On Visual Similarity Based 3D Model Retrieval

Ming Ouhyoung, Ding-Yun Chen . Xiao-Pei Tian
Edward Yu-Te Shen .

Department of Computer Science and Information Engineering
National Taiwan University
November 18, 2003

1

Outline

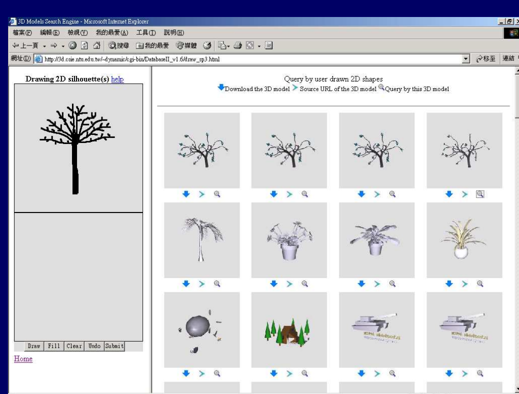
- Introduction
- Previous works
- Our proposed approach
- 3D shape search engine
- Experimental results
- Conclusion & future work

2

Introduction

Online demo

- 3D models in the database
- http://3d.csie.ntu.edu.tw



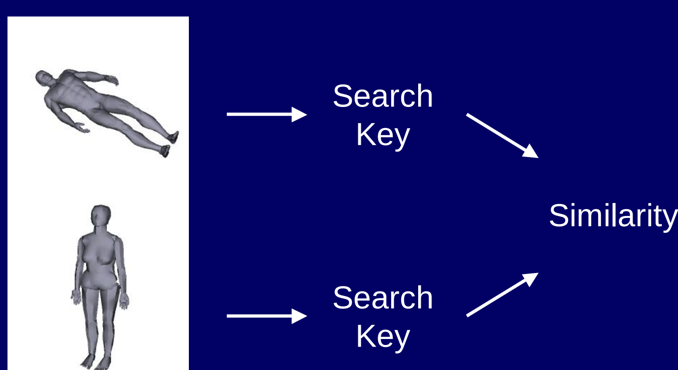
Motivation

- Content-based retrieval for multimedia has become important
- The lack of meaningful description for automatic matching
- Development of 3D modelling and digitizing technology
- Fast 3D model generation via modifying retrieved models: Virtual Sculptor DEMO

5

Problem statement

- Similarity between two given 3D objects
  - Features and distance metric



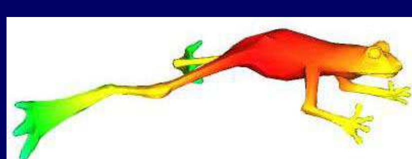
6

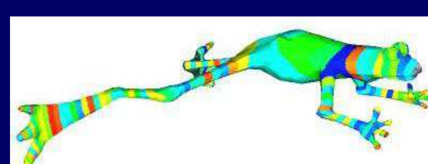
Previous Works

7

Topology matching
for fully automatic similarity estimation of 3D shapes

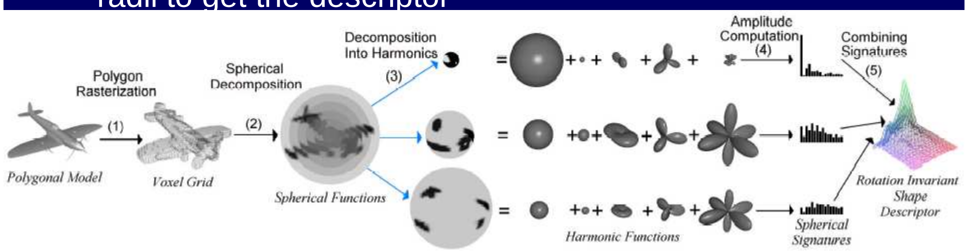
- M. Hilaga et al., *ACM SIGGRAPH*, 2001.
- Geometry-based and topology-based approach
- For each vertex of the object, calculate a sum of the geodesic distance from this vertex to others
- Reeb graph, where each node represents a region according to the value
- Construct multi-resolutional Reeb graph by grouping adjacent regions





