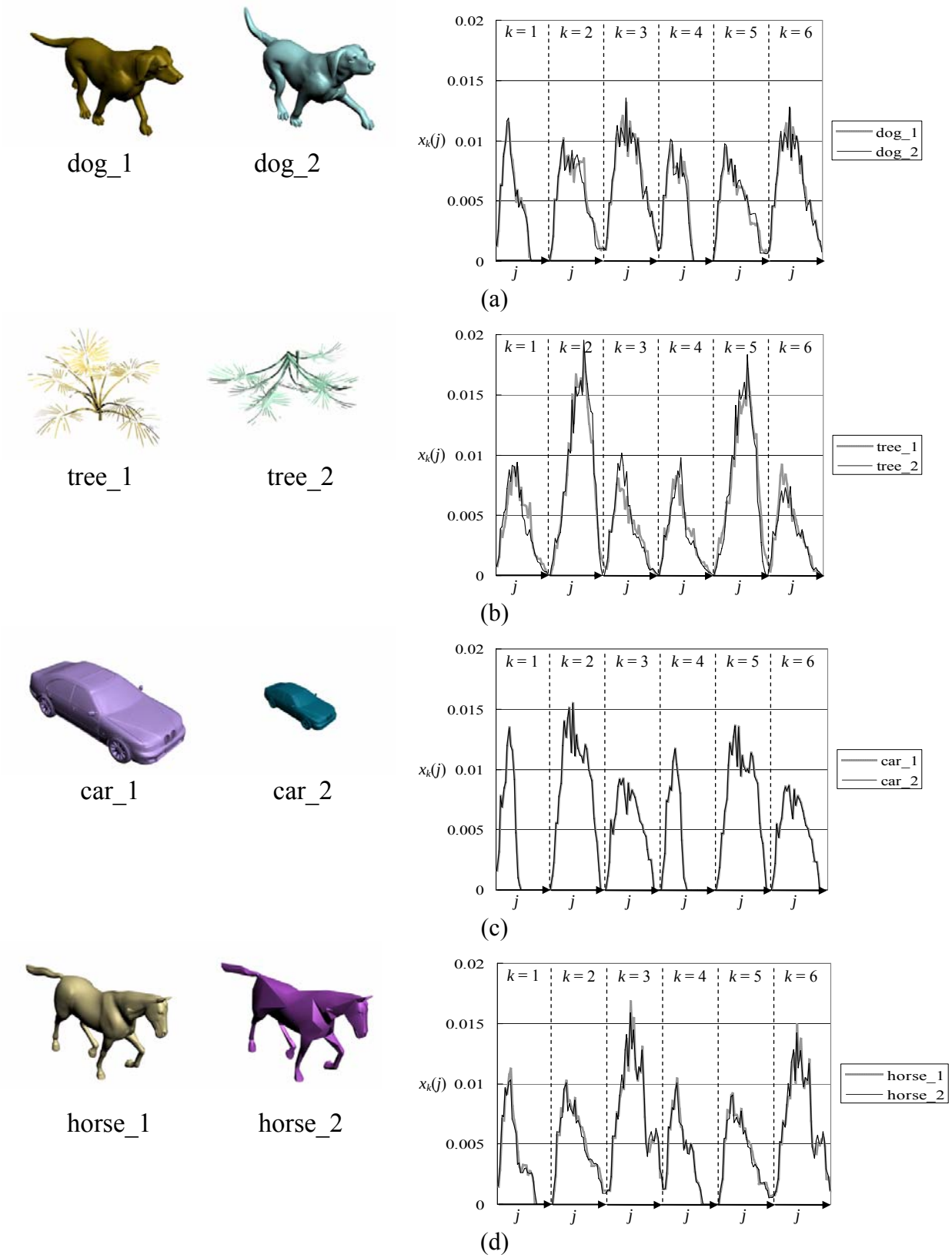


Fig. 8 Deformed 3D models and their corresponding elevation descriptors. (a) Geometric deformation (b) Rotation (c) Scaling (d) Various resolution.

