Content Based 3D Shape Retrieval
A Survey of State of the Art

Ilknur Icke

`http://www.cs.gc.cuny.edu/~iicke`

November 2004

The Graduate Center,

City University of New York

Pattern Recognition Laboratory / Room 4-321

PhD Program in Computer Science

365 Fifth Avenue

New York, NY 10016

Motivation

Large databases of 3D Data are becoming available in various fields such as:

- Computer Aided Design(CAD)
- Computer Graphics and Vision
- Moleculer Biology, Chemistry, Medicine
- Archeology and other fields

Problem:How to find what we want in these large databases efficiently?
File name or textual annotation based classical searching methods do not work.

Solution: Use the content, namely the Shape.

What is Shape?

- Merriam-Webster Dictionary
  (a)The visible makeup characteristic of a particular item or kind of item.
  (b)spatial form or contour.

- Definition by D.G. Kendall
  Shape is all the geometrical information that remains when location, scale and rotational effects are filtered out from an object.

Outline

- Representing 3D Shapes in Digital World

  A closer look at how the 3D Shapes are organized in computer environment.

- Shape Similarity and Matching Concepts

  How to measure similarity of two shapes?

- 3D Shape Matching for Retrieval

  The review of the literature.

- Retrieval Performance and Related Issues

  How to measure the retrieval performance? How to compare different 3D shape matching techniques?

- Discussion

  What else could be done?

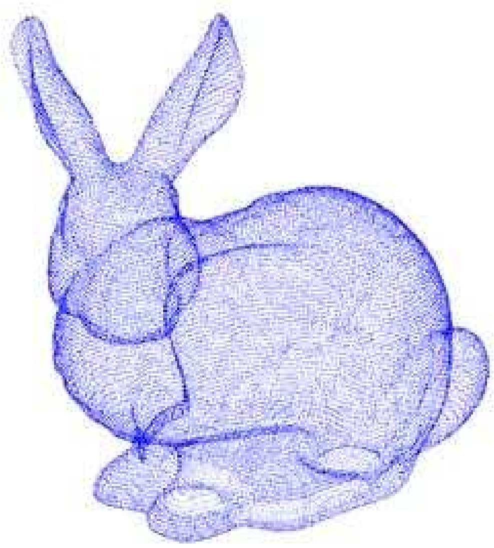
3D Shape Representations

- Static Shapes-Rigid Shapes that do not change in time by articulation or deformation.
  - Model Based Representations
    - Point Based models
    - Surface models
    - Volumetric(Solid) models
  - View Based Representations
    A set of 2D Projections(views) of the 3D Shape.
- Dynamic Shapes-Change in time by articulation or deformation.
  - Snakes: Active Contour Models (Kass et al. (1987))
  - Balloon Models (Chen and Medioni (1995))
  - Deformable Volumetric Models (Park et al. (1996))

Point Based Representations

Point Clouds

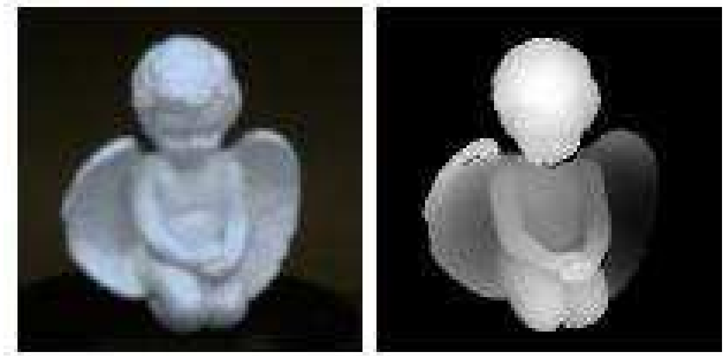
A set of points
$\mathbf{P} = \{\mathbf{p_1}, \mathbf{p_2}, ..., \mathbf{p_N}\}$
where $\mathbf{P} \epsilon \mathbf{R^3}$ and $\mathbf{p_i} = (x_i, y_i, z_i)^T$.



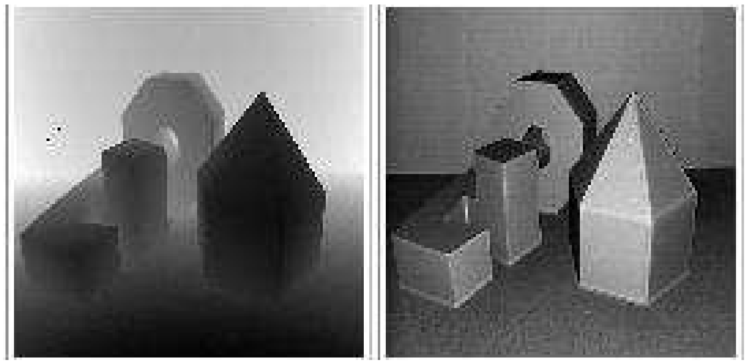
A snapshot of Stanford Bunny Point Cloud

Range Images

Each pixel represents the depth information.



Angel image from Ohio State University

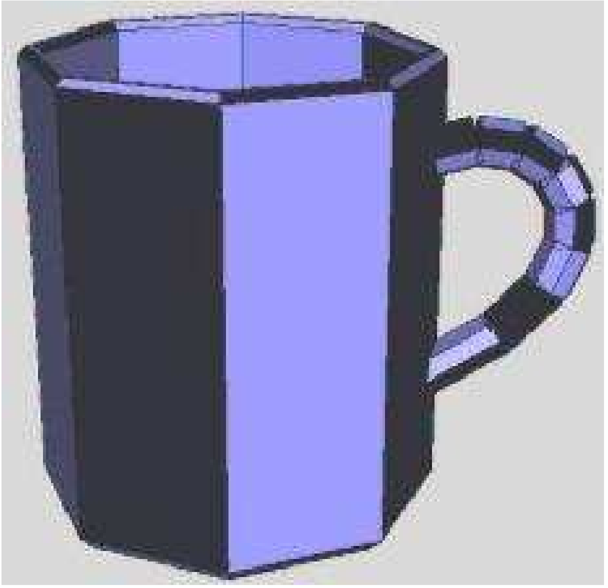


Polyhedral objects image from USF Database

Surface Representations-Polygon Models

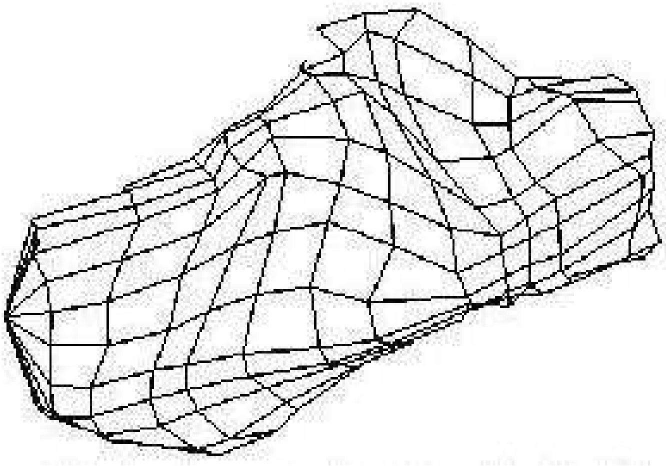
Polygon Soups

- The polygons might not be connected
- Ill-defi ned models
- Very common representation



A coffee cup CAD model

Polygon Meshes



A foot bone model (Campbell and Flynn(2001))

$\mathbf{M} =< \mathbf{P}, \mathbf{V} >$ ordered lists.

$\mathbf{V} = \{\mathbf{v_1}, \mathbf{v_2}, ..., \mathbf{v_N}\}$ list of vertices

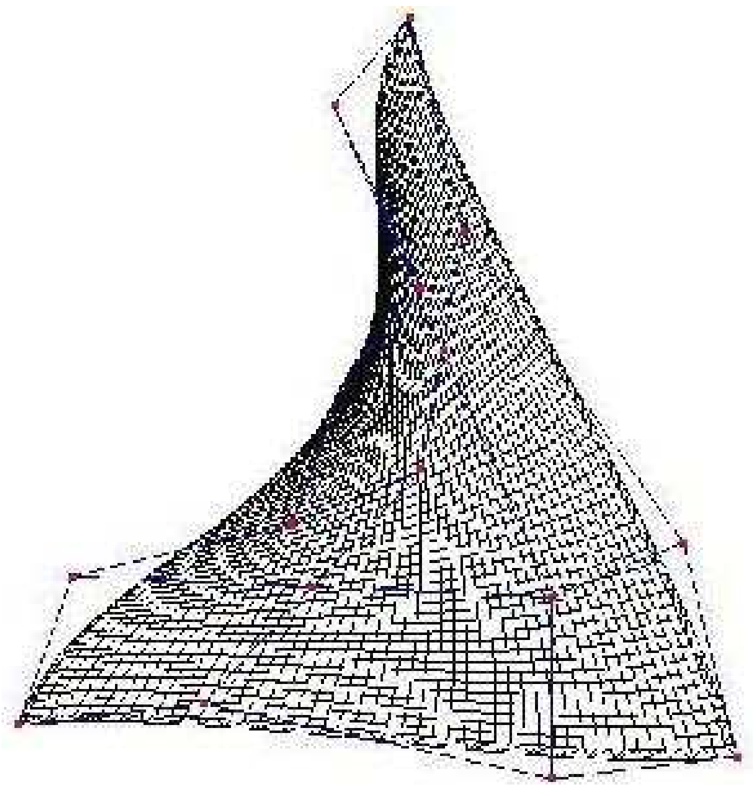
$\mathbf{v_n} = (x_n, y_n, z_n)^T$

$\mathbf{P} = \{\mathbf{p_1}, \mathbf{p_2}, ..., \mathbf{p_R}\}$ list of planar polygons

