



McGill School of
Computer Science

Indexing and Matching Articulated 3-D Models Using Medial Surfaces

Juan Zhang

School of Computer Science,
McGill University, Montréal

August 2005

A Thesis submitted to McGill University in partial fulfilment
of the requirements for the degree of Master of Science

© Juan Zhang, 2005



Library and
Archives Canada

Bibliothèque et
Archives Canada

Published Heritage
Branch

Direction du
Patrimoine de l'édition

395 Wellington Street
Ottawa ON K1A 0N4
Canada

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

ISBN: 978-0-494-22780-0

Our file Notre référence

ISBN: 978-0-494-22780-0

NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protègent cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.


Canada

Abstract

We consider the use of medial surfaces to represent symmetries of 3-D objects. This allows for a qualitative abstraction based on a directed acyclic graph (DAG) of components and also a degree of invariance to a variety of transformations including the articulation and deformation of parts. We demonstrate the use of this representation for both indexing and matching 3-D object models. Our formulation uses the geometric information associated with each node along with an eigenvalue labeling of the adjacency matrix of the subgraph rooted at that node. We compare our algorithm with the techniques of harmonic spheres (Kazhdan et al., 2003c) and shape distributions (Osada et al., 2002). The results demonstrate the significant potential of medial surface-based representations and their graph spectra in the context of 3-D model retrieval in computer graphics.

Résumé

Nous considérons l'utilisation de surfaces médiales afin de représenter les symétries d'objets en trois dimensions. Cette approche permet une abstraction qualitative basée sur un graphe orienté sans circuit (DAG) des composantes. Elle permet aussi une certaine invariance vis-à-vis plusieurs transformations géométriques, telles que l'articulation et la déformation des parties. Nous démontrons l'utilisation de cette représentation autant pour l'indexation que le regroupement de modèles d'objets tridimensionnels. Notre formulation emploie l'information géométrique associée à chaque noeud ainsi que les valeurs propres de la matrice de contiguïté du sous-graphe basé à ce noeud. Nous comparons notre algorithme aux techniques des sphères harmoniques (Kazhdan et al., 2003c) et de distributions des formes (Osada et al., 2002). Nos résultats démontrent le potentiel important des représentations fondées sur les surfaces médiales et de leur spectres de graphe quant à la recherche de modèles en trois dimensions dans le domaine de l'infographie.

Acknowledgments

First and foremost, I would like to thank my supervisor, Prof. Kaleem Siddiqi, for his tireless guidance and constructive suggestions on my research. His ideas have kept me motivated, confident, challenged and productive at all times and his support has gone far beyond the expectation of a Master's thesis supervisor. I am certain that his instruction will benefit me in the course of my career.

I am indebted to Ali Shokoufandeh, Diego Macrini and Sven Dickinson, for their collaboration in this research. I am especially thankful to Diego Macrini who not only provided me with his implementation of the indexing and matching algorithms but also offered prompt answers to my questions. I am thankful to Ali Shokoufandeh and Sven Dickinson for providing valuable suggestions on detailed problems for their efforts in composing the conference article that grew out of our collaboration. I also want to mention John Novatnack who helped me test the EMD algorithm for node similarity between two 3D parts.

Special thanks go to my colleagues, Sylvain Bouix, Peter Savadjiev, Scott McCloskey and Maxime Descoteaux, who shared their algorithms, including the computation of medial surfaces of 3D objects, the voxelization of 3D mesh models and the implementation of mean curvature. I also would like to thank Svetlana Stolpner and Ryan Eckbo, who offered their support when I faced difficulties. I am grateful to Svetlana Stolpner for helping me review this thesis. I am thankful to Marc Gendron-Bellemare for his assistance in preparing the french version of the abstract.

Finally, I would like to thank my dear Dad and Mum for their financial and spiritual support and their understanding in times of significant stress. A huge thank you to Ming Li, who gave me confidence and encouragement from far away.

I would like to thank all the kind-hearted staff at the Center for Intelligent Machines and the School of Computer Science.

Contributions of Authors

Parts of this thesis will appear in the EMMCVPR 2005 conference. These parts have involved collaborations between myself, Kaleem Siddiqi and Diego Marcini, Ali Shokoufandeh and Sven Dickinson which proceeded in four phases.

First, the decomposition of medial surfaces, the interpretation of directed acyclic graph and the node similarity were introduced by Kaleem Siddiqi and myself. Parts of chapter 3 and chapter 4 of the thesis are based on these developments.

Second, the computation of medial surfaces based on the average outward flux and the labeling of medial surfaces was studied by Sylvain Bouix and Kaleem Siddiqi. These lead to their publication in the sixth European Conference on Computer Vision. Chapter 3 of the thesis is based on these developments.

Third, a system for indexing and matching directed acyclic graph was designed by Ali Shokoufandeh, Sven Dickinson, Kaleem Siddiqi and Diego Marcini. The implementation and the programming was done by Diego Marcini. Chapter 4 of the thesis is based on this system.

Finally, I worked on the integration of the system and on using medial surfaces for indexing and matching 3-D objects. This is described in chapter 4, with experimental results presented in chapter 5.

Contents

1	Introduction	1
1.1	Background	3
1.2	Method Overview	4
1.3	Contributions	5
1.4	Organization	6
2	Previous Work	7
2.1	Translation and Rotation Normalization Needing Methods	8
2.2	Translation Invariant Methods	13
2.3	Rotation Invariant Methods	13
2.4	Translation and Rotation Invariant Methods	15
2.5	Articulation Invariant Methods	17
3	Medial Surfaces and Directed Acyclic Graphs	21
3.1	Medial Surfaces	22
3.1.1	Computation of Medial Surfaces	22
3.1.2	Classification and Segmentation of Medial Surfaces	24
3.2	Directed Acyclic Graph	27
3.2.1	Definition and Properties of a Directed Acyclic Graph	28
3.2.2	Forming Directed Acyclic Graphs from Segmented Medial Surfaces	28

4	Indexing and Matching	31
4.1	Indexing	32
4.1.1	Criteria for an Effective Index	32
4.1.2	Formulating an Index	35
4.2	Matching	36
4.2.1	Node Similarity	36
4.2.2	Bi-partite Graph Matching Algorithm	37
5	Experiments	41
5.1	Construction of the Database	42
5.2	Indexing Experiments	43
5.3	Matching Experiments	44
6	Conclusions	51
6.1	Discussion	52
6.2	Contributions and Summary	52
6.3	Future Work	53
	Appendices	56
A	Useful links for online 3D shape retrieval	57
	Bibliography	63

List of Figures

1.1	Articulating Parts	2
1.2	Framework	5
3.1	Locus points of a hand	23
3.2	Medial Surfaces and DAGs	25
3.3	Medial Surfaces and DAGs	26
3.4	Segmented Medial Surfaces	27
3.5	Medial Surfaces and DAGs	30
4.1	Forming a Low-Dimensional Vector Description of Graph Structure. At node a , we compute the sum of the magnitudes of the k_1 largest eigenvalues of the adjacency sub-matrix defined by the subgraph rooted at a . The sorted sums S_i become the components of $\chi(V)$, the <i>topological signature vector</i> (or TSV) assigned to V	35
4.2	Indexing Mechanism. Each non-trivial node (whose TSV encodes a topological abstraction of the subgraph rooted at the node) votes for models sharing a structurally similar subgraph. Models receiving strong support are candidates for a more comprehensive matching process.	36

4.3	Matching Algorithm. Given two graphs to be matched (a), form a bipartite graph (b) spanning their nodes but excluding their edges. Edge weights ($W(i, j)$) not only encode node content similarity (see Section 4.2.1), but the structural similarity of their underlying subgraphs, as encoded by the difference in their respective TSV's. The best matching pair is identified, the two nodes are removed from their respective graphs and added to the solution set of correspondences, and the process applied recursively to their subgraphs (c). . .	39
5.1	Database Exemplars. 20 members are shown from each of the object classes (with the exception of the class dolphins which has fewer exemplars). Exemplars from classes on the left have significant part articulation of a complexity not seen in the Princeton Shape Benchmark. Note that we treat the dinosaurs and the four-legged animals as members of a single object class "four-limbs".	46
5.2	Indexing Results: Percentage Recall	47
5.3	Indexing Results: Average Ranks	48
5.4	49
5.5	50

List of Algorithms

3.1	Average Outward Flux.	24
3.2	Topology Preserving Thinning.	25

List of Tables

3.1	The topological classification of (Malandain et al., 1993).	27
5.1	Average Matching Results Using MS. Each object in the database is matched against all the other objects in the database. Each cell shows the average similarity between objects selected from two fixed object classes. In each row red and blue boxes are drawn, respectively, around the two highest average similarity scores. In all cases the highest score coincides with the correct object class. In most cases there is also a very significant difference between the top two average similarity scores.	45