Spherical harmonic descriptor

- Funkhouser et al., *ACM Trans. on Graphics*, Jan. 2003
- Geometry-based and shape-based approach
- Transform 3D model to concentric spheres
- Frequency components by spherical harmonics (Fourier transformation in spherical coordinates)
- Combining these different signatures over the different radii to get the descriptor



Overview of 3D model retrieval systems

- Geometry-based approach
  - Shape-based
  - Topology-based
- Projected image-based approach — Shape-based — Geometry alignment and then matching by image

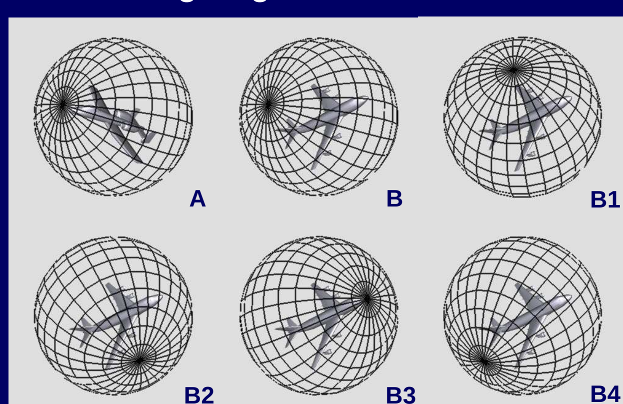
10

Our proposed approach

11

The main idea of our approach

- If two 3D models look similar, they look similar from all viewing angles



12

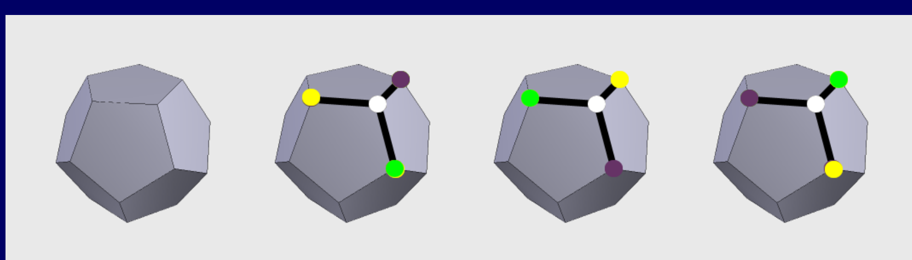
Design issues

- Correctness, discrimination
- Efficiency, automation
- Robustness
  - Similarity transformation: translation, rotation and scaling
  - Connectivity changes: re-meshing, sub-division and simplification
  - Model degeneracy: missing, wrongly oriented normals, intersecting, disjoint and overlapping polygons
  - Noise, deformation, etc.
- Scope
  - Static objects
  - Not for articulated creatures

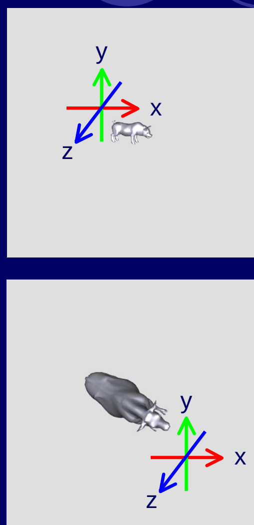
13

Cameras on a regular dodecahedron

- A dodecahedron is a polyhedron with 12 faces and 20 vertices
- A regular dodecahedron has most vertices from all regular convex polyhedrons
- 60 different rotations when matching (3 * 20 = 60)



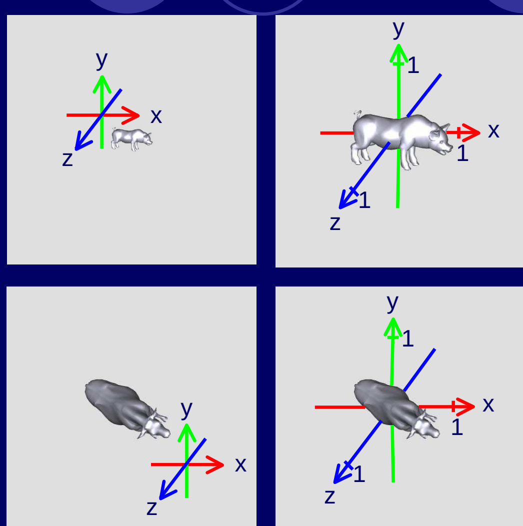
Matching between two 3D models (I)