$\mathbf{p_r} = (\mathbf{v_{n,1}}, \mathbf{v_{n,2}}, ..., \mathbf{v_{n,k_r}})$

k_r the number of vertices in polygon $\mathbf{p_r}$

Surface Representations-Parametric Forms



A NURBS surface shown as mesh

(from Campbell and Flynn (2001))

A generic parametric form for a 3D surface is given as follows:

$$S(u,v) = \begin{bmatrix} x = f(u,v) \\ y = g(u,v) \\ z = h(u,v) \end{bmatrix}$$

where u and v are the parametric variables.
Domain of (u,v) can be taken as $[0,1] \times [0,1]$

A surface is generated from the Cartesian product of two curves.
A nonuniform rational B-spline (NURBS) is one type of parametric form which is defi ned as :

$$S(u,v) = \sum_i \sum_j B_{i,j} N(u), M(v)$$

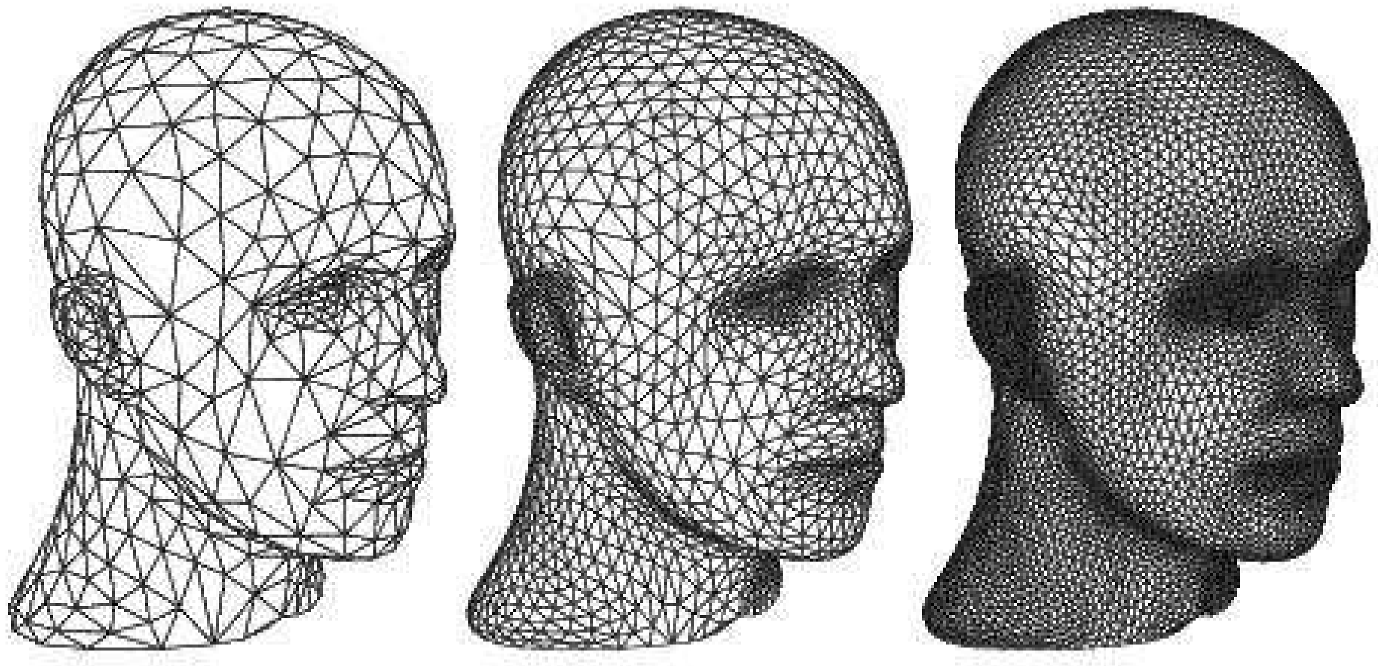
where N and M are B-spline basis functions

and $B_{i,j}$ are the coordinates of the control points.

These models can be easily converted to polygon meshes.

Surface Representations-Subdivision Surfaces

A detailed(and smoother) surface is generated starting from a coarser representation.

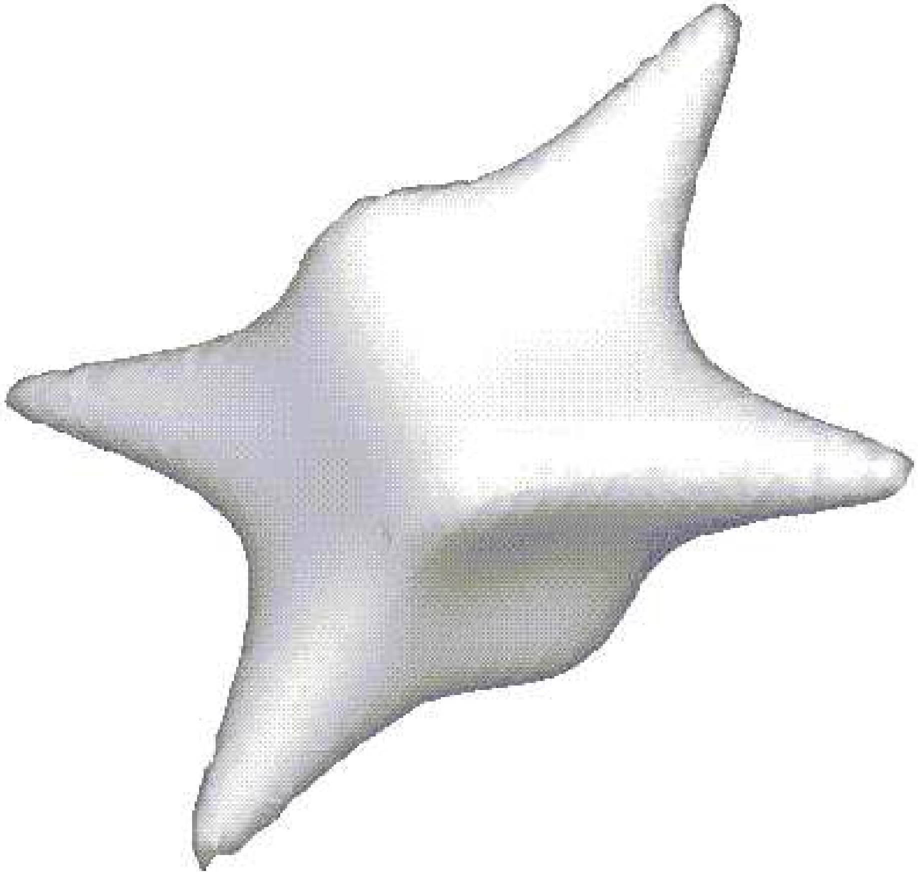


A head model from Zorin and Shroeder (1999)

Each triangle in the left mesh is divided into four triangles according to a subdivision rule, the result is the model in the middle. The rightmost model is generated from the middle one repeating the same process.

Surface Representations-Implicit Surfaces

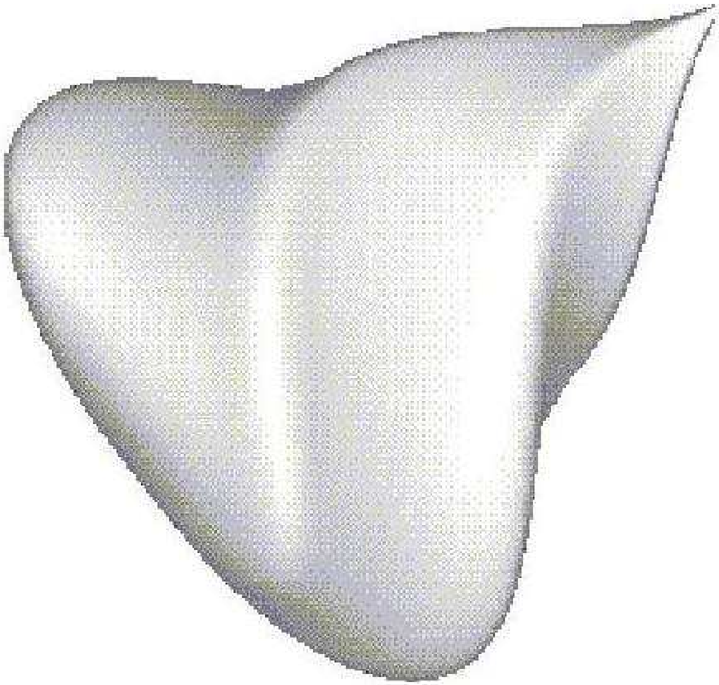
A 3D surface can be defi ned as the zero-set of an arbitrary function: $S = \{(x, y, z) | f(x, y, z) = 0\}$



A model generated by equation $f(x, y, z) = 2x^4 - 3x^2y^2 + 3y^4 - 3y^2z^2 - 2x^3z + 6z^4 - 1 = 0$

(from Campbell and Flynn (2001))

Surface Representations-Superquadrics



A deformed Superquadric

(from Campbell and Flynn (2001))

A superquadric is a closed surface spanned by a vector in which x, y, z are specifi ed as functions of the angles η and ω using a spherical product of two 2D parameterized curves.