Chapter 1

Introduction

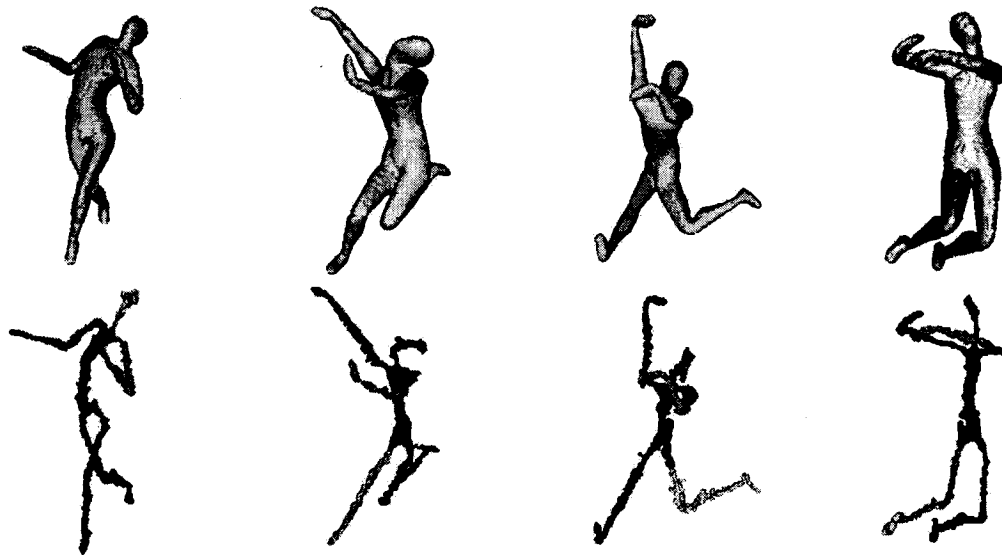


FIGURE 1.1: Exemplars of the object class “human” created by changes in pose and articulations of parts (top row). The medial surface (or 3-D skeleton) of each is computed using the algorithm of (Siddiqi et al., 2002) (bottom row). The medial surface is automatically partitioned into distinct parts, each shown in a different color.

With an explosive growth in the number of 3-D object models stored in web repositories and other databases, the graphics community has begun to address the important and challenging problem of 3-D object retrieval and matching, a problem which traditionally falls in the domain of computer vision research. Recent advances include query-based search engines (Funkhouser et al., 2003a) which employ promising measures including spherical harmonic descriptors and shape distributions (Osada et al., 2002). Such systems can yield impressive results on databases including hundreds of 3-D models, in a matter of a few seconds.

Thus far the emphasis in the computer graphics community has broadly been on the use of qualitative measures of shape that are typically global. Such measures are robust in the sense that they can deal with noisy and imperfect models, and at the same time are simple enough so that efficient algorithmic implementations can be sought. However, such methods are sensitive to deformations of objects and articulation of their parts.

As a motivating example, consider the 3-D models in Fig. 1.1. These four exem-

plars of an object class were created by articulations of parts and changes of pose. For such examples, the very notion of a center of mass or an origin, which is crucial for the computation of descriptions such as shape histograms (sectors or shells) (Ankerst et al., 1999a) or spherical extent functions (Vranic and Saupe, 2001), can be nonintuitive and arbitrary. In fact, the centroid of such models may actually lie in the background. To complicate matters, it is unclear how to obtain a global alignment of such models, and hence signatures based on a Euclidean distance transform (Borgefors, 1984; Funkhouser et al., 2003a) have limited power in this setting. As well, measures based on reflective symmetries (Kazhdan et al., 2003a), and signatures based on 3-D moments (Elad et al., 2001) or chord histograms (Osada et al., 2002) are not invariant under such transformations.

In this thesis, we build on a recent technique to compute medial surfaces (Siddiqi et al., 2002) by proposing an interpretation of its output as a directed acyclic graph (DAG) of parts. We then suggest refinements of algorithms based on graph spectra to tackle the problems of indexing and matching 3-D object models. These algorithms have already shown promise in the computer vision community for category-level view-based object indexing and matching using 2-D skeletal graphs (Siddiqi et al., 1999b; Shokoufandeh et al., 1999). We demonstrate their significant potential for 3-D object retrieval with experimental results on a database of 320 models representing 11 object classes, including exemplars of both rigid objects and ones with significant deformation and articulation of parts.

1.1 Background

The computer vision community has grappled with the problem of *generic* or category-level object recognition by suggesting representations based on volumetric parts, including generalized cylinders, superquadrics and *geons* (Binford, 1971; Marr and Nishihara, 1978; Pentland, 1986a; Biederman, 1987). Such approaches build a degree of robustness to deformations and movement of parts, but their representa-

tional power is limited by the vocabulary of geometric primitives that are selected. Motivated in part by such considerations there have been attempts to encode 3D shape information using probabilistic descriptors. These allow intrinsic geometric information to be captured by low dimensional signatures. An elegant example of this is the geodesic shape distribution of (Hamza and Krim, 2003) where information theoretic measures are used to compare probability distributions representing 3D object surfaces. In the domain of graph theory there have also been attempts to address the problem of 3D shape matching using representations based on Reeb graphs (Hilaga et al., 2001). These allow for topological properties to be captured, at least in a coarse sense.

An alternative approach is to use 3-D medial loci (3-D skeletons), obtained by considering the locus of centers of maximal inscribed spheres along with their radii (Blum, 1973). As pointed out by Blum, this offers the advantage that a graph of parts can be inferred from the underlying local mirror symmetries of the object. To motivate this idea, consider once again the human forms of Fig. 1.1. A medial surface-based representation (bottom row) provides a natural decomposition, which is largely invariant to the articulation and bending of parts.

1.2 Method Overview

We develop an integrated framework for indexing and matching 3-D objects using medial surfaces and their graph spectra. The key idea is to decompose a 3-D object model into parts and to interpret the parts into a directed acyclic graph so that indexing and matching techniques can be applied (Siddiqi et al., 1999a; Macrini, 2003; Shokoufandeh et al., 2005).

An integrated framework for understanding the indexing and matching 3-D system using medial surfaces is demonstrated in Fig. 1.2. We begin by voxelizing a 3D mesh model into volumetric data. Then, medial surfaces are abstracted from the volumetric data and decomposed into parts. The directed acyclic graph obtained

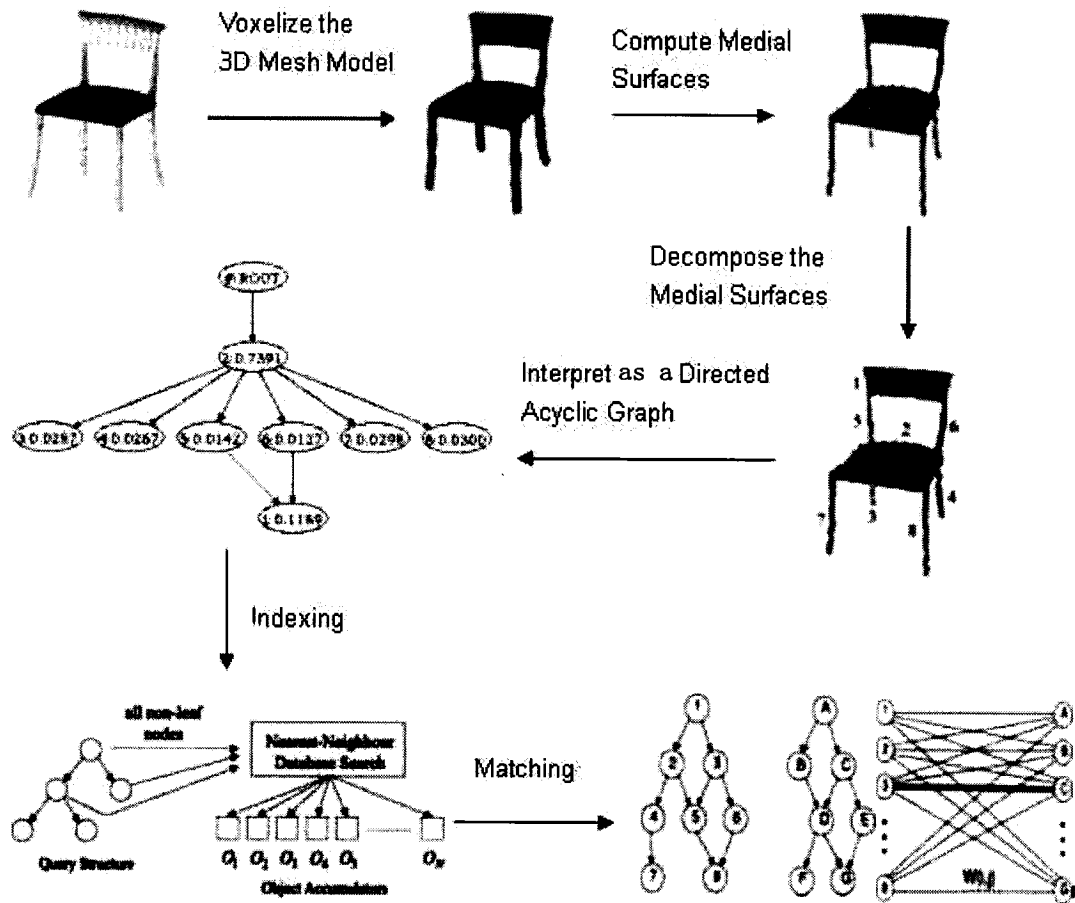


FIGURE 1.2: An integrated framework of indexing and matching 3D objects in a large 3D object database.

from the decomposed medial surfaces is added to a database, on which indexing and matching is applied.