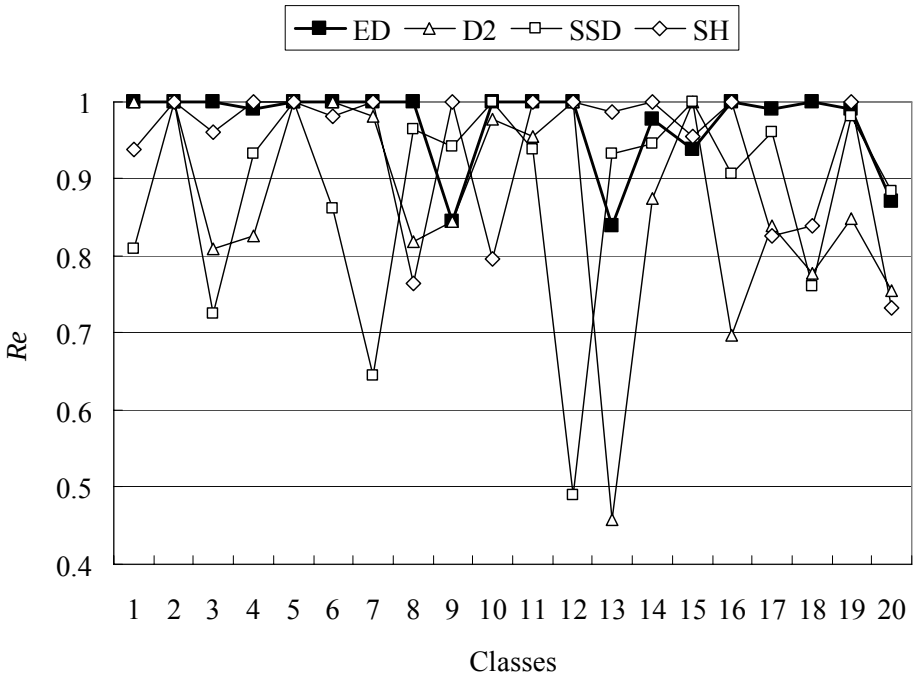


Fig. 9 Comparison of the average recall (*Re*) for each class on Database 1(K = 15).

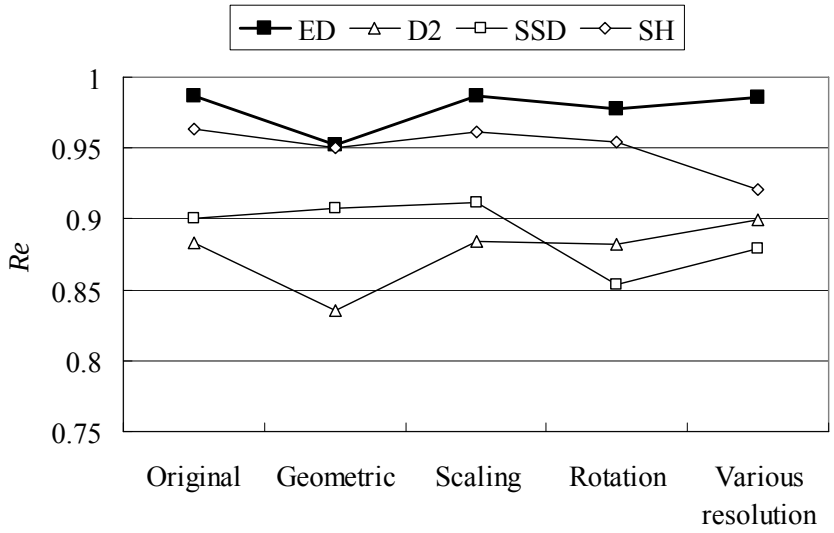
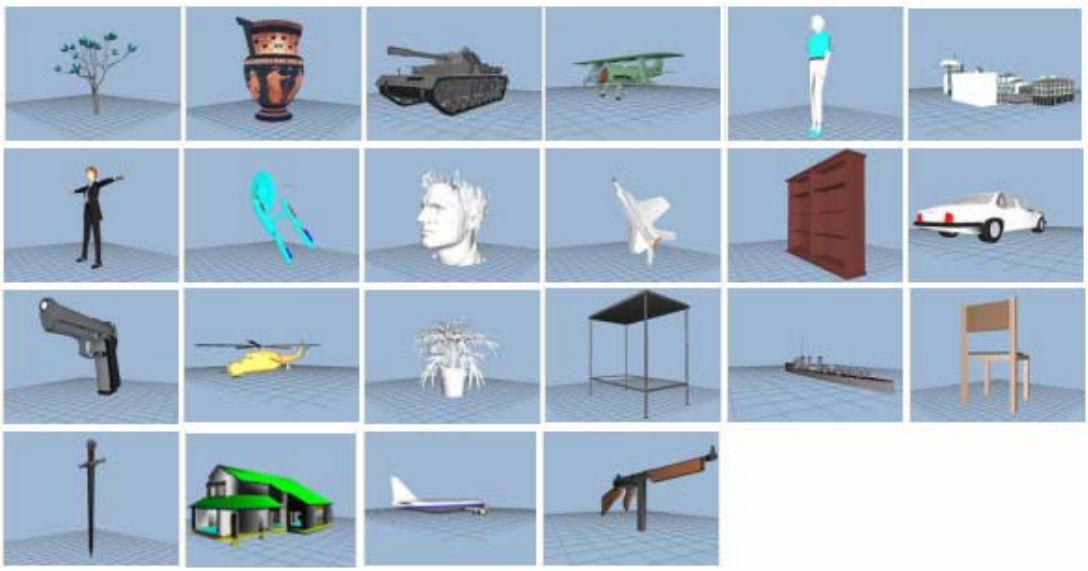