A Superellipsoid is one type of superquadric which has the parametric form:

$$S(\eta,\omega) = \begin{bmatrix} x(\eta,\omega) \\ y(\eta,\omega) \\ z(\eta,\omega) \end{bmatrix} = \begin{bmatrix} a_1 cos^{\epsilon 1}(\eta) cos^{\epsilon 2}(\omega) \\ a_2 cos^{\epsilon 1}(\eta) sin^{\epsilon 2}(\omega) \\ a_3 sin^{\epsilon 1}(\eta) \end{bmatrix},$$

$$-\frac{\pi}{2} \leq \eta \leq \frac{\pi}{2}, -\pi \leq \omega \leq \pi$$

and implicit form:

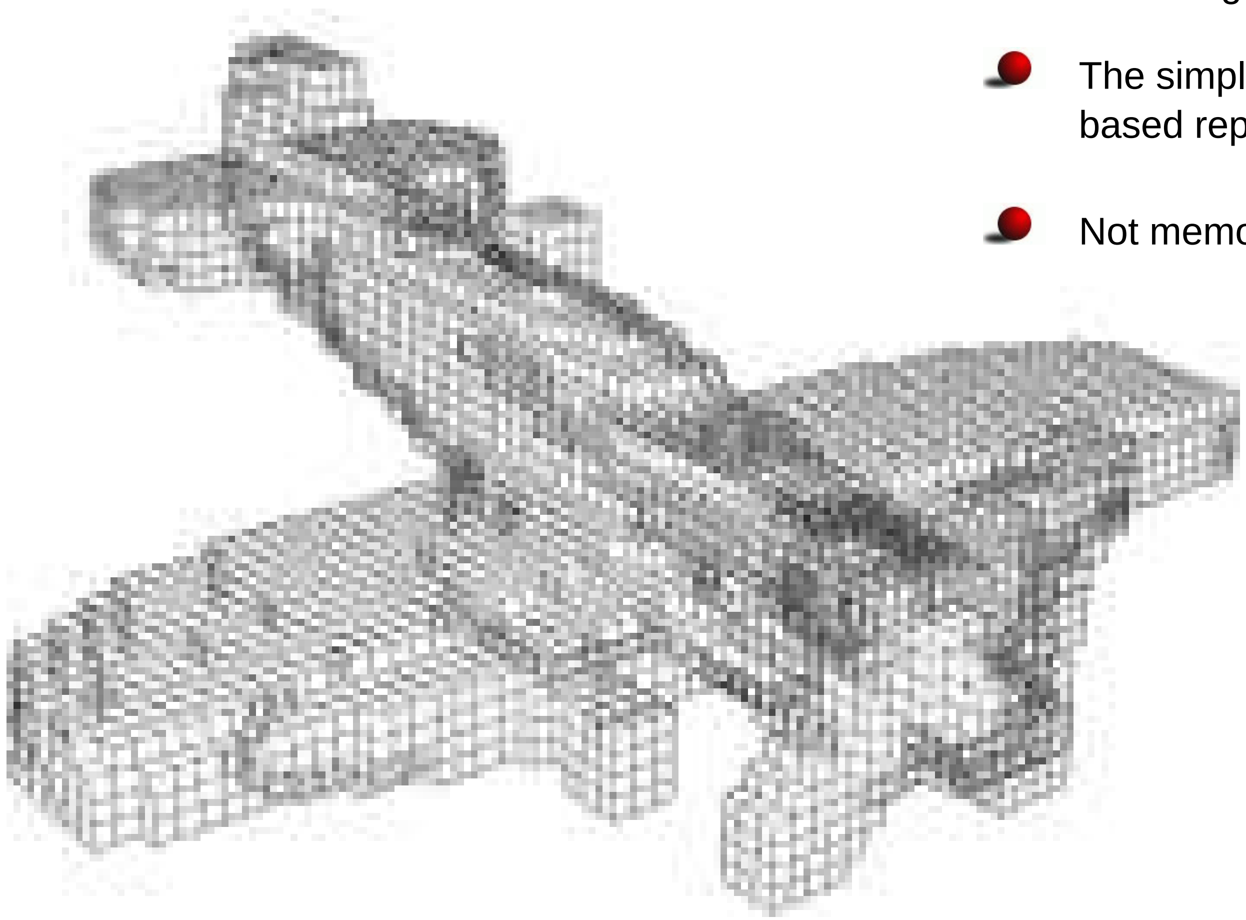
$$S(x,y,z) = \left[\left[(\frac{x}{a_1})^{\frac{2}{\epsilon 1}} + (\frac{y}{a_2})^{\frac{2}{\epsilon 2}} \right]^{\frac{\epsilon 2}{\epsilon 1}} + (\frac{z}{a_3})^{\frac{2}{\epsilon 1}} \right]^{\epsilon 1} = 0$$

where $(a1, a2, a3)^T$ is a scaling vector and ϵ_1,ϵ_2 represent the degree of squareness.

Superquadrics can also be deformed via tapering, twisting and bending transformations.

Volumetric(Solid) Representations-Space Subdivision

Voxels



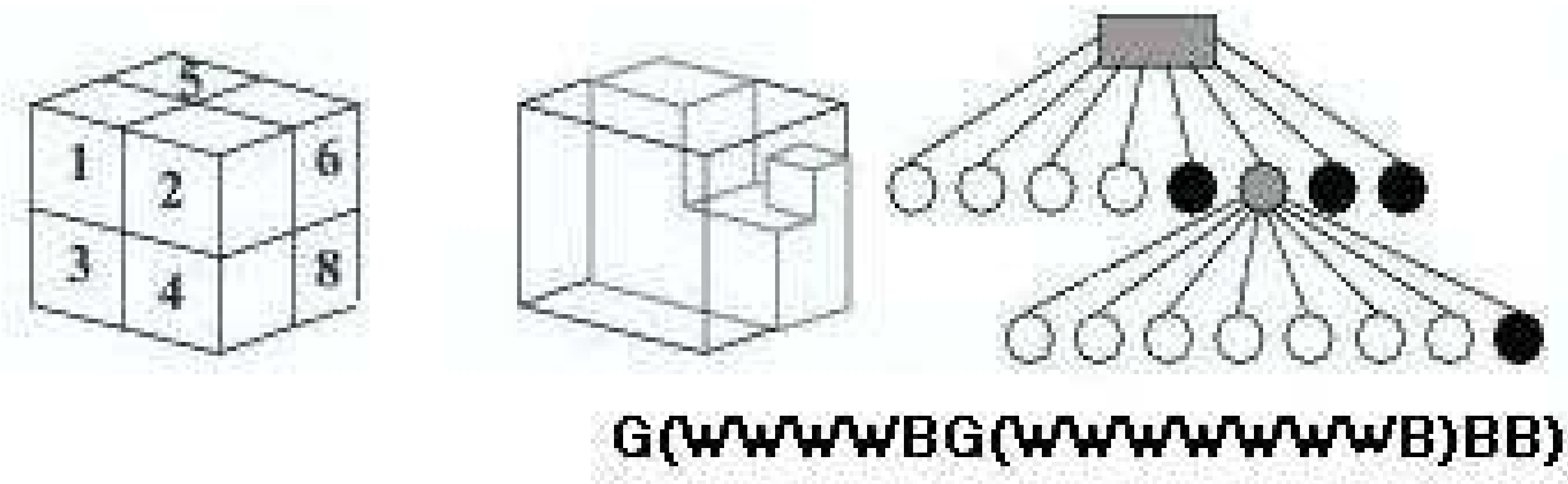
An Airplane model (from Kazhdan et al.(2003))

- A voxel is the minimum unit in volume rendering equivalent to a pixel in 2D.
- The simplest form of space subdivision based representation.
- Not memory effi cient.

Volumetric(Solid) Representations-Space Subdivision

Octree

A space subdivision based representation in which a cubic space is recursively divided into smaller cubic volumes and a hierarchical data structure is built.