1.3 Contributions

We advance the state-of-the-art in 3-D object model retrieval by: 1) introducing a modification of a Euclidean distance function-based method for computing and segmenting medial surfaces, 2) proposing a DAG representation of the medial surface which captures a notion of part saliency, 3) building on algorithms in the com-

puter vision literature to address the problem of 3-D model indexing and matching in a uniform framework and 4) presenting indexing and matching results on a database of object models organized according to an *entry* level of categorization, with categories having significant part articulation. Whereas all the pieces of this system have been developed in past work, putting them together and demonstrating them in the context of 3-D model retrieval with comparative results against competing methods is the main contribution of this thesis. Some of these results are also presented in the article "Retrieving Articulated 3-D Models Using Medial Surfaces and their Graph Spectra" (Zhang et al., 2005).

1.4 Organization

An outline of the thesis is follows. Chapter 2 contains a review of relevant work on shape based 3D object matching. Chapter 3 introduces medial surfaces and their computation, classification and segmentation. It also develops an interpretation of the decomposed medial surfaces as a directed acyclic graph for indexing and matching, which is discussed in Chapter 4. Chapter 5 presents experimental results evaluating medial surface based graph matching and comparing it against the alternative techniques of shape distributions (Osada et al., 2001) and spherical harmonics (Kazhdan et al., 2003c). Finally, a summary of contributions and a discussion of further work appears in Chapter 6.

Chapter 2

Previous Work

Numerous 3-D shape matching schemes can be found in the literature of computer vision and computer graphics (Alt and Guibas, 1996; Arman and Aggarwal, 1993; Besl and Jain, 1985; Loncaric, 1998; Pope, 1994; Veltkamp and Hagedoorn, 1999). A few links to 3-D shape retrieval systems are presented in Appendix A (Bochuan et al., 2004). At first glance, since shape based 3-D object matching avoids the challenges of projection, occlusion and segmentation, it might seem easier than 2-D shape matching. Unfortunately, since 3-D surfaces do not have a natural arc length parameterization and the topology of the 3-D objects can be complex, the pose registration and feature correspondence problems in three dimensions are difficult to address. Many shape matching approaches perform pose normalization as a preprocessing step before extracting a shape descriptor characterizing intrinsic properties of the object. Discriminative matching methods without pose estimation and registration have not been well explored. Matching articulated and deformed objects is another significant challenge for 3-D shape based object matching techniques. The following subsections give a detailed discussion of the methods used in 3-D shape matching. The approaches are characterized according to whether the matching methods need normalization or registration of objects and whether they can cope with articulated objects: 1)translation and rotation normalization needing methods, 2)translation invariant methods, 3)rotation invariant methods, 4) translation and rotation invariant methods, 5)articulation invariant methods. A discussion of each method's advantages and disadvantages is presented after a short introduction.

2.1 Translation and Rotation Normalization Needing Methods

3-D object models, given in the form of polygonal meshes or volumetric data, are created with various poses in arbitrary reference frames. A common process of

shape comparison between two 3-D objects begins by extracting feature vectors and then computing the distance between the feature vectors. However, feature vectors of the same 3-D model abstracted based on the distribution of the data in different reference frames may result in distinct orders of feature vector elements. Hence, to obtain a uniform order of feature vectors for similar models, one possibility is to estimate the pose of the 3-D objects and then optimally align the 3-D models into canonical coordinates.

Principal Component Analysis is the most prominent technique for pose normalization and has been tested in pose registration for 2-D images, for which a detailed description can be found in (Elad et al., 2002). It begins with translating the origin of the 3-D models to the center of mass and then rotating around the origin so that the orientation of the largest variance is along the x-axis. Another rotation is then performed around the x-axis in order that the orientation of the second largest variance is consistent with the y-axis. The PCA techniques enable the alignment of the models into canonical coordinates for which matching algorithms can later be applied. However, as experience reveals (Funkhouser et al., 2003b), PCA does not offer a perfect alignment for objects within the same class because of the unstable orientation of the second largest variance. A modification of PCA associating weights with triangles for 3-D mesh models is introduced in (Vranic et al., 2001). Also, Vranic et al. (2001) comes up with a continuous principal component analysis for an infinite number of points. Equivalent to PCA, SVD decomposition on the covariance matrix of the 3-D data is also a common pose registration method for 3-D objects. A large number of the shape descriptors use or need PCA or SVD for pose normalization. Typical examples include cords-based shape descriptors, reflective symmetry descriptors (Kazhdan et al., 2003b), statistic moments (Elad et al., 2002), parameterized statistics (Ohbuchi et al., 2002), shape histograms (Ankerst et al., 1999b) and spherical extent functions (Vranic et al., 2001). After translation and rotation of 3-D models, Elad et al. (2002) suggests a flipping using the heaviness

factor in each octant to confirm the orientation of the 3-D objects.

Paquet and Rioux (1997) describe a cords-based shape descriptor. A cord is defined as a ray shooting from the center of mass of the 3-D object to the center of mass of a triangle. Obviously, the number of cords is the same as the number of the triangles in the mesh model. A cords-based representation consists of three histograms: 1) an angle histogram defined by the angles of the cords with the first principal axis, 2) an angle histogram figured out by the angles of the cords with the second principal axis and 3) a length histogram obtained by the norm of the cords. The Euclidean distance between the histograms gives a measure of shape similarity. In practice, this approach is very sensitive to the sampling and level of detail of the mesh model. Furthermore, a PCA normalization is needed to determine the principal axes.

The reflective Symmetry Descriptor proposed in (Kazhdan et al., 2003b) generates a spherical function where each point on the sphere represents a reflective symmetry similarity of the 3-D shape against a 2-D plane through the center of the mass of the 3-D shape. The reflective symmetry similarity of the 3-D shape against a 2-D plane measures the L_2 -difference between the 3-D shape and its reflection. The peaks of the spherical function illustrate the 2-D planes possessing highly reflective symmetry and the valleys suggest the 2-D planes with poor reflective symmetry. The distance between two reflective symmetry descriptors is measured by taking the L_∞ -distance of the two spherical functions. The reflective symmetry descriptor is scale invariant and scale normalization is unnecessary. The spherical function relies on knowing the origin of the model and rotates with rotation of the object. Thus it is neither translation, rotation nor articulation invariant. It is a global measure of 3-D shapes and it is orthogonal to many other shape measures. It can be efficiently computed in a matter of 5 seconds on a standard desktop PC. Kazhdan et al. (2003b) suggest to use it in combination with other existing measures.

Statistical moments have been broadly used for pose registration for 2-D images and 2-D object matching. In (Elad et al., 2002), the authors extend them to the 3-D domain, where the $(p, q, r)^{th}$ moment for an object D is defined as:

$$m_{p,q,r} = \int_{\partial D} x^p y^q z^r dx dy dz,$$

with ∂D denoting the surface. From lower to higher order moments, these statistics provide coarse and detailed shape information. The first order moments are used to translate the center of mass of the object to the origin of the coordinate system. SVD decomposition of the matrix of the second order moments is used for rotating the 3-D models in the canonical coordinates. Experimental results demonstrate that the first six order moments are typically enough for discriminating different objects. To further improve the retrieval system, they associate the SVM, a learning and classification system, to refine the search by allowing the user to select their preferred results and use the feedback to figure out the weights for the SVM. Iterative refinements of the search implicitly picks up the useful orders of moments for different types of models.

The technique of (Ohbuchi et al., 2002) identifies a set of parameterized statistic features of VRML-like models, including degenerated mesh models. Three feature vectors are abstracted along the principal axes of inertia by evaluating the statistics within analysis windows: 1) the moment of inertia about the axis, 2) the average distance to the surfaces from the axis, 3) and the variance of distance to the surfaces from the axis. The technique assumes that the object has been normalized to the canonical axes and that the axes have been divided into equal intervals. The Euclidean distance between the feature vectors gives a rigid comparison and the elastic distance allows a degree of deformation and elongation of portions of the objects.

Ankerst et al. (1999b) introduce 3-D shape histograms as a means of determining similarity for 3-D molecular surfaces. After pose registration of the 3-D objects,

the objects are divided into isometric sectors and isometric shells around the centroid of the objects. The geometric properties, as well as thematic attributes such as physical and chemical properties are analyzed within each partition. Two kinds of distance measures are evaluated on the shape histograms of the database, the Euclidean distance and the quadratic distance. As the experimental results show, quadratic distance accounts for some degree of the measurement error, numerical rounding and rotation.

Vranic et al. (2001) employ a shape descriptor based on spherical functions. A series of dense concentric spheres with increasing radius are constructed centered at the origin of the 3-D object. The spherical functions are gained from the intersections of the concentric spheres with the 3-D shape. An analog of the fourier transform using spherical harmonics on the spherical functions gives a set of feature vectors composed of the complex coefficients. L_1 distance is computed for searching the nearest neighbors of the query model. This ray based spherical harmonics is rotation dependent, while an improved approach presented in (Funkhouser et al., 2003b) modifies it into a rotation invariant one, which is described in 2.3.