- Two 3D models with different translation, rotation and scaling

15

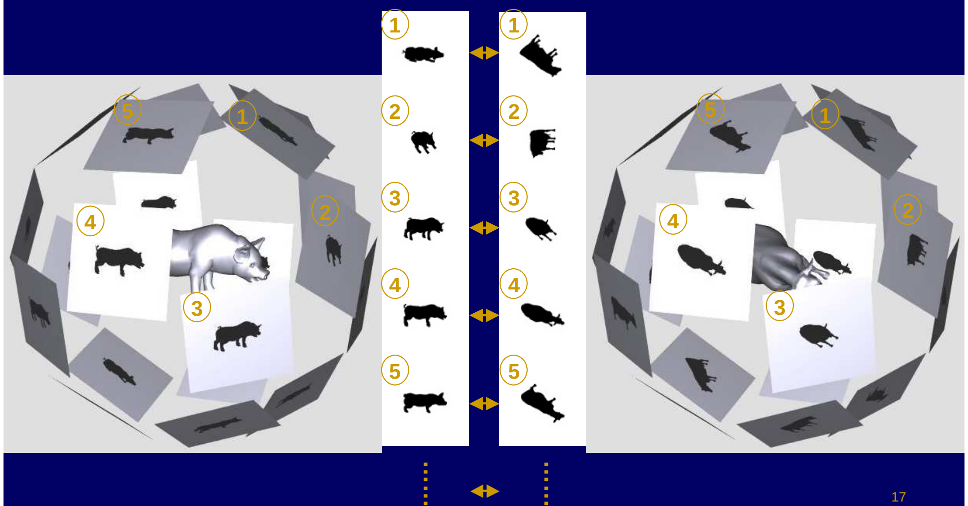
Matching between two 3D models (II)



- Translation and scaling roughly in 3D space
- The error in translation and scaling will be diminished by 2D image matching

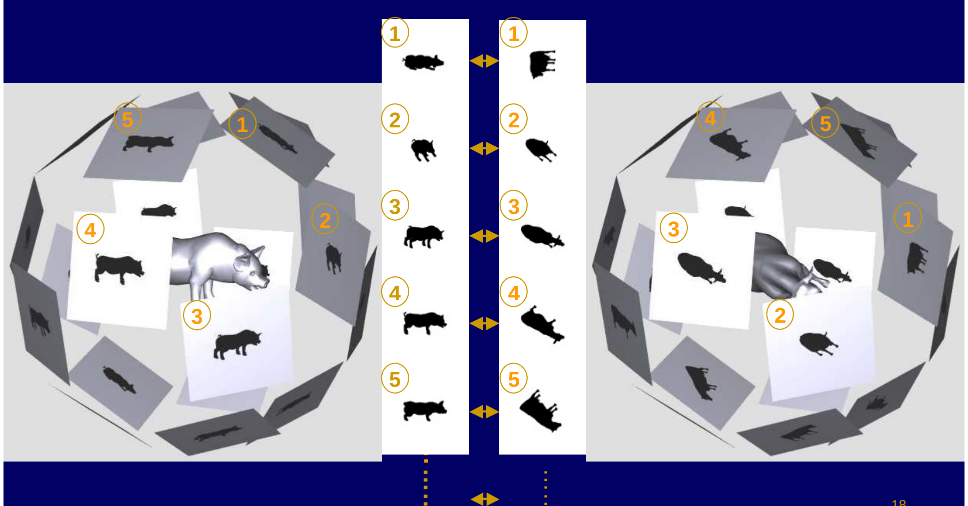
16

Matching between two 3D models (III)