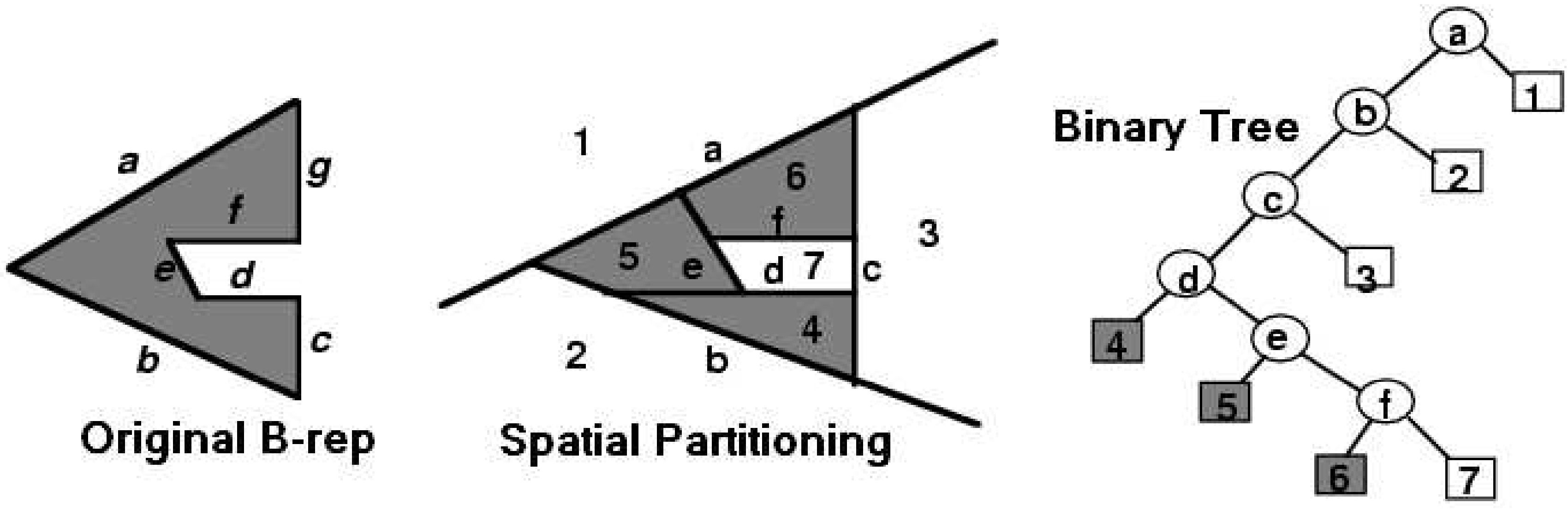


- white node(W) : completely empty subvolume
- gray node (G) : partially occupied subvolume
- black node(B) : completely occupied subvolume

Volumetric(Solid) Representations-Space Subdivision

Binary Space Partitioning(BSP) Tree

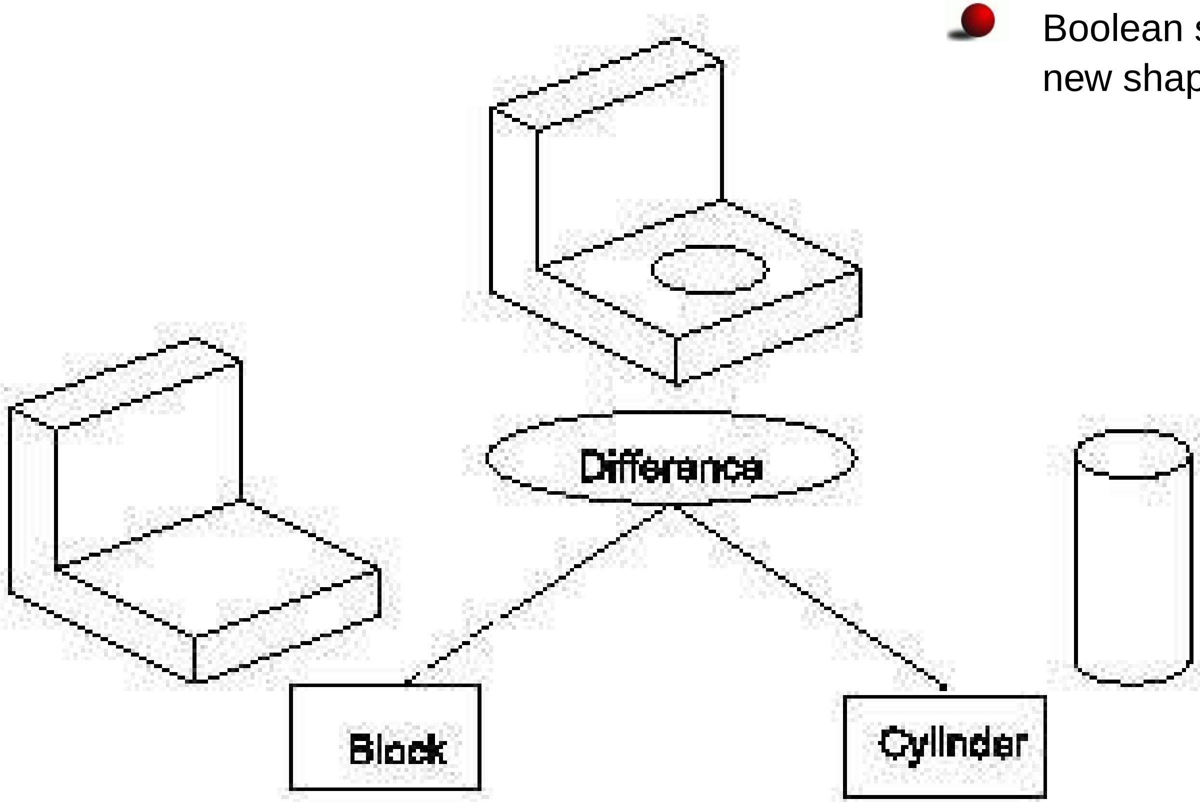
- An alternative to Octree representation.
- Provides a search structure as well as a representation of the geometry of one or multiple objects.



A BSP tree example from Naylor (1996)

Volumetric(Solid) Representations-CSG

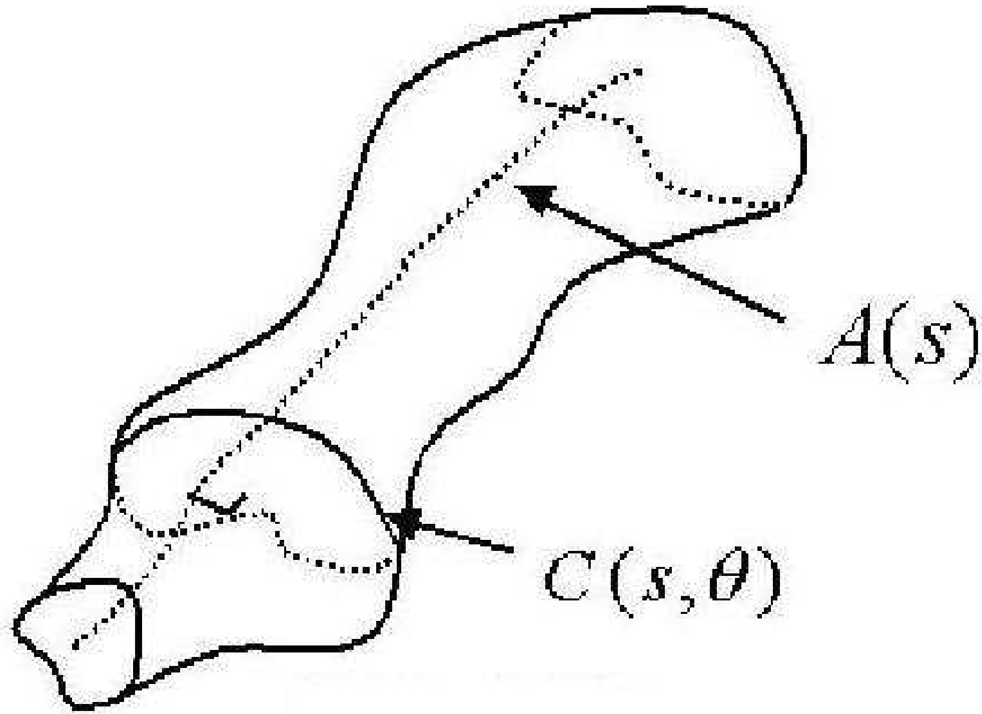
Constructive Solid Geometry(CSG)



- A hierarchical representation
- Each shape is made up of primitive shapes
- Boolean set operators are used to create new shapes from the primitive shapes

Volumetric(Solid) Representations-Generalized Cylinders

- A cross section contour $C(s,\theta)$ is swept along a space curve A(s) which is the spine(axis) of the model.
- The cross section contour may or may not be constant while sweeping.
- A cylinder is has a straight line segment as the spine and a circle with constant radius as the cross section.
- A torus has a a circle as the spine and another circle with constant radius as the cross section.
- A cone has a a straight line as the spine and a circle with an increasing radius $0 \ldots r_{max}$ as the cross section.