Körtgen et al. (2003) consider using 3-D shape contexts as a search key for 3-D object retrieval. The 3-D shape contexts capture the local properties of N uniformly distributed boundary points on the 3-D shape. The representation of shape feature is N histogram vectors, where each histogram defines a distribution of $N-1$ other boundary points with respect to a reference boundary point and each element in a histogram vector estimates the number of $N-1$ other boundary points in a bin with respect to that reference point. The organization of the bins is based on three kinds of models: a Shell Model, a Sector Model and a Combined Model. Only the Shell Model is rotation invariant. All the models are defined by the center of gravity of the object and thus are translation invariant. The matching process consists of local matching and global matching. The local matching differentiates two histograms by considering shape, appearance and position. The global matching tries to find

the best correspondences for two 3-D shapes.

2.2 Translation Invariant Methods

Translation normalization is easily realized by finding the center of the mass of the 3-D object and then applying a translation transformation to move the centroid of the object to the origin of the coordinates. It is known that when the centroid of the object coincides with the origin, the sum of the squared distance of the points on the object is minimized. Thus, there are not too many approaches focusing merely on translation invariance.

Extended Gaussian Images (Horn, 1984) project the points on the object to points on the Gaussian Sphere (unit sphere) having the same unit normal and associate each mapped point on the Gaussian Sphere with the inverse of the Gaussian curvature of the points at the object. The extended Gaussian image is not a rotation invariant shape descriptor. To get every unit normal for each point on the object surface, it is assumed that the object is closed and convex, which is a big limitation. Each 3-D object defines a unique gaussian sphere, while a gaussian sphere may correspond to a number of 3-D objects. This fact implies that extended gaussian images are not discriminative enough to be good shape descriptors.

2.3 Rotation Invariant Methods

As mentioned in the last section, normalization is a way to address the fact that 3-D models are created with arbitrary position and orientation. Proposing a shape descriptor which is insensitive to the rotation of the models is an alternative solution. Examples include spherical harmonics and 3-D Zernike moments. These methods just need to move the object to its center of mass and leave out the step of rotating the object to a canonical coordinate frames.

Kazhdan et al. (2003c) introduce a powerful rotation invariant spherical harmonics descriptor on a voxel grid. This descriptor characterizes frequency information encoding on the concentric spheres with different radii. The spherical functions are decomposed as the sum of their harmonics. The L_2 -norm for each frequency is obtained by summing the harmonics within each frequency. The first L bands of spherical harmonics for each sphere with increasing radius capture the important shape property encoding on the spheres and thus are used to form the shape histogram. It is shown that the norm of L_2 -norm remains the same even if the spherical function rotates. As a result, the 3-D rotation invariant shape descriptor is represented as a 2-D histogram indexed by the frequency and radius. An L_2 -distance comparing two 2-D histograms gives a similarity measure between two 3-D shapes. This spherical harmonics shape descriptor is argued to be the most discriminative technique which supports rotation invariance so far. One limitation is that the spherical functions are drawn on a voxel grid and thus are dependent on the sampling resolution of the 3-D models. The determination of the center of gravity of the spherical functions makes them sensitive to the articulation and deformation of the objects.

Funkhouser et al. (2003b) proposed sketch queries similar to view based matching techniques. 3-D object model is aligned with cartesian axes and is then projected onto 2-D images in 13 orthographic view directions. Hand drawn sketches entered by a user are compared against the 13 projected images for each 3-D model using image matching methods analogous to the 2-D spherical harmonics based on the Euclidean distance transform of the sketch and rendered image. This provides a rotation and reflection invariant descriptor for 2-D images. It offers a brand new interface to allow the user to input the query by drawing sketches (Igarashi et al., 1999), which is different with the traditional way of giving a 3-D model in a file format as a query. This property allows the user to search for a desired model without having to possess a 3-D model in advance. However, drawing a complex

projection of a 3-D model on the sketch interface is not an easy job for many users.

3-D Zernike moments investigated in (Novotni and Klein, 2003) are natural extensions of moments and spherical harmonics based descriptors. The basis functions of the moments are defined by the radius coefficients and spherical harmonics on voxel grids. The orthogonality property of the basis functions enable a more compact representation of the objects than spherical harmonics. Not only are the features of the object along the spherical direction encoded but also the frequency coherence in the radius direction is captured. A vector of 122 scalar values provides discriminating power and saves lots of storage. As the experimental results indicate, 3-D Zenike moments compare favorably to spherical harmonics. In the implementation, the numerical problems of accurate computation of geometric moments and of handling long floating point numbers are considered.

2.4 Translation and Rotation Invariant Methods

Approaches that ignore the origin and main axis of the 3-D models provide matching algorithms that are robust to translation and rotation of the 3-D models and independent of the origin of the object. Many algorithms make an assumption that the models are well defined, have no holes and no cracks, no self-intersections and no missing polygons. The Shape Distribution method proposed in (Osada et al., 2001) is a simple but robust algorithm that can be applied to all kinds of nondegenerate or degenerate models, mesh models or volume objects. The idea is to represent the signature of the 3-D models as a distribution by sampling random points on the uniformly sampled surface of the models, compute shape functions on these points, and distribute shape functions into continuous divided bins. Shape functions include $D2$ space(Euclidean distance between two fixed points), $A3$ space(angles between three fixed points) and some other geometric measurements based on angles, distances and volumes. The L_∞ distance of two distributions is shown to be the best for comparing two models. Though the distribution

abstracted in this way is robust to translation and rotation of objects, it is not robust to articulation and deformation. The sampling process is fast but too simple to provide a discriminative shape distribution to distinguish complex models. Moreover, the histogram of shape distribution also depends largely on the number of bins and the number of sample points. The authors show that 1024 bins and 64 intervals realize the best results.

Graph based methods are a category of shape matching methods that can be performed independently of the position and orientation of objects. These methods represent the shape as a graph structure by decomposing the 3-D objects into parts, following a particular rule. Not only geometric properties but also the topological structure of the objects are taken into account. Regional geometry properties are stored in graph nodes and regional relationships between graph nodes are reflected in the arcs. This type of organization of the overall impression of the object agrees with the way that the human brain memorizes and recognizes objects (Biederman, 1987). The problem of measuring the similarity between two models is equivalent to estimating the minimum distance between two graphs. It allows for partial matching by applying subgraph isomorphism techniques. By searching the correspondences among the graph nodes, matches for the regional geometry are also retrieved. By considering the topological structure of models, the problem breaks into smaller pieces to compensate for the complexity of matching. Examples include generalized cylinders, superquadrics and geons (Binford, 1987), (Marr and Nishihara, 1978), (Pentland, 1986b) (Biederman, 1987). Such approaches build a degree of robustness to deformations and movement of parts, but their representational power is limited by the vocabulary of geometric primitives that are selected.

Hilaga et al. (2001) introduced a Multiresolutional Reeb Graph Based matching method on a 3-D mesh model with triangles evenly scattered. The function μ used to partition the 3-D object for each point is the sum of the geodesic distances of

that point to all points on the surface of the object. The surface of the object is divided into patches according to the ranges of the function μ . Each patch represents a node in the Reeb graph and edges are created for connected patches. By varying the ranges of the function μ , coarse to fine reeb graphs are created at various levels of resolution. The matching focuses on topology matching using a coarse-to-fine strategy while preserving the topology consistency. In general, the reeb graph defined by the geodesic distances provides a translation and rotation robust shape descriptor but cannot avoid being computational expensive. The construction of the reeb graph is quite sensitive to the range of the function μ . Moreover, the topological structure of the reeb graph appears to be quite different from what the human visual system builds, so it does not offer a natural decomposition of a 3-D model. The matching considers a simple node similarity measure based on the area and length of the attributes of the patch.

2.5 Articulation Invariant Methods

One of the most challenging and interesting issues for research in 3-D object matching is the question of how to come up with an articulation invariant method. Articulation allows the models to move around joints. This supports the creation of diverse, natural and interesting objects for the applications in education, entertainment and many other areas.

Sundar et al. (2003) simulates a skeleton graph based matching. The stick-figure of a 3-D model is constructed by a volume thinning strategy. A clustering process is followed to group skeleton parts and then a directed acyclic graph is formed based on a minimum spanning tree. The matching is formulated based on the eigen space based characterization of the graph. This particular skeleton is a coarse abstraction of the 3-D object. The volume thinning process is heuristic and can be sensitive to noise and perturbation on the boundaries. It does not necessarily converge to locating exact skeletal points and the skeleton turns out to have many

unwanted branches, which affects the structure of the graph. The clustering based on proximity is also heuristic and does not offer a unique group of skeletal points. The formulation of the directed acyclic graph based on the length of the edges is very sensitive to the shape of the object. The matching emphasizes the topological structure of the graph, but little emphasis is placed on the geometry of the implied parts.