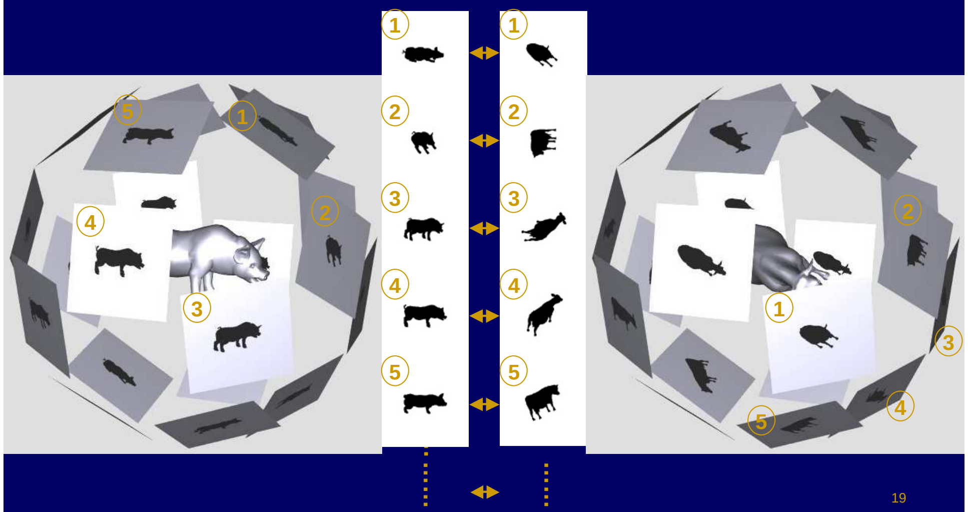
17

Matching between two 3D models (IV)

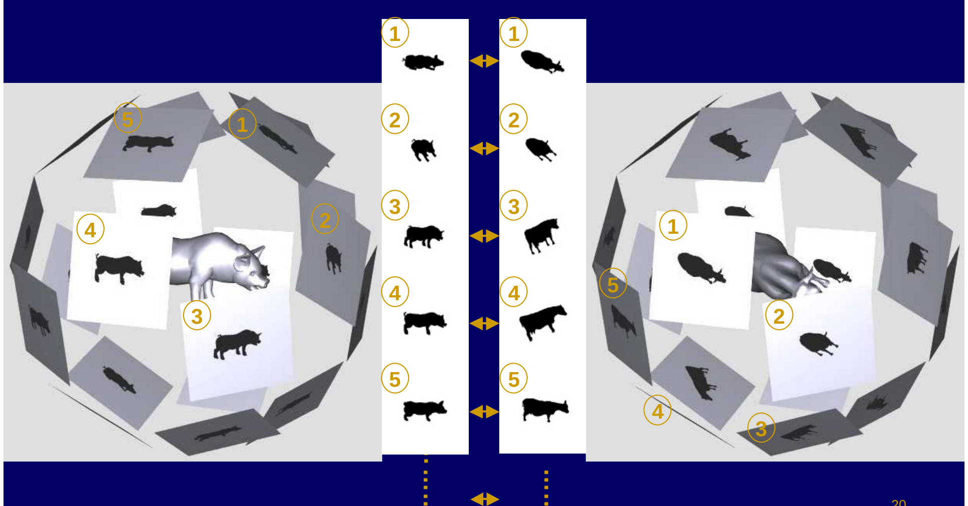


18

Matching between two 3D models (V)

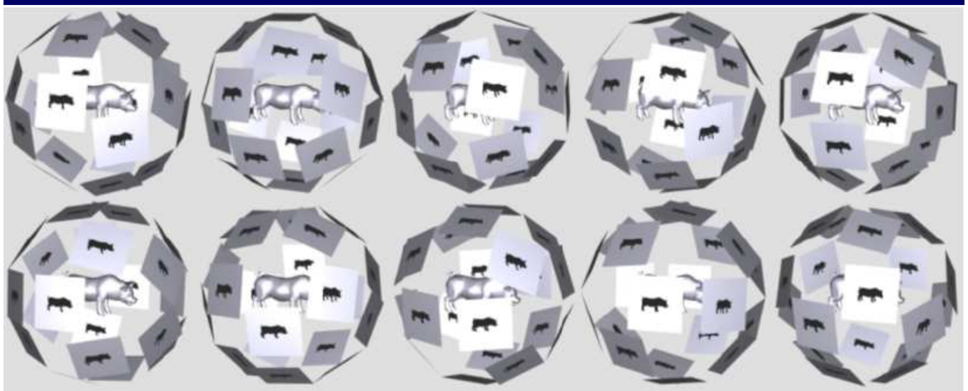


Matching between two 3D models (VI)