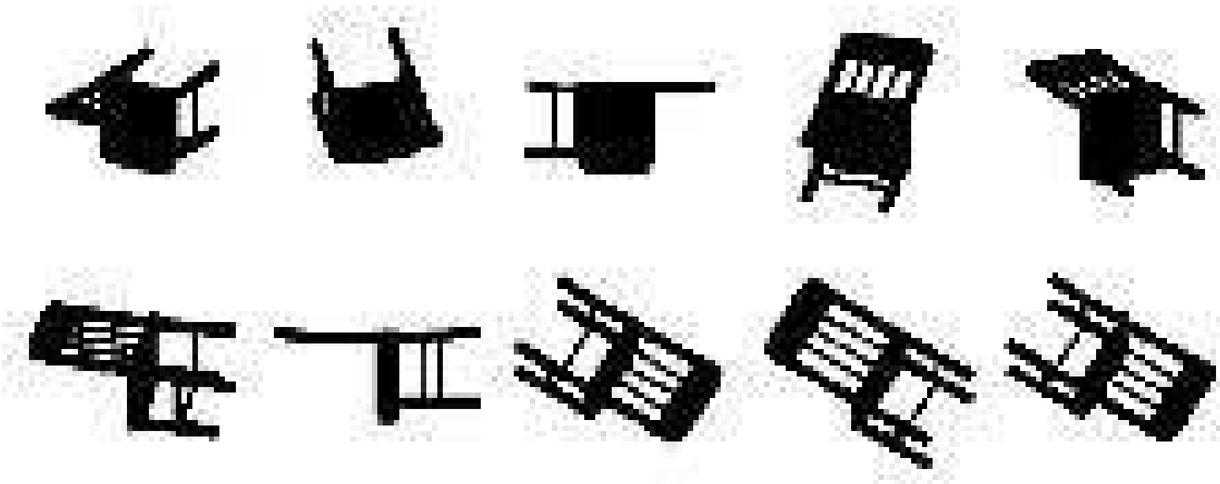
An example model from Campbell and Flynn (2001)

Summary of Model Based 3D Shape Representations

- Point Based Representations
  - Point Clouds
  - Range Images
- Surface Representations
  - Polygon Soups
  - Polygon Meshes
  - Parametric Forms
  - Subdivision Surfaces
  - Implicit Surfaces
  - Superquadrics
- Volumetric(Solid) Representations
  - Voxels
  - Octree
  - Binary Space Partitioning-BSP Tree
  - Constructive Solid Geometry(CSG)
  - Generalized Cylinders

View Based Representations-Silhouettes

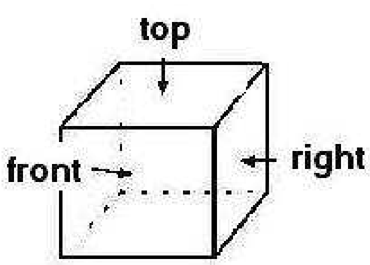
- They contain the boundary of a shape from one view point.
- A set of silhouettes are generated and stored to represent a 3D Shape.
- Useful when matching is done between one view of a 3D shape and a database of 3D shapes in which each shape is represented as a set of silhouettes.
- In theory different shapes might have the same set of silhouette images.

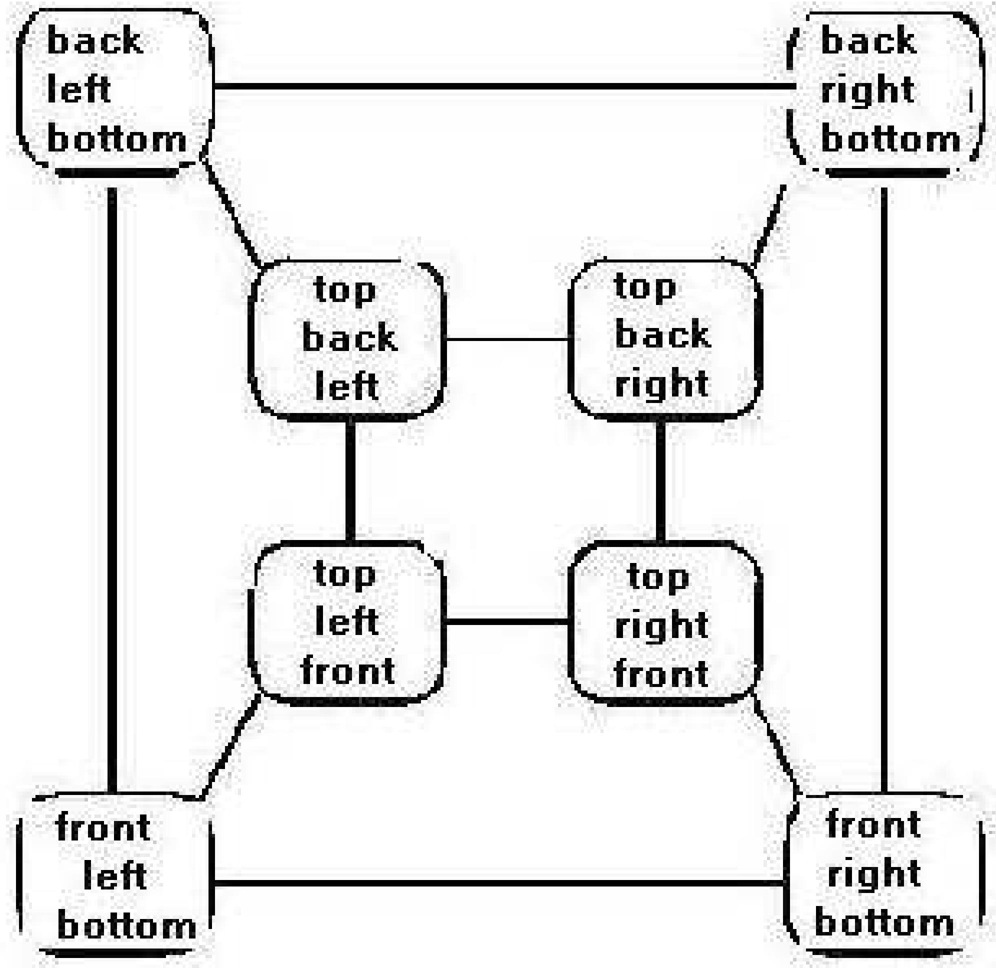


10 silhouettes of a chair (from Chen et al. (2003))

View Based Representations-View Classes

- 3D shapes look different from different points of views. The space of views can be partitioned into view classes (characteristic views).
- A view class representation named Aspect Graph was proposed by Koenderink and Van Doorn in 1979. Each aspect represents a class of views and neighboring nodes differ in one certain aspect.





Aspect Graph

Shape Similarity and Matching Concepts

- Shape Matching is a form of spatial pattern matching.
- It is the process of determining how similar/dissimilar two shapes are.
- Determining the similarity of shapes is generally done by computing a distance where small distance means more similarity, less dissimilarity.

Definition: Given a set of shapes $S = \{s_1, s_2, ..., s_N\}$, the *similarity distance* is defi ned as $d(s_i, s_j) : S \times S \rightarrow R^+ \cup 0$ where $s_i, s_j \epsilon S$. Function d may have some of the following properties:

- (i) Identity: $\forall s_i \epsilon S$, $d(s_i, s_i) = 0$
- (ii) Positivity: $\forall s_i, s_j \epsilon S, s_i \neq s_j$, $d(s_i, s_j) > 0$
- (iii) Symmetry: $\forall s_i, s_j \epsilon S$, $d(s_i, s_j) = d(s_j, s_i)$
- (iv) Triangle Inequality: $\forall s_i, s_j, s_k \epsilon S$,$d(s_i, s_k) \leq d(s_i, s_j) + d(s_j, s_k)$
- (v) Transformation Invariance: Given a transformation group G, $\forall s_i, s_j \epsilon S, g \epsilon G$, $d(s_i, g(s_j)) = d(s_i, s_j)$

If (i)-(iv) hold then d is called a *metric*.

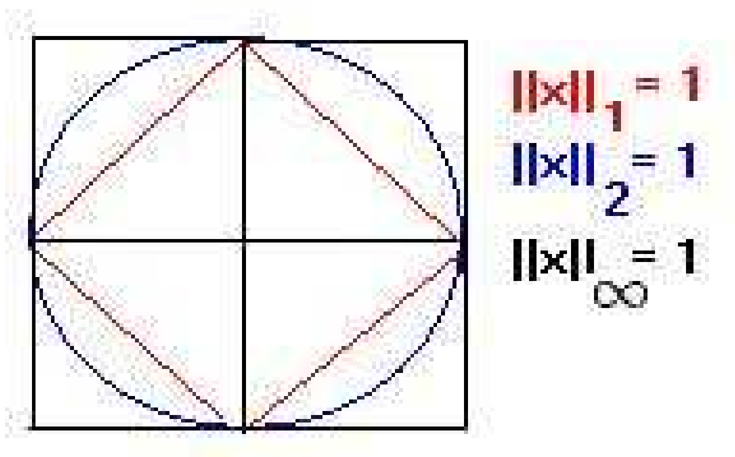
Example Distance Functions-Matching Points

Definition (L_p Norm-Minkowski Distance): Given two points $x, y \epsilon R^k$, the L_p distance is defined as:

$$L_p = (\sum_{i=1}^{k} |x_i - y_i|^p)^{1/p}$$

For $p \geq 1$, L_p distance is a metric.
If $p = 1$, it is called L_1 norm or Manhattan distance or city block distance.
If $p = 2$, it is called L_2 norm or Euclidean distance.

L_p norm is *not a transformation invariant* dissimilarity measure.