Tam et al. (2004) consider topology points as features for topological and geometric similarity matching. The topology points are extracted based on the Level Set Diagram algorithm that simulates the marching of a waterfront to locate the critical points: minima, maxima and saddles, which significantly describe the topology change of a smooth model surface. A heuristic rule is opted for reduction of incorrect points marked as saddles. The 3-D mesh is partitioned into N parts with respect to a topology point and a curvature distribution vector for this topology point is obtained from a list of curvature distributions for each part. A bipartite graph is formulated for a query and a model by topology points as nodes with edges measuring the geometric distance of the sum of geodesic distances and curvature distribution vectors. This approach provides a robust measure of similarity for deformable models. However, the localization of topology points is too local and the reduction rule is too heuristic to be robust. The computation of the sum of geodesic distances for each vertex adds a heavy computational cost for the implementation. Though it considers the use of the topological structure of the object, the matching process does not take into account the relationships between the topology points.

View based 3-D object matching can be classified into any one of the above types of 3-D object matching methods, depending on the matching metric used for the 2-D views. Cyr and Kimia (2004) identify a representation of a 3-D object by selecting a set of aspects of the object using a region-growing approach. An unknown view of an unknown object is matched against all the aspects of the known

objects in the database based on a shock graph. The shock graph based matching allows this approach to handle articulation of the object. Their approach can be extended to incorporate various viewpoints of a object rather than "ground views".

Macrini et al. (2002) extend the shock graph matching described in (Siddiqi et al., 1998) to address the problem of 3-D object recognition. A number of different viewpoints of a 3-D model is taken and their shock graphs constructed based on (Siddiqi et al., 1998) are stored in the database. A shock graph of a view of a 3-D query model is generated as a search key for the whole database. Indexing based on the topological structure of the graph and matching adding the geometry information in the nodes, are both explored, accommodating variations due to occlusion and noise.

In this thesis, we propose a medial surface graph based shape matching method which allows matches invariant to a large degree of articulation and deformation of 3-D objects. It also supports translation and rotation invariance. The topology structure is encoded in the graph and geometric properties are stored in the nodes of the graph. It therefore provides an indexing scheme by filtering the similarity of topology structure with respect to the graphs. High ranking models obtained through the indexing are later verified by a graph matching method. The following chapters give a detailed description of the construction of the medial surface based shape descriptor, the indexing, and the matching incorporating geometry and topological information.

Chapter 3

Medial Surfaces and Directed Acyclic Graphs

3.1 Medial Surfaces

The computation of the skeleton is a fundamental problem which the computer vision community has focused on for several years. The skeleton of an object is the locus of the centers of all maximal inscribed balls of the object (Blum, 1973). It was first populated by Blum's grass fire theory: Given the assumption that an object is composed of dense grass, flammable and isotropic homogeneous, place it in a non flammable surrounding space and set a fire from its boundary at the same time; the fire will propagate inward from the boundary until two or more flame fronts meet and the location where the fire stops is the locus of the medial set (see Fig. 3.1)(Bouix, 2003). For further reading, we refer the reader to the upcoming book (Siddiqi and Pizer, 2005) which provides an in-depth account of the mathematics, algorithms and applications of skeletal representations.

3.1.1 Computation of Medial Surfaces

Recent approaches for computing 3-D skeletons include the *power crust* algorithm (Amenta et al., 2001), the *shock scaffold* (Leymarie, 2003) and *average outward flux-based* skeletons (Siddiqi et al., 2002). The first two methods have the advantage that they can be employed on input data in the form of points sampled from an object's surface, and theoretical guarantees on the quality of the results can be provided. Unfortunately, automatic segmentation of the resulting skeletons remains a challenge. The last method assumes that objects have first been voxelized, and this adds a computational burden. However, once this is done the limiting behavior of the average outward flux of the Euclidean distance function gradient vector field can be used to characterize 3-D skeletal points. We choose to employ this latter method since it has the advantage that the digital classification of (Malandain et al., 1993) allows for the taxonomy of generic 3-D skeletal points (Giblin and Kimia, 2004) to be interpreted on a rectangular lattice, leading to a graph of parts.

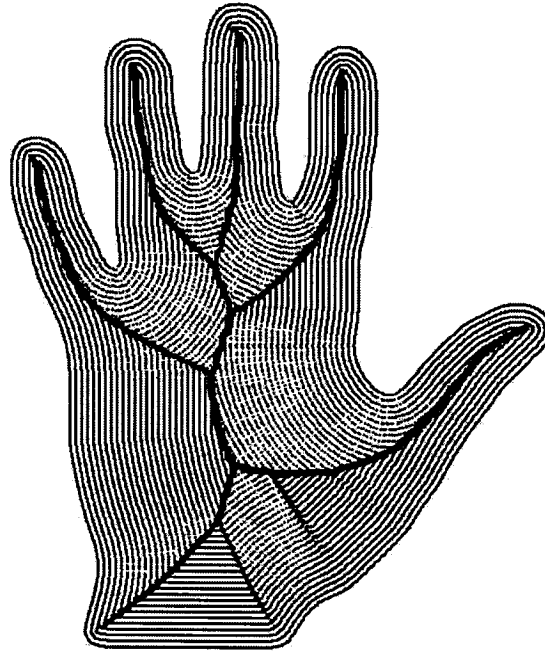


FIGURE 3.1: The medial set (axis in 2-D) of X is the set of points of X simultaneously reached by grass fires initiated from at least two different points of ∂x , the boundary of X . Taken from (Siddiqi et al., 1999b)

Under the assumption that the initial model is given in triangulated form, we begin by scaling all the vertices so that they fall within a rectangular lattice of fixed dimension and resolution. We then sub-divide each triangle to generate a dense intersection with this lattice, resulting in a binary (voxelized) 3-D model. The average outward flux of the Euclidean distance function's gradient vector field is computed through unit spheres centered at each rectangular lattice point, using Algorithm 3.1. This quantity has the property that it approaches a negative number at skeletal points and goes to zero elsewhere (Siddiqi et al., 2002), and thus can be used to drive a digital thinning process, for which an efficient implementation is described in Algorithm 3.2. This thinning process has to be implemented with some care, so that the topology of the object is not changed. This is done by identifying each *simple* or removable point x , for which a characterization based on the 26-neighborhood of each lattice point x is provided in (Malandain et al., 1993). With O being the set of points in the interior of the voxelized object and N_{26}^* being

the 26-neighborhood of x , not including x itself, this characterization is based on two numbers:

1. C^* : the number of 26-connected components 26-adjacent to x in $O \cap N_{26}^*$, and
2. \bar{C} : the number of 6-connected components 6-adjacent to x in $\bar{O} \cap N_{18}$.

It can be shown that a digital point x is *simple* if $C^*(x) = 1$ and $\bar{C}(x) = 1$.

Algorithm 3.1: Average Outward Flux.

Data : Voxelized 3-D Object Model.

Result : Average Outward Flux Map.

Compute the Euclidean distance transform D of the model ;

Compute the gradient vector field ∇D ;

Compute the average outward flux of ∇D :

For (each point x) $AOF(x) = \frac{1}{26} \sum_{i=1}^{26} \langle \hat{N}_i, \nabla D(x_i) \rangle$;

(where x_i is a 26-neighbor of x in 3-D and \hat{N}_i is the outward normal at x_i of the unit sphere centered at x)

The average outward flux based method along with the topology preserving thinning process is applied for every model in the database. Exemplars of a cup, an airplane, a human and a deer are shown along with their corresponding medial surfaces in Fig. 3.2.

3.1.2 Classification and Segmentation of Medial Surfaces

The taxonomy of generic 3-D skeletal points in the continuum, i.e., those which are stable under small perturbations of the object, is provided in (Giblin and Kimia, 2004). Using the notation A_n^k , where n denotes the number of points of contact of the maximal inscribed sphere with the surface and k the order of these contacts, the taxonomy includes: 1) A_1^2 points which form a smooth medial manifold, 2) A_3 points which correspond to the rim of a medial manifold, 3) A_1^3 points which

Algorithm 3.2: Topology Preserving Thinning.