Fig. 10 Recall (*Re*) for various deformations in Database 1 (K = 15).



(a)

(b)

Fig. 11 Query models in Database 2 derived from the Princeton Shape Benchmark database. (a) The 22 query classes in Database 2**.** (b) All models belong to the tank class.

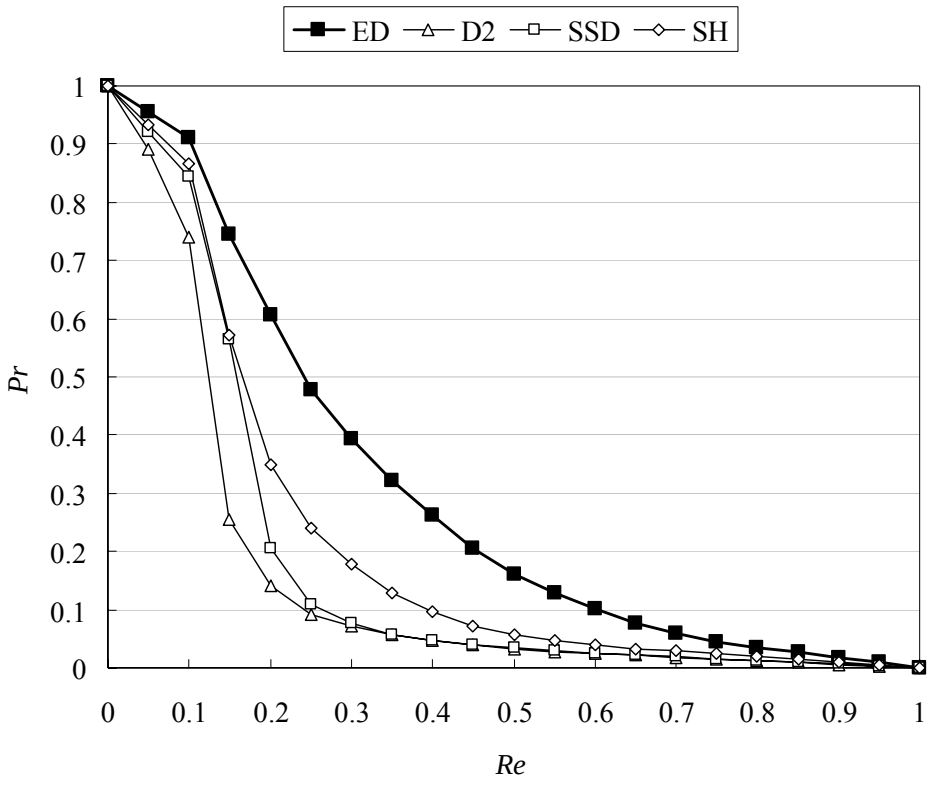


Fig. 12 Precision vs. recall curves on Database 2.