The plot of points in 2D satisfying $||x||_p = 1$

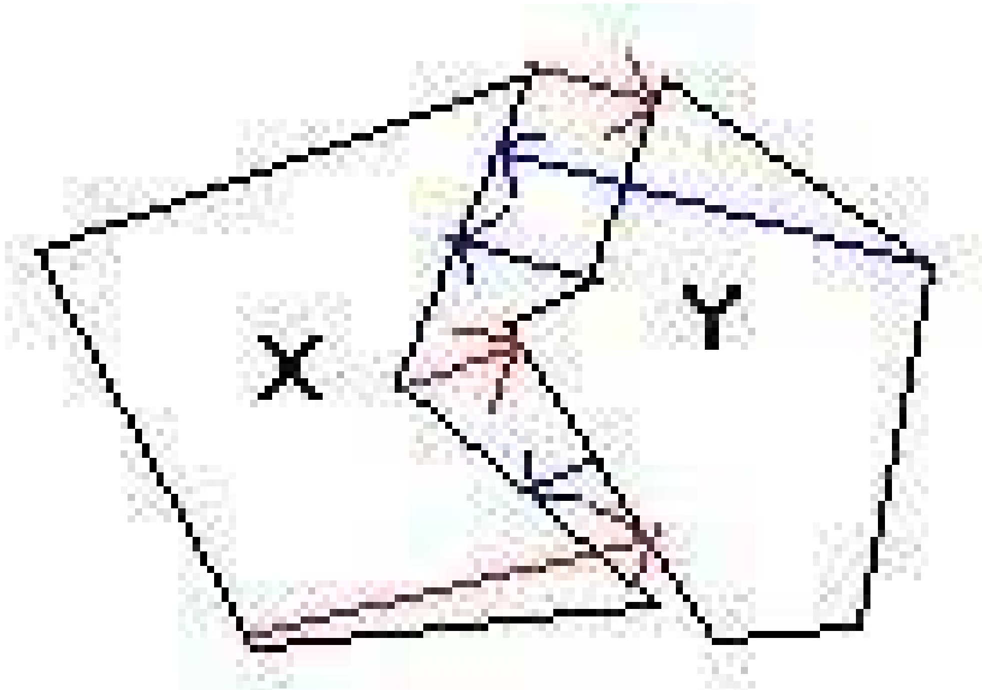
Example Distance Functions-Matching Point Sets

Definition (Hausdorff Distance): Given two shapes represented by two sets of points: $X = \{x_1, x_2, ..., x_M\}$ and $Y = \{y_1, y_2, ..., y_N\}$ the Hausdorff distance between X and Y is defined as:

$$H(X,Y) = max(h(X,Y), h(Y,X))$$

where $h(X,Y) = \max_{x \epsilon X} \min_{y \epsilon Y} ||\mathbf{x} - \mathbf{y}||$ and $||.||$ is usually Euclidean distance.

The Hausdorff distance is a metric but *not transformation invariant*.



Matching 3D Shapes: Transformation Invariance

- By defi nition, the shape of a 3D object is independent of any **translation, scaling and rotations** applied to it.
- The distance function is desired to be *invariant* to these transformations.
- All possible transformations need to be considered:

$$D(s_i, s_j) = \min_{g \in G} d(s_i, g(s_j))$$

where G is the group of transformations.

- This is not an effi cient similarity measure for 3D shape matching.

Transformation Invariance

Definition(Pose normalization):
Given a set of shapes $S = \{s_1, s_2, ..., s_N\}$, a metric $d(s_i, s_j)$ and a group of transformations G.

Let n be a many-to-one function where $\forall g \in G, s_i \in S, n(g(s_i)) = \hat{s_i}$ and $\forall s_i, s_j \in S, d(s_i, s_j) \sim d(\hat{s_i}, \hat{s_j})$.

Then

$$d(s_i, s_j) \sim d(\hat{s_i}, \hat{s_j}) = d(g(s_i), g(s_j))$$

In 3D shape matching, the group of transformations G contain any combination of translation, scaling and rotation.

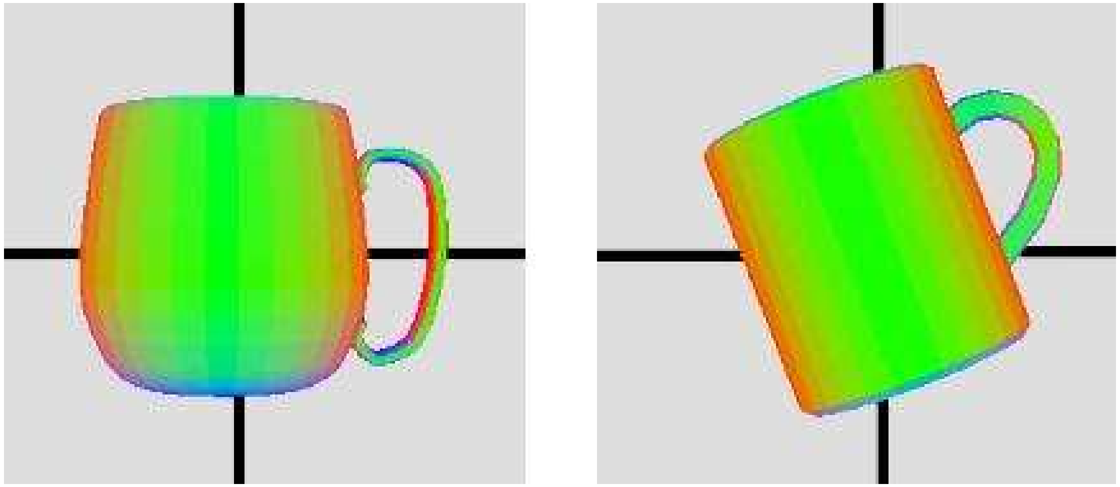
The function n defi ned on this G is called a **pose normalization function**.

Pose Normalization of 3D Shapes

- **Translation Invariance**
  Center of mass of the model is computed and translated to the origin.

$$x' = x - \bar{x}, y' = y - \bar{y}, z' = z - \bar{z}$$

- **Scaling Invariance**
  All the models are scaled to the same dimensions.
- **Rotational Invariance**
  Principal axes of the model is computed using the Principle Components(PCA) Algorithm. Then the model is rotated so that the principle axes would align with the canonical coordinate system for all models. This is not a robust method.



Transformation Invariance

Definition(Invariant features):
Given a set of shapes $S = \{s_1, s_2, ..., s_N\}$, a metric $d(s_i, s_j)$ and a group of transformations G.

Let f^+ be a function where $\forall g \in G, s_i \in S, f^+(g(s_i)) = f^+(s_i)$ and $d(s_i, s_j) \sim d(f^+(s_i), f^+(s_j))$.
Then

$$d(s_i, s_j) \sim d(f^+(s_i), f^+(s_j)) = d(g(f^+(s_i), g(f^+(s_j))$$

The function f^+ is called an **invariant feature extraction function**.

Shape Descriptors for 3D Shape Matching

- 3D Shapes in their representation forms are not well suited for matching. Therefore simplifi ed descriptions (*shape descriptors*) capturing the signifi cant features of the shapes are needed.

  Definition(Shape descriptor generation):
  Given a set of shapes $S = \{s_1, s_2, ..., s_N\}$, a metric $d(s_i, s_j)$.

  Let f be a function where $\forall s_i, s_j \in S, d(s_i, s_j) \sim d(f(s_i), f(s_j))$.

  The function f is called a **shape descriptor generation function**.

  If f is also invariant to translation, scaling and rotations then it is called an **invariant shape descriptor generation function**.

- Shape descriptors could be numerical or structural.
  Numerical shape descriptors generate a mapping $\mathbf{X} \rightarrow \mathbf{R^n}$ where $\mathbf{X}$ is the space of original shape representations.

3D Shape Based Retrieval Problem

Definition (3D Shape Based Retrieval Problem) : Given a database of 3D shapes $S = \{s_1, s_2, ..., s_N\}$ and a query shape q, retrieve the shapes that are similar to q.

Solution:

- (Decision Problem) Given a threshold similarity value t, retrieve all the shapes where $d(f(q), f(s_i)) < t$.
- (Computation Problem) Retrieve the top k shapes where $d(f(q), f(s_i))$ are minimum where d is a distance function, preferably a metric.

where f is a **shape descriptor generation function**.

- If f is not invariant to translation, scale and rotations then the shapes have to be pose normalized fi rst.

Overview of 3D Shape Based Retrieval Techniques

- Model Based Techniques
  - Geometric Techniques-Numerical shape descriptors
    - Global geometry
    - Local geometry
  - Structural and Topological Techniques
    Numerical or Structural shape descriptors
    - Based on the connected components (parts) or the number of holes and tunnels in 3D shapes.
    - Shape descriptors are generally structural i.e graph and shape matching is more complex.
    - Noise in the data can lead to incorrect structures, therefore these techniques are not robust.
- View Based Techniques-Numerical shape descriptors
  - The set of 2D views of the 3D model should be proven to cover all possible viewpoints if the models are not going to be made rotationally invariant.
  - Generally a large number of 2D views are required.

Overview of Geometric Techniques

- Global Shape Descriptors
  - They describe the shape as a whole.
  - Generally easy to implement.
  - Not robust against noise.
  - Partial matching is not possible.
  - Initial pose normalization is generally necessary.
  - In most techniques, a fi xed number of points are sampled from every model.
- Local Shape Descriptors
  - Instead of the whole shape, they try to capture information about a neighbourhood of some points on the boundary of the shape.
  - More robust against noise.
  - Not very effi cient to compute and match.
  - Partial matching is possible.
  - Translation and Rotation invariant, scale is normalized.
  - These methods are generally used in recognizing objects in the scenes in presence of possible occlusion.