Data : 3-D Object Model, Average Outward Flux Map.
Result : 3-D Skeleton (Medial Surface).
for (each point x on the boundary of the object) **do**
 if (x is simple) **then**
 insert(x , maxHeap) with $\text{AOF}(x)$ as the sorting key for insertion;
while (maxHeap.size > 0) **do**
 $x = \text{HeapExtractMax}(\text{maxHeap})$;
 if (x is simple) **then**
 if (x is an end point) and ($\text{AOF}(x) < \text{Thresh}$) **then**
 mark x as a medial surface (end) point;
 else
 Remove x ;
 for (all neighbors y of x) **do**
 if (y is simple) **then**
 insert(y , maxHeap) with $\text{AOF}(y)$ as the sorting key for insertion;



FIGURE 3.2: A voxelized cup, an airplane, a human and a deer (top row), and their corresponding medial surfaces (bottom row)

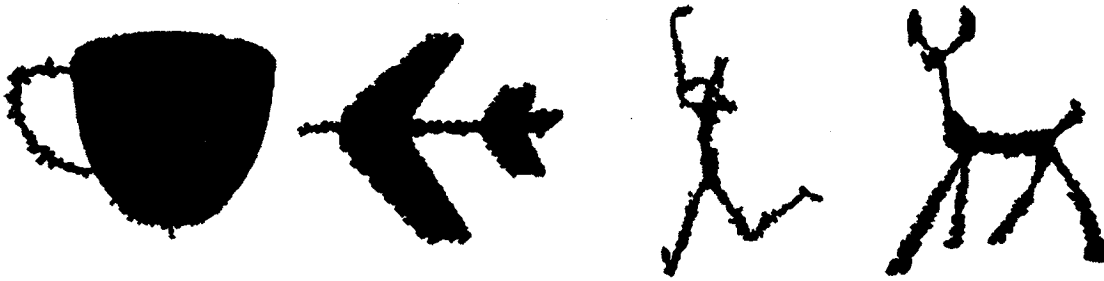


FIGURE 3.3: Classification of medial surface points. Grey, red, yellow, green, cyan and blue corresponds, respectively, to border points, curve points, curve junctions, surface points, surface-curve junction points and surface junctions.

represent the intersection curve of three medial manifolds, 4) an A_1^4 point at the intersection of four A_1^3 curves, and 5) an A_1A_3 point at the intersection between an A_3 curve and an A_1^3 curve.

It is clear from this classification that 3-D skeletons are essentially comprised of medial manifolds, their rims and intersection curves, and this is why we refer to this as a *medial surface* representation. As shown in (Malandain et al., 1993), the numbers C^* and \bar{C} , defined in section 3.1.1, can also be used to classify surface points, rim points, junction points and curve points on a rectangular lattice. These results are summarized in Table 3.1 with examples shown in Fig. 3.3. This suggests the following 3-step approach for segmenting the (voxelized) medial surface into a set of connected parts:

1. Identify all manifolds comprised of 26-connected surface points and border points.
2. Use junction points to separate these manifolds, but allow junction points to belong to all manifolds that they connect.
3. Form connected components with the remaining curve points, and consider these as parts as well.

This process of automatic skeletonization and segmentation is illustrated for four object classes, a cup, an airplane, a human form and a deer in Fig. 3.4.

\bar{C}	C^*	TYPE
0	any	interior point
any	0	isolated point
1	1	border (simple) point
1	2	curve point
1	> 2	curves junction
2	1	surface point
2	> 2	surface-curve(s) junction
> 2	1	surfaces junction
> 2	≥ 2	surfaces-curves junction

TABLE 3.1: The topological classification of (Malandain et al., 1993).



FIGURE 3.4: Thin set of medial surfaces of a cup, an airplane, a human and a deer form (top row), with segmented medial surfaces shown in the bottom row.

3.2 Directed Acyclic Graph

The segmented medial surfaces are now interpreted as directed acyclic graphs (DAGs). The topological properties of the 3-D models are encoded in the structure of the directed acyclic graph along with the geometrical properties of the parts of the 3-D models stored in the nodes of the graph. In the remainder of the thesis, we shall use these DAGs for 3-D model matching and indexing.

3.2.1 Definition and Properties of a Directed Acyclic Graph

Definition: A directed graph with no path that starts and ends at the same vertex, is also known as a acyclic directed graph or an oriented acyclic graph.

Directed acyclic graphs are an important class of graphs and have many applications. In the context of 3-D object retrieval, many good properties render the directed acyclic graph a powerful tool :

1. **TOPOLOGICAL STRUCTURE.** A node in the graph can correspond to a part in the 3-D object. Edges in the graph indicate the connectivity between the nodes.
2. **GEOMETRIC PROPERTIES.** Geometric information about a part of a 3-D object is contained in a node, permitting detailed comparisons when necessary but also allowing for invariance to part articulation.
3. **GRAPH MATCHING.** Much pioneering work has been done on the subject of graph matching. By comparing two directed acyclic graphs based on topological structure and geometric properties, one can develop a similarity measure between the two 3-D objects.

3.2.2 Forming Directed Acyclic Graphs from Segmented Medial Surfaces

We now propose an interpretation of the segmented medial surface as a directed acyclic graph (DAG). We begin by introducing a notion of *saliency* which captures the relative importance of each component. Consider that the envelope of maximal inscribed spheres of appropriate radii placed at all skeletal points reconstructs the original object's volume (Blum, 1973). The contribution of each component to the overall volume can thus be used as a measure of its significance. Since the spheres associated with adjacent components can overlap, an objective measure of

component j 's saliency is given by:

$$Saliency_j = \frac{Voxels_j}{\sum_{i=1}^N Voxels_i}.$$

Here we assume that there are N components and $Voxels_i$ is the number of voxels *uniquely* reconstructed by component i . We propose the following construction of a DAG, using each component's saliency. Consider the most salient component as the root node (level 0), and place components to which it is connected as nodes at level 1. Components to which these nodes are connected are placed at level 2, and this process is repeated in a recursive fashion until all nodes are accounted for. The graph is completed by drawing edges between all pairs of connected nodes, in the direction of increasing levels. However, to allow for 3-D models comprised of disconnected parts we introduce a single dummy node as the parent of all DAGs for a 3-D model. This process is illustrated in Fig. 3.5 (bottom row) for the human and chair models, with the saliency values shown within the nodes. Note how this representation captures the intuitive sense that the human is a torso with attached limbs and a head, a chair is a seat with attached legs and a back, etc. Our DAG representation of the medial surface is quite different than the graph structure that follows from a direct use of the taxonomy of 3-D skeletal points in the continuum (Giblin and Kimia, 2004). The latter is more complex and does not naturally lend itself to hierarchical structure indexing and matching algorithms, which we describe next in the following chapter.

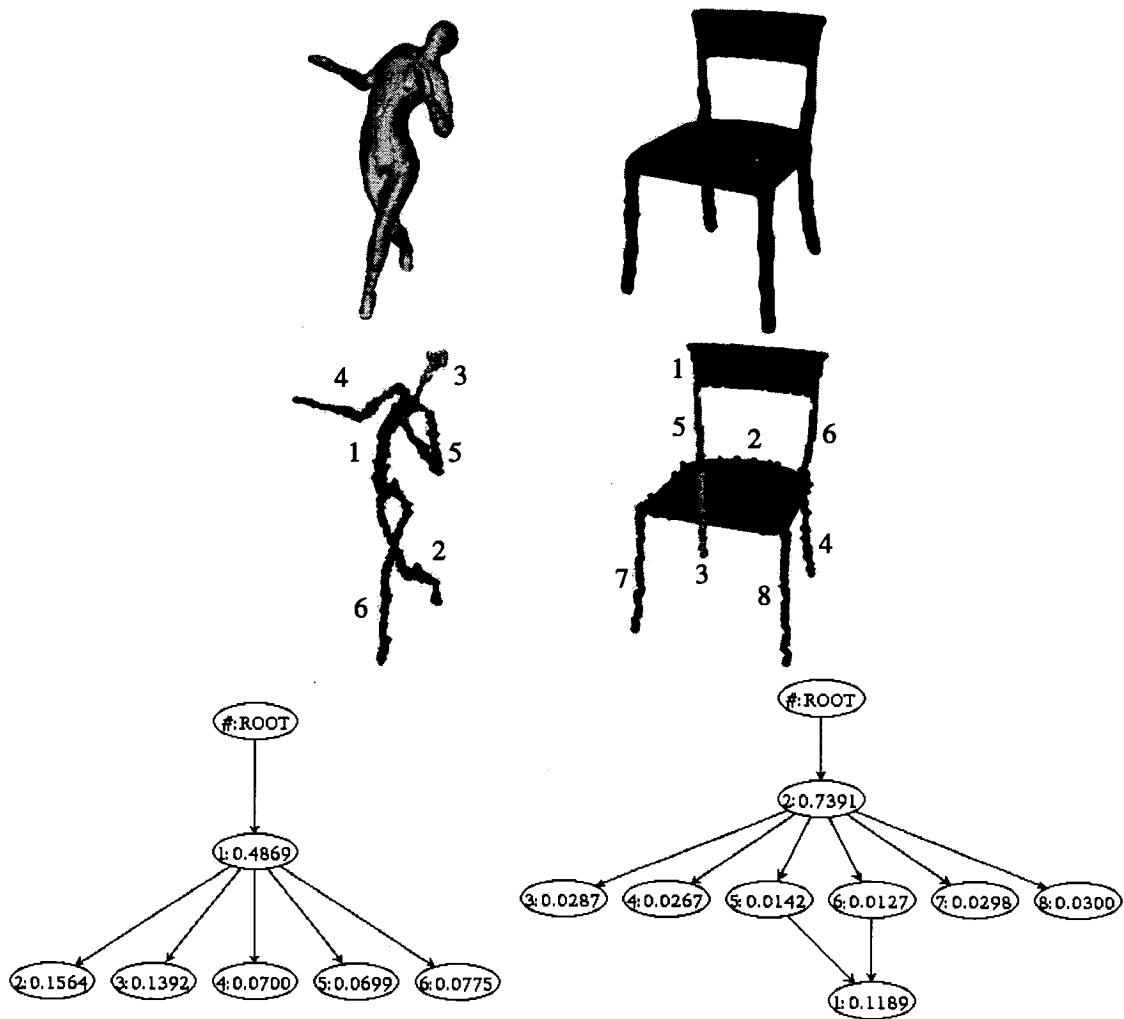


FIGURE 3.5: A voxelized human form and chair (top row), their segmented medial surfaces (middle row). A hierarchical interpretation of the medial surface, using a notion of part saliency, leads to a directed acyclic graph DAG (bottom row). The nodes in the DAGs have labels corresponding to those on the medial surface, and the saliency of each node is also shown.