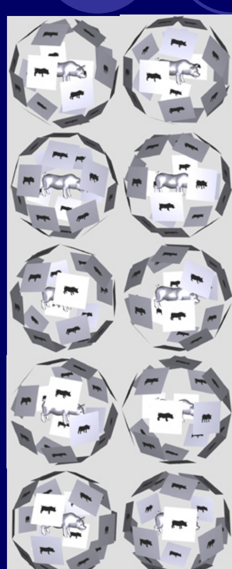
Minimum error from all corresponding 2D shapes

To be more robust in rotation(I)

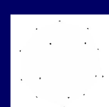


21

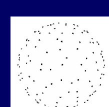
To be more robust in rotation(II)



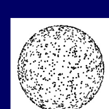
1-1
60



1-10
600

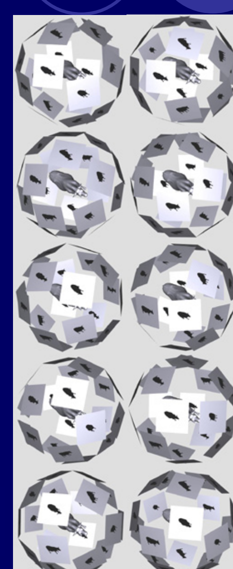


10-10
5460



Different rotation:
(Mx(N-1)+1)x60

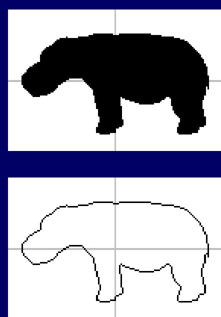
$$\frac{180^\circ}{x} \times \frac{360^\circ}{x} = 5460 \Rightarrow x \cong 3.4^\circ$$

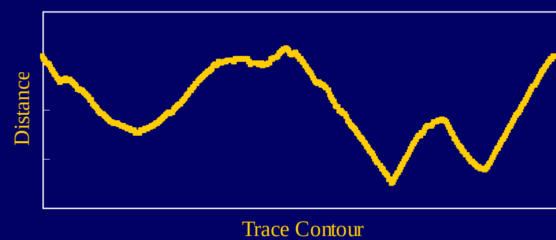


22

Image metrics

- Integrated region-based and contour-based approach
  - Zernike moment (35 coefficients)
  - Fourier Descriptor (10 coefficients)





- Robust from translation, rotation and scaling

23

2D Zernike Moments

- Teague proposed Zernike moments based on orthogonal Zernike polynomials
- Teague, J. Opt. Soc. Am. 70 (8), 1980
- Rotation invariant
- Khotanzad *et al*, IEEE PAMI, 12 (5), 1990
- Translation invariant
- Chong *et al* Pattern Recognition 36, 2003, pp.1765-1773

24

2D Zernike moments

Zernike moments

$$Z_{pq} = \frac{p+1}{\pi} \int_0^{2\pi} \int_0^1 V_{pq}^*(r,\theta) f(r,\theta) r\,dr\,d\theta$$

where $p - |q|$ is even and non-negative

25

Zernike polynomials of order p, with repetition q

$Vpq(r, \theta) = R_{pq}(r)\ e^{jq\theta}$

$R_{pq}(r) = \sum_{k=0}^{(p-|q|/2)} (-1)^k (p-k)!/\{k!((p+|q|/2-k)!(p-|q|)/2-k)!\}\ r^{p-2k}$

$R_{00}(r) = 1, \quad R_{11}(r) = r,$

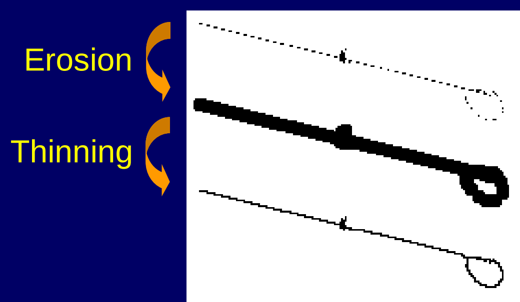
$R_{20}(r)= 2r^2-1, \quad R_{22}(r) = r^2$

$R_{31}(r) = 3r^3-2r, \quad R_{33}(r) = r^3$

26

Trace contour for rendered image

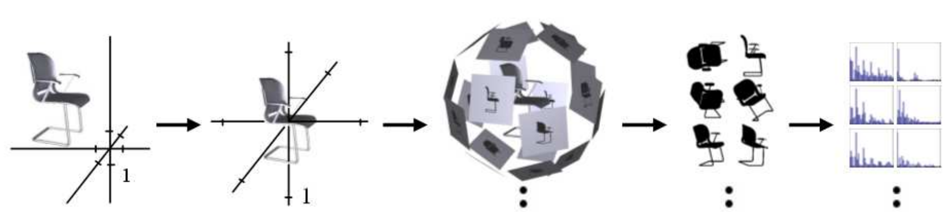
- Erosion operation to connect the separated parts
- Thinning, but don’t remove the rendered pixels



27

Steps Extracting the Descriptors

- Translation and scaling
- Render images from the camera positions
- Totally 100 images are rendered because of 10 Descriptors.
- Descriptors for a 3D model are extracted from the 100 images as Zernike Descriptors and Fourier Descriptors



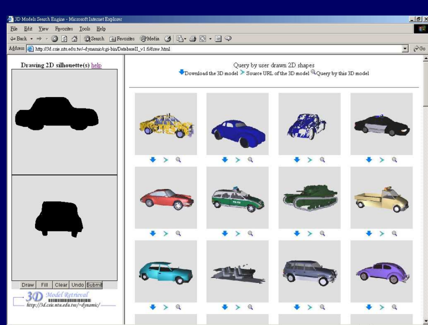
Retrieval from a large database

- Iterative early rejection of non-relevant 3D models
- Only parts of the images and coefficients are compared
- 6 iterations in total
  - 1st~5th iteration: Zernike moment
  - 6th iteration: Fourier descriptor
  - The threshold of removing models is set as the mean of the similarity

29

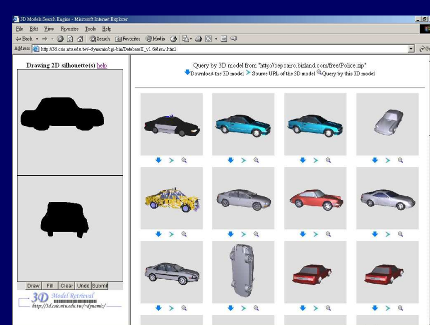
Speed of 3D shape search engine

(>10,000 3D models)



Retrieval results from user drawn 2D shapes

0.1 seconds



Retrieval results from interactively searching by selecting a 3D model

2 seconds

30

Experimental results

31

Ground truth data set

- 3DCAFE, downloaded in Dec. 2001
- All 1833 3D models
- Classified class
  - 47 class (category of airplane, car, chair, etc.)
  - 549 3D models
- Miscellaneous class
  - 1284 3D models

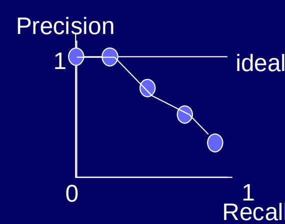
64 models | 52 models | 61 models | 15 models | 13 models

32

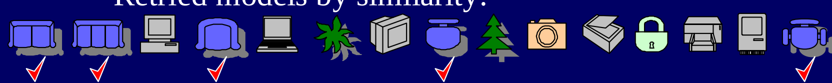
“Precision” vs. “Recall” diagram

$$\text{Precision} = \frac{\text{relevant correctly retrieved}}{\text{all retrieved}}$$

$$\text{Recall} = \frac{\text{relevant correctly retrieved}}{\text{all relevant}}$$



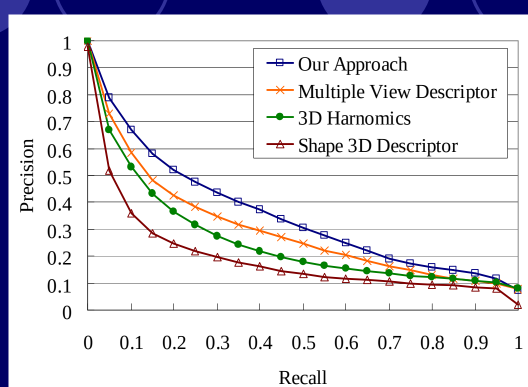
Input: Retried models by similarity:



Precision	1	3/4	4/8	5/15
Recall	1/4	2/4	3/4	4/4

33

Performance evaluation on NTU data



22% better
43% better
95% better

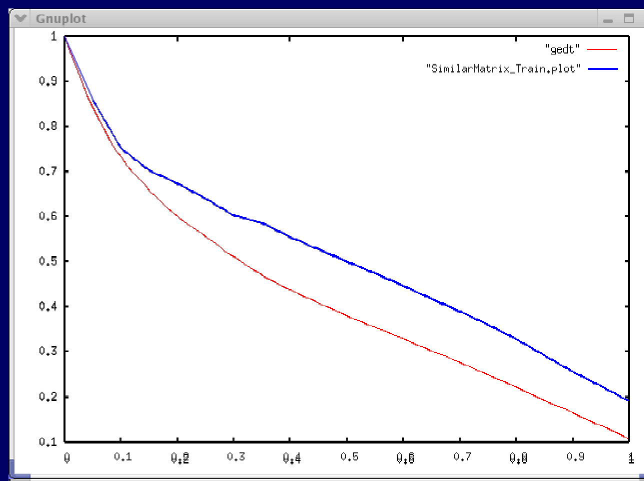
- Our Approach
- Multiple View Descriptor: MPEG-7 international standard
- 3D Harmonics: *ACM Trans. on Graphics*, Jan. 2003.
- Shape 3D Descriptor: MPEG-7 international standard

34

Precision Recall Plot for PSB

blue line:ours, red: the spherical harmonics

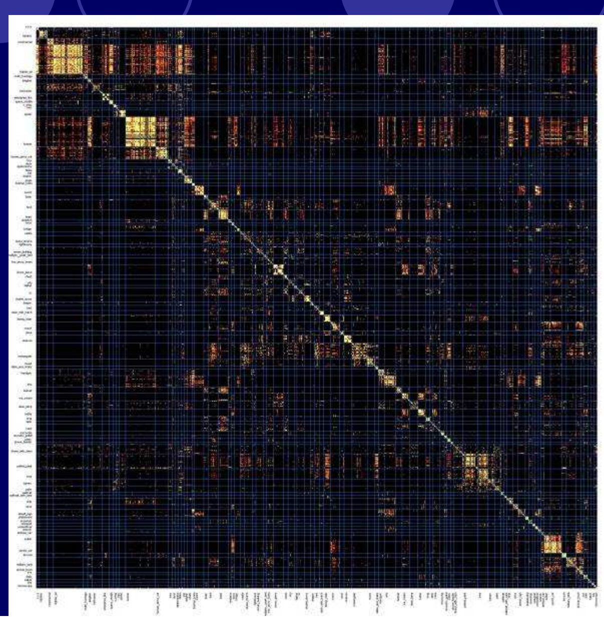
Data provided by Phil Shilane



35

Visualization of the similarity matrix, Data provided by Phil Shilane

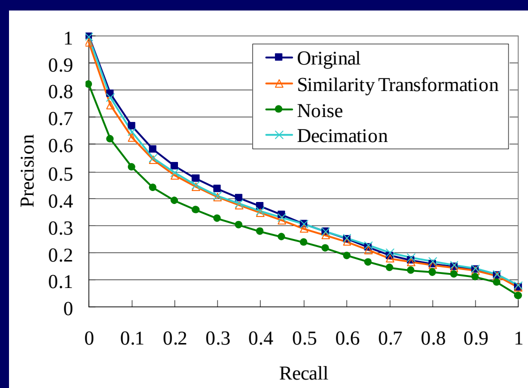
YELLOW, RED, BLACK: similarity values in decreasing order

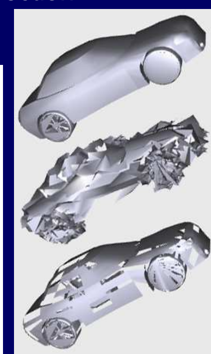


36

Robustness evaluation

- Similarity transformation: rotations(0º ~ 360º), translation(-10 ~ +10 times of ℓ) and scaling(a factor between -10 and +10)
- Noise: translating -3%~+3% times of ℓ in x, y, or z direction
- Decimation: deleting 20% of the polygons

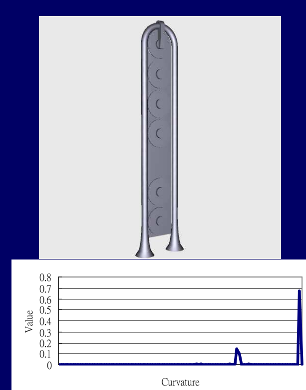




37

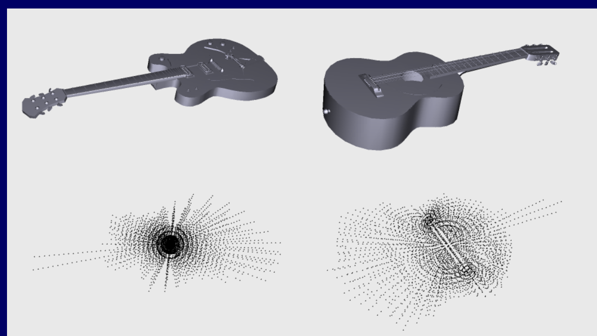
Shape 3D Descriptor

- Curvature histogram
- Discrimination: different models have the same curvature histogram
- Model degeneracy

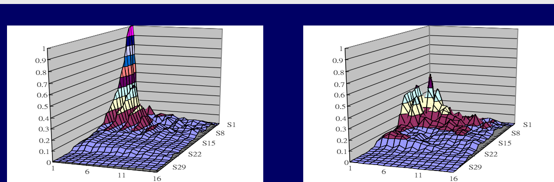


38

3D Harmonics



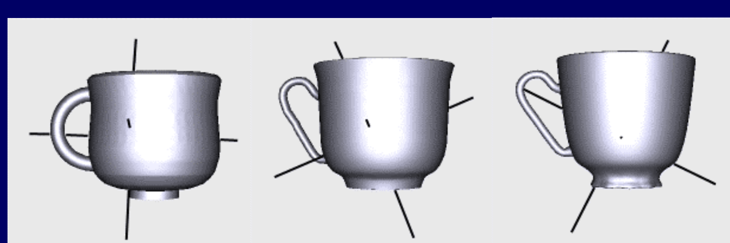
- Voxelization
- Shape near the center of mass has larger effects



39

Multiple View Descriptor (PCA based)

- Based on Principal Component Analysis (PCA)
- Principal axes are not good enough at aligning orientations of different models within the same class



40

Timing statistics in computing the similarity matrix

- The average time is 0.0013 second to compare two shape descriptors
- The average storage is 4700 bytes for a single shape descriptor
- The average time is 3.25 seconds to compute each shape descriptor
- on a PC with Pentium 4 CPU 2.4GHz, 256MB RAM and NVIDIA GeForce 2 MX

Future works

- More image metrics
- Cocktail approach
- Training / learning mechanism
- Partial matching

Conclusion

- A 3D model retrieval system is proposed based on visual similarity
- The new metric for 3D models is proposed
- Robustness, efficiency, discrimination, etc.
- 3D shape search engine for practical use
  - http://3d.csie.ntu.edu.tw

Thank You