Global Shape Descriptors-Features

- One can describe a shape in terms of some invariant measurements (features) like the volume, surface area, aspect ratio and statistical moments.
  - **Elad et al. (2000)**. Moments based method that works on polygon mesh models. N points are sampled from the surface of the pose normalized model. A number of moments are computed using an approximated moment formula:

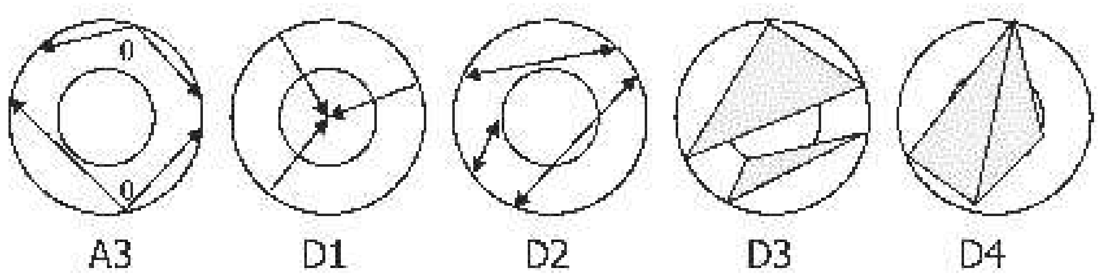
    $$\hat{M}[k_1, k_2, k_3] = \frac{1}{N}\sum_{j=1}^{N} x_j^{k_1} y_j^{k_2} z_j^{k_3}$$

    These moments form a feature vector to be used as a shape descriptor. The similarity measure is the L_2 distance between the feature vectors.
  - **Zhang and Chen (2001)** describe effi cient ways to compute the volume, surface area and moments of a 3D model represented as polygon mesh.
- These features are rough descriptions of the shapes therefore they are not very discriminative. They can be used as an initial fi ltering in stage retrieval.

Global Shape Descriptors-Feature Distributions

- **Osada et al. (2002)**. Shape Distributions on polygon models.
  A number of points are sampled from the model surface. A *shape distribution* is estimated by creating equi-width binned histogram of some measurements on the sampled points.
  Various *translation and rotation invariant* geometric properties on the model surface are measured:
  - A3: The angle between three randomly chosen points.
  - D1: The distance between a fi xed point and one randomly chosen point.
  - D2: The distance between two randomly chosen points.
  - D3: The square root of the area of the triangle formed by three randomly chosen points.
  - D4: The cube root of the volume of the tetrahedron formed by four randomly chosen points.

  Shape matching reduces to histogram matching. Scale invariance is achieved by normalizing the histograms.

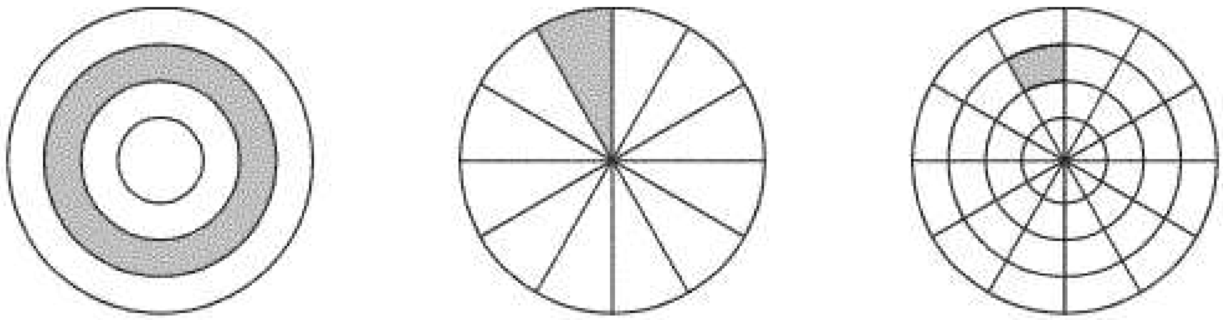


- **Obhuchi et al. (2003)** and **Ip et al. (2002)** describe extensions to D2 method.

Global Shape Descriptors-Spatial Maps

- **Ankerst et al. (1999)** 3D Shape Histograms for protein models.

  The idea is to divide the 3D space into sections and approximate the distributions of model points or other features in each section. The neighbourhood relations between these sections are also considered in similarity matching process.



Shell, sector and spiderweb models

Shell model is rotationally invariant while the others are not. Pose normalization is needed.
A quadratic distance function between two models given their feature vectors q and m is :

$$d(\mathbf{q},\mathbf{m})^2_A = (\mathbf{q}-\mathbf{m})\cdot A\cdot(\mathbf{q}-\mathbf{m})^T = \sum\sum a_{ij}(q_i - m_i)(q_j - m_j)$$

Matrix A contains the degree of similarity between each component which is a measure of how close the sectors are in the space decomposition. If A is identity matrix d is equal to Euclidean distance and the spatial relationships of the sectors are ignored.

Global Shape Descriptors-Transforms and Special Functions

- **Zaharia and Preteux (2001)**. 3D Hough Transform Based Shape Descriptors.

  Works on polygon mesh models. Invariance to translation and scaling is required.
  For each model, 3D Hough Transforms on all possible coordinate systems form the feature vector for that model. $\mathbf{x} = (x_1, x_2, ...x_N)^T (N = 48)$.
  The similarity measure is :

  $$d(\mathbf{q}, \mathbf{m}) = \min_{1 \le i \le N} ||x_{q,i} - x_{m,i}||$$

  where q and m are feature vectors of two models and $||.||$ is L_2 or L_1 distance.

- **Vranic and Saupe (2001)**. Discrete 3D Fourier Transform (3DDFT) based shape descriptors.
  Input models are pose normalized mesh models. Discrete 3D Fourier transform is applied to the voxelized models. The real valued coeffi cients of the transform are taken to form a feature vector as the shape descriptor for each model. L_1 and L_2 distances are used as similarity measures.

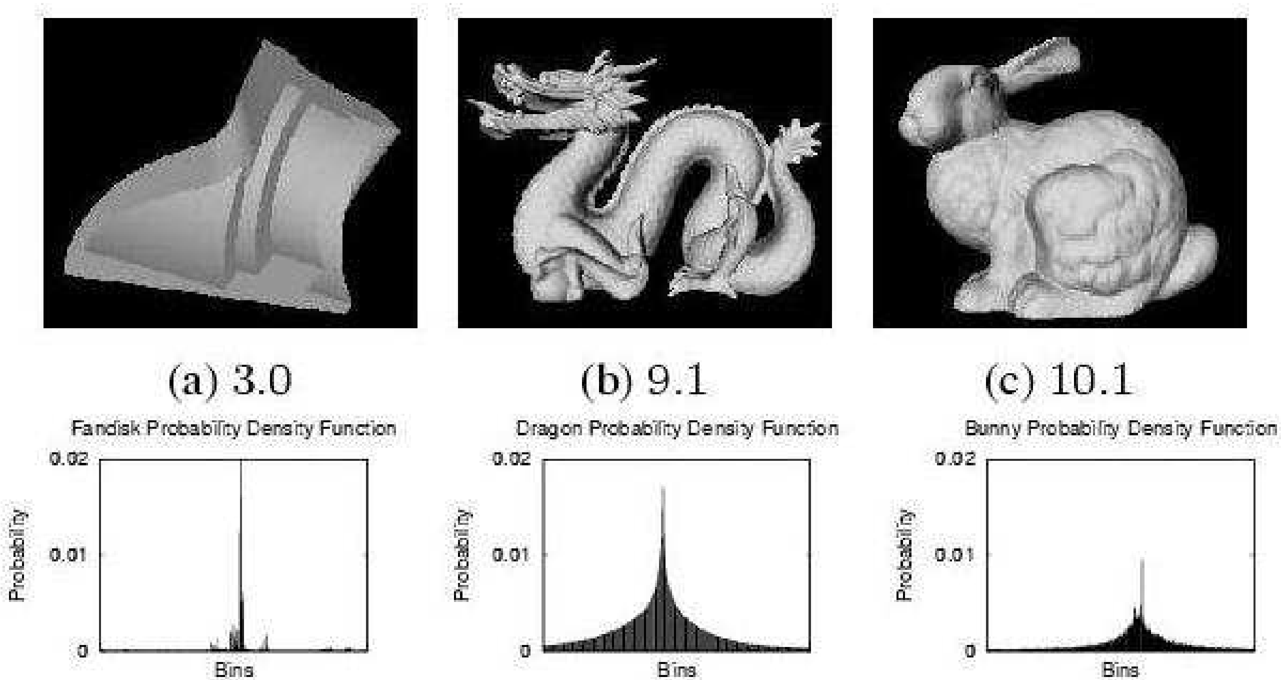
- **Paquet et al. (2000)**. Wavelet transform based shape descriptors for 3D model retrieval.

- **Daras et al. (2004)**. 3D Radon transform based shape descriptors.

- **Kazhdan et al. (2003)**. Spherical Harmonics.
  A rotationally invariant shape descriptor based on taking samples in concentric spheres intersected with voxelized models. Then each sphere is described in terms of spherical functions which can be transformed into frequency domain and represented using less amount of coeffi cients to form a shape descriptor for each model. Similarity measure is L_2 distance between the shape descriptors.

Global Shape Descriptors-Information Theory Approach

- **Page et al. (2003)**. A method to measure shape complexity in terms of surface curvature information.
  - A set of points are sampled on the polygon mesh
  - An estimate of Gaussian Curvature on each of the sampled points are computed
  - An equi-width binned(M bins) histogram is computed on the curvature estimates as an approximation to the probability density function.
  - Based on the this histogram the *shape information* of a model is computed in terms of entropy: $H = -\sum_{i=1}^{M} p_i \; log_2 p_i$



Global Shape Descriptors-Other Techniques

- **Volumetric Difference**

  Shapes occupy the volume in different ways though the two shapes might have the same total volume. A simple volume difference can not really discriminate between to shapes.