Chapter 4

Indexing and Matching

In previous chapters we have proposed a medial-surface based DAG as a shape descriptor which is designed to be robust to articulation. In this chapter we shall use this graph representation for indexing and matching. The implicit assumption we make is that similar 3-D models have similar topologies. Comparing the hierarchical structures of two DAGs yields an efficient preliminary measure for the similarity between two 3-D models. With this idea, we use an indexing method based on the DAG topologies and select a small number of candidates for the matcher to verify (Siddiqi et al., 1999a; Shokoufandeh et al., 2005). This facilitates the retrieval process when the database is very large.

Articulations of parts of 3-D models are often associated with junctions of medial surface manifolds. Therefore, when we partition the medial surfaces at junctions, variations due to articulation are handles. The topological structure of the 3-D models is reflected by the connectivity of the graphs. The algorithm we use as had some success in 2-D shape matching using skeletal graphs (Siddiqi et al., 1999a) and is essentially a greedy best-first search bi-partite graph matching technique. The method considers both node information (part geometry) and the sub-structure rooted at each node (part connectivity). We begin by discussing the indexing problem.

4.1 Indexing

The method we use is described in much greater detailed in (Shokoufandeh et al., 2005).

4.1.1 Criteria for an Effective Index

The goal of indexing is to search a small set of candidates which share similar properties with the query. The criteria for an effective index can be concluded as follows: i) provide a small number of candidates carrying similar shape properties

with the query. ii) take possession of a large proportion of the desired candidates over all the retrieved candidates. iii) establish computationally efficient search over a large portion of database.

A linear search of the 3-D model database, i.e., comparing the query 3-D object model to each 3-D model and selecting the closest one, is inefficient for large databases. An indexing mechanism is therefore essential to select a small set of candidate models to which the matching procedure is applied. When working with hierarchical structures, in the form of DAGs, indexing is a challenging task, and can be formulated as the fast selection of a small set of candidate model graphs that share a subgraph with the query. But how do we test a given candidate without resorting to subgraph isomorphism and its intractability? The problem is further compounded by the fact that due to perturbation and noise, no significant isomorphisms may exist between the query and the (correct) model. Yet, at some level of abstraction, the two structures (or two of their substructures) may be quite similar. Thus, our indexing problem can be reformulated as finding model (sub)graphs whose structure is *similar* to the query (sub)graph.

Choosing the appropriate level of abstraction with which to characterize a DAG is a challenging problem. We seek a description that, on the one hand, provides the low dimensionality essential for efficient indexing, while on the other hand, is rich enough to prune the database down to a tractable number of candidates. We adopt the approach of Siddiqi et al. (1999b), which draws on the eigenspace of a graph to characterize the topology of a DAG with a low-dimensional vector that will facilitate an efficient nearest-neighbor search in a database. The approach begins by noting that any graph can be represented as an antisymmetric $\{0, 1, -1\}$ node-adjacency matrix (which we will subsequently refer to as an adjacency matrix), with 1's (-1's) indicating a forward (backward) edge between adjacent nodes in the graph (and 0's on the diagonal). The eigenvalues of a graph's adjacency matrix encode important structural properties of the graph, characterizing the degree

distribution of its nodes. Moreover, it has been shown that the magnitudes of the eigenvalues (and hence their topological characterization) are stable with respect to minor perturbations of graph structure due to, for example, noise, segmentation error, or minor within-class structural variation.

One simple structural abstraction would be a vector of the sorted magnitudes of the eigenvalues of a DAG's adjacency matrix¹. However, for large DAGs, the dimensionality of the index would be prohibitively large (for efficient nearest-neighbor search), and the descriptor would be global (prohibiting effective indexing of query graphs with added or missing parts). This problem can be addressed by exploiting eigenvalue sums rather than the eigenvalues themselves, and by computing both global and local structural abstractions Siddiqi et al. (1999b). Let V be the root of a DAG whose maximum branching factor is Δ , as shown in Fig. 4.1. Consider the subgraph rooted at node a , the first child of V , and let the out-degree of a be k_1 . We compute the sum S_1 of the magnitudes of the k_1 largest eigenvalues of the adjacency sub-matrix defined by the subgraph rooted at node a , with the process repeated for the remaining children of V . The sorted S_i 's become the components of a Δ -dimensional vector $\chi(V)$, called a *topological signature vector* (TSV), assigned to V . If the number of S_i 's is less than Δ , the vector is padded with zeroes. We can recursively repeat this procedure, assigning a vector to each nonterminal node in the DAG, computed over the subgraph rooted at that node.

In summing the magnitudes of the eigenvalues, some uniqueness has been lost in an effort to reduce dimensionality. The k_i largest eigenvalues are chosen for two reasons: 1) the largest eigenvalues are more informative of subgraph structure, and 2) by summing k_i elements, the sums are effectively normalized according to the local complexity of the subgraph root, thereby distinguishing subgraphs that have richer part structure at coarser levels. The dimensionality of the TSV, χ , is bounded by the maximum branching factor in the graph, which is typically small, and not

¹Since the eigenvalues of an antisymmetric matrix are complex we utilize the magnitude of an eigenvalue.

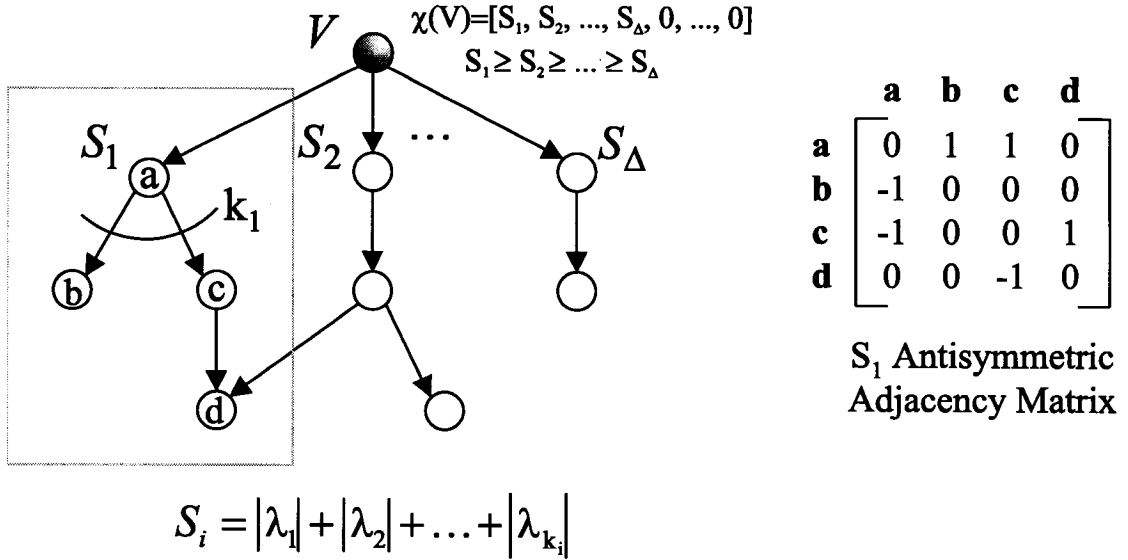


FIGURE 4.1: Forming a Low-Dimensional Vector Description of Graph Structure. At node a , we compute the sum of the magnitudes of the k_1 largest eigenvalues of the adjacency submatrix defined by the subgraph rooted at a . The sorted sums S_i become the components of $\chi(V)$, the *topological signature vector* (or TSV) assigned to V .

by the size of the graph, which can be large for complex 3-D models.

4.1.2 Formulating an Index

Indexing now amounts to a nearest-neighbor search in a model database, as shown in Fig. 4.2. The TSV of each non-leaf node in each model DAG defines a vector location in a low-dimensional Euclidean space (the model database) at which a pointer to the model containing the subgraph rooted at the node is stored. At indexing time, a TSV is computed for each non-leaf node, and a nearest-neighbor search is performed using each “query” TSV. Each TSV “votes” for nearby “model” TSVs, thereby accumulating evidence for models that share the substructure defined by the query TSV. Indexing could, in fact, be accomplished by indexing solely with the root of the entire query graph. However, in an effort to accommodate large-scale perturbation (which corrupts all ancestor TSVs of a perturbed subgraph), indexing is performed locally (using all non-trivial subgraphs, or “parts”) and evidence

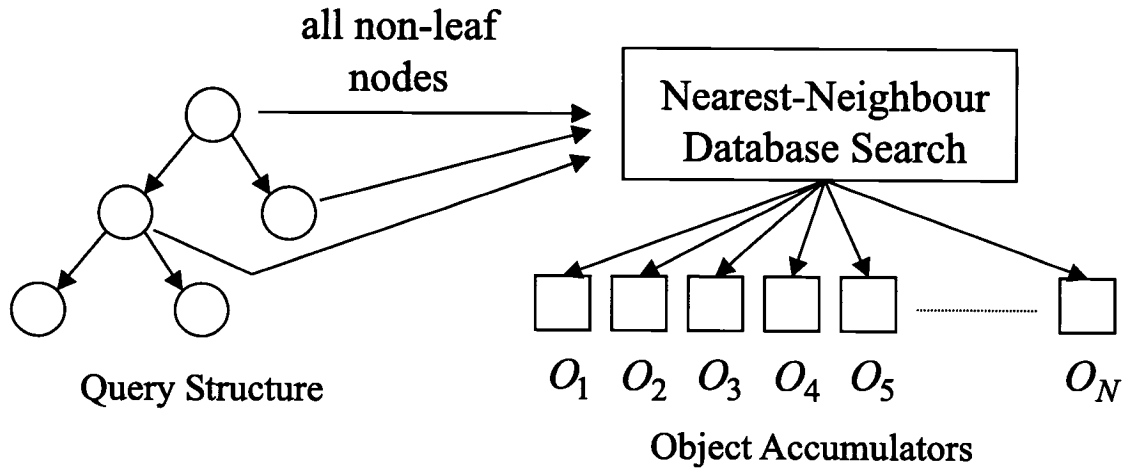


FIGURE 4.2: Indexing Mechanism. Each non-trivial node (whose TSV encodes a topological abstraction of the subgraph rooted at the node) votes for models sharing a structurally similar subgraph. Models receiving strong support are candidates for a more comprehensive matching process.

combined. The result is a small set of ranked model candidates which are verified more extensively using the matching procedure described next.

4.2 Matching

The matching algorithm uses: a) a measure of topological structure given by the sum of the magnitudes of the eigen values of the adjacency matrix rooted at each node and b) a node similarity measure based on a mean curvature signature vector.

4.2.1 Node Similarity

The graph matching algorithm requires a node similarity function that compares the shapes of the 3-D parts associated with two nodes. A variety of the measures used in the literature as signatures for indexing entire 3-D models could be used to compute similarities between two parts (nodes) Osada et al. (2002); Ankerst et al. (1999a); Vranic and Saupe (2001); Elad et al. (2001); Kazhdan et al. (2003a). Some care would of course have to be taken in the implementation of methods which

require a form of global alignment. We have opted for a much simpler measure, which is based on the use of a mean curvature histogram.

First, consider the volumetric part that a node i represents, along with its Euclidean distance function D . At any point within this volume, the mean curvature of the iso-distance level set is given by $\text{div}(\frac{\nabla D}{\|\nabla D\|})$. On a voxel grid with unit spacing the observable mean curvatures are in the range $[-1, 1]$. We compute a histogram of the mean curvature over all voxels in the volumetric part, over this range, using a fixed number of bins N . A mean curvature histogram vector \vec{M}_i is then constructed with entries representing the fraction of total voxels in each bin. The similarity between two nodes i and j is then given by:

$$\text{Similarity}(i, j) = [1 - \underbrace{\sqrt{\sum_{k=1}^N [\hat{M}_i(k) - \hat{M}_j(k)]^2}}_{\text{Distance}(i,j)}].$$

By construction, this similarity function is in the interval $[0, 1]$. This measure could be further modified to take into account overall part sizes. In our experiments we choose not to do this since our object models have undergone a global size normalization.

4.2.2 Bi-partite Graph Matching Algorithm

Each of the top-ranking candidates emerging from the indexing process must be verified to determine which is most similar to the query. If there were no noise our problem could be formulated as a graph isomorphism problem for vertex-labeled graphs. With limited noise, we would search for the largest isomorphic subgraph between query and model. Unfortunately, with the presence of significant noise, in the form of the addition and/or deletion of graph structure, large isomorphic subgraphs may simply not exist. This problem can be overcome by using the same eigen-characterization of graph structure used for indexing Siddiqi et al. (1999b).

As we know, each node in a graph (query or model) is assigned a TSV, which reflects the underlying structure in the subgraph rooted at that node. If we simply discarded all the edges in our two graphs, we would be faced with the problem of finding the best correspondence between the nodes in the query and the nodes in the model; two nodes could be said to be in close correspondence if the distance between their TSVs (and the distance between their domain-dependent node labels) was small. In fact, such a formulation amounts to finding the maximum cardinality, minimum weight matching in a bipartite graph spanning the two sets of nodes. At first glance, such a formulation might seem like a bad idea (by throwing away important graph structure) until one recalls that the graph structure is effectively encoded in the node's TSV. Is it then possible to reformulate a noisy, largest isomorphic subgraph problem as a simple bipartite matching problem?

Unfortunately, in discarding all the graph structure, the underlying hierarchical structure has also been discarded. There is nothing in the bipartite graph matching formulation that ensures that hierarchical constraints among corresponding nodes are obeyed, i.e., that parent/child nodes in one graph don't match child/parent nodes in the other. This reformulation, although softening the overly strict constraints imposed by the largest isomorphic subgraph formulation, is perhaps too weak. Since no polynomial-time solution is known to exist for enforcing the hierarchical constraints in the bipartite matching formulation, an approximate solution to finding corresponding nodes between two noisy, occluded DAGs, subject to hierarchical constraints, is sought Siddiqi et al. (1999b); Shokoufandeh et al. (2002).

The key idea is to use a modification of Reyner's algorithm Reyner (1977), that combines the above bipartite matching formulation with a greedy, best-first search in a recursive procedure to compute the corresponding nodes in two rooted DAGs, as shown in Fig. 4.3. As in the above bipartite matching formulation, the maximum cardinality, minimum weight matching in the bipartite graph spanning the two sets of nodes from the query and model graphs, is computed, as shown in Fig. 4.3(a).

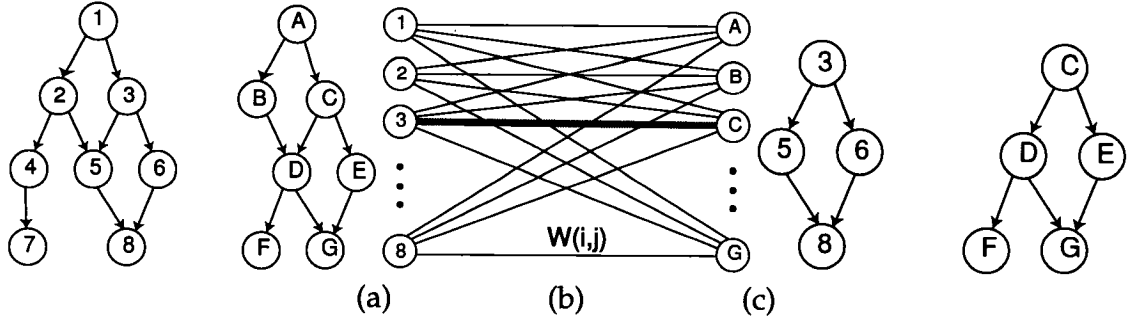


FIGURE 4.3: Matching Algorithm. Given two graphs to be matched (a), form a bipartite graph (b) spanning their nodes but excluding their edges. Edge weights ($W(i, j)$) not only encode node content similarity (see Section 4.2.1), but the structural similarity of their underlying subgraphs, as encoded by the difference in their TSV's. The best matching pair is identified, the two nodes are removed from their respective graphs and added to the solution set of correspondences, and the process applied recursively to their subgraphs (c).

Edge weight encodes a function of both topological similarity as well as domain-dependent node similarity, described in the following paragraph. The result will be a selection of edges yielding a mapping between query and model nodes. As mentioned above, the computed mapping may not obey hierarchical constraints. They therefore greedily choose only the best edge (the two most similar nodes in the two graphs, representing in some sense the two most similar subgraphs), as shown in Fig. 4.3(b), add it to the solution set, and recursively apply the procedure to the subgraphs defined by these two nodes, as shown in Fig. 4.3(c). Unlike a traditional depth-first search, which backtracks to the next statically-determined branch, this algorithm effectively recomputes the branches at each node, always choosing the next branch to descend in a best-first manner. In this way, the search for corresponding nodes is focused in corresponding subgraphs (rooted DAGs) in a top-down manner, thereby ensuring that hierarchical constraints are obeyed. The structural abstraction offered by the TSV effectively unifies the indexing and matching procedures, providing an efficient model retrieval mechanism.

Chapter 5

Experiments

5.1 Construction of the Database

In order to test the power of our indexing and matching algorithms using medial surface-based DAGs, we have considered using the Princeton Shape Benchmark Shilane et al. (2004). This standardized database, which contains 1,814 3-D object models organized by class, is an effective one for comparing the performance of a variety of methods including those in Funkhouser et al. (2003a); Kazhdan et al. (2003a); Osada et al. (2002); Ankerst et al. (1999a); Vranic and Saupe (2001); Elad et al