  - **Kaku et al. (2004)**. Oriented Bounding Box(OBB) Tree based volumetric difference
  - **Leifman et al. (2003)**. Octree based volumetric difference.

- **Projection(Morphing) into a Canonical Shape**

  The idea is that the amount of energy required to morph a shape into another can be a measure of similarity between the two shapes. In 3D shape retrieval each model is morphed into a canonical model and the amount of energy is taken as the shape descriptor. Pose normalization is required.

  - **Leifman et al. (2003)**. Projection into a sphere. Computes the enery to morph a model into its bounding sphere in terms of the distances from the sampled points on the model surface to the sphere. minimum Euclidean distance from the sphere to the model is used.
  - **Yu et al. (2003)**. Projection into a sphere. Generates a map of distances from the object to its bounding sphere. The similarity measure is a weighted Euclidean distance on the distance maps.

- **Weighted Point Sets**

  - **Tangelder and Veltkamp (2003)**. A number of methods for generating weighted point sets from the pose normalized polygon models placed in a 3D grid. Comparing the sets of weighted point sets is done using a variation of earth mover's distance.

Geometric Techniques-Local Shape Descriptors

- **Kortgen et al. (2003)**. 3D Shape Contexts.

  N points are sampled on the given mesh model.
  The space is decomposed into shells or sectors and for each sampled point a histogram of the relative coordinates of the remaining $N-1$ points is computed.

  The histogram is called a shape context for that point.
  Shape matching is done by comparing the shape contexts to fi nd the corresponding points on two models.

- **Johnson and Hebert (1999)**. Spin images.

  A spin image is a 2D histogram that is computed at a chosen point on the surface of a model.
  For a mesh model, a spin image can be computed for every vertex on the mesh.
  A surface normal can be estimated at each vertex which is picked as the oriented point.

  A set of points within the maximum distance D to the oriented point which satisfy the condition that the angle between their normal and the normal of the oriented point is within allowed angle values are selected as the contributing points.

  Then a 2D histogram is calculated based on the perpendicular distances to the surface normal and to the tangent planes at the to the oriented point. This histogram can also be used as an image.

- **De Alarcon et al.(2002)**. Spin images method in 3D Shape retrieval.

  For each 3D model given as a polygon mesh, a large number of spin images are generated. A k-means clustering algorithm is used to cluster these spin images as an indexing technique.

Structural and Topological Shape Descriptors

- Numerical Shape Descriptors

  Yu et al. (2003). Surface Penetration Map.
  An imaginary ray is shot from the center of a sectioned bounding sphere of a model.
  Depending on the number of holes in the model it may pass through multiple surfaces.
  The average number of these surfaces per section forms a shape decriptor for the model.

- Graph Structures
  - **Hilaga et al. (2001)**. Topology Matching.
    Multiresolutional Reeb Graphs(MRG) for comparing 3D models.
  - **Tung and Schmit (2004)**. Augmented Reeb Graph based algorithm.
    The graph also includes geometrical properties such as volume and curvature.
  - **Sundar et al. (2003)**. Skeletal Graph.
    The graphs encode topological and geometrical information(i.e radius) about the models.

- Relational Structures.

  Each 3D shape is thought of as a combination of primitives and a set of relationships among these primitives. Each primitive is described in terms of some geometric attributes like radius, area and so on.
  - **Vosselman (1992).** Relational matching framework.
  - **Haralick and Shapiro(1993).** Consistent-labeling framework based on a relational distance.

View Based Shape Descriptors

The idea is that if two shapes are similar then they should look similar from all viewing angles

- **Chen et al. (2003)**.
  Works on polygon mesh models that are made invariant to translation and scaling. $N = 10$ silhouettes are generated from $V = 10$ uniformly distributed viewpoints on the bounding sphere of the model. Rotating the sphere 10 times, a total of 100 silhouettes are generated. A 2D shape descriptor is computed for each silhouette. The similarity measure between two models q and m is :

  $$D(q, m) = \min_{1 \leq i \leq V} \sum_{k=1}^{N} d(I_{q_{ik}}, I_{m_{ik}})$$

  where $I_{q_{ik}}$,$I_{m_{ik}}$ are the 2D descriptors computed on the silhouettes and d is the L_1 distance.

- **Obhuchi et al. (2003).**
  Works on polygon mesh models that are made invariant to translation and scaling. $N = 42$ depth-buffer rendered (range image) images of the model are generated. The set of images are claimed to discretely cover all possible view aspects of the model. Then for each image a 2D shape descriptor is computed. The similarity measure between two models q and m is:

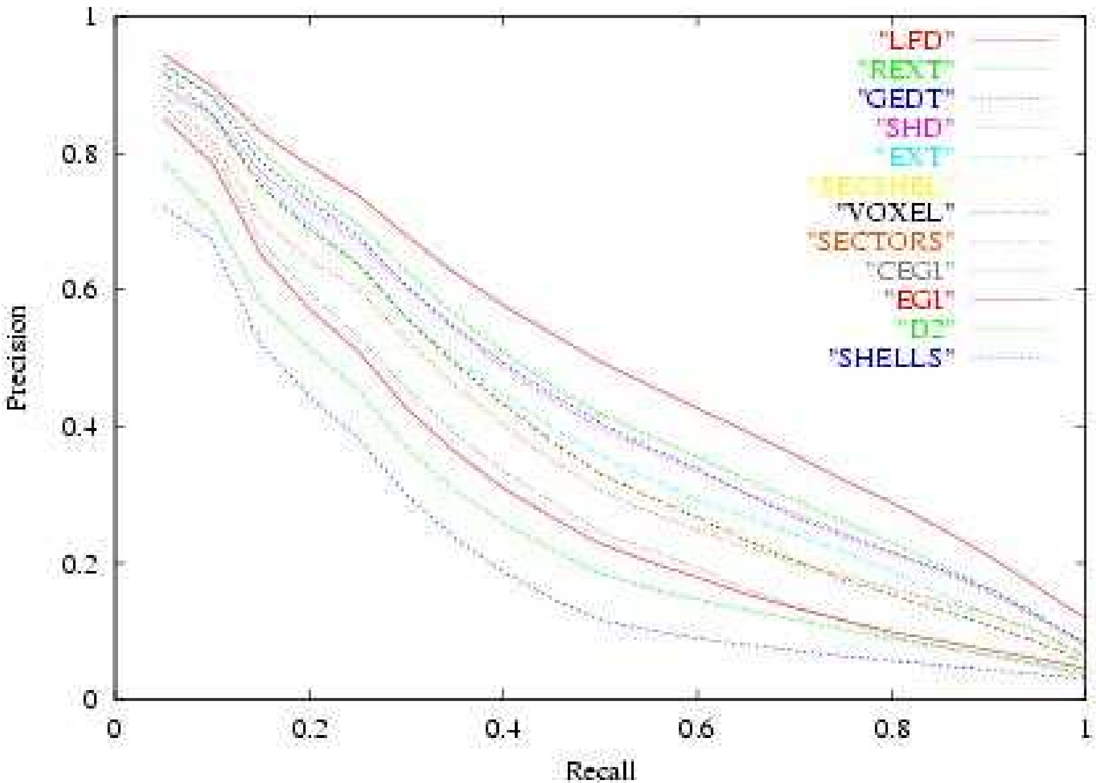
  $$D(q, m) = \frac{1}{N} \sum_{i=1}^{N} \min_{1 \leq j \leq N} d(I_{q_i} - I_{m_j})$$

  where I_{q_i}, I_{m_j} are the computed 2D descriptors and the d is the L_1 distance.

Retrieval Performance and Related Issues

- The retrieved models are displayed in descending order on their similarity distance to the query model.
- If each model has an assigned class label then precision-recall plots can be used to visualize the retrieval performance and compare different techniques. Let the q be the query model, C be the class that q belongs to.

$$Precision(P) = \frac{\text{\# models retrieved that are in class C}}{\text{\# number of models retrieved}},\ Recall(R) = \frac{\text{\# models retrieved that are in class C}}{\text{\# number of models in class C}}$$



Shape Descriptor Selection

- Some researchers in 3D shape retrieval area defi ne multiple shape descriptors to match 3D shapes. Generally, they use these descriptors individually and compare them based on the retrieval performance.
- A combination of the shape descriptors might perform better than using individual descriptors.
- **Vandeborre et al. (2002)**.
  Three types of numerical shape descriptors are computed for each model. Each descriptor returns N models which have rankings based on their L_1 distance to the query model. Each model's new ranking can be approximated in two ways:
  - OR method : Considers the highest ranking using any of the 3 descriptors.
  - MEAN method : Takes the average value of the rankings using all 3 descriptors.

  Better retrieval results are reported using the combination methods.
- Different descriptors might perform better depending on the query shape.
- **Bustos et al. (2004)**.
  A query dependent method for estimating the retrieval performance of a shape descriptor. The idea is that the result set should be as coherent as possible. This is measured in terms of entropy impurity.
  If the retrieved models belong to one class, the performance is considered as good. If the retrieved models fall into varioud classes the performance is considered as worse.
  The model database needs to be preclassifi ed.

Discussion-Similarity Measure Selection

- L_1 and L_2 distances assume specific underlying distribution of the data.
- The distribution assumption may not be true in real 3D Shape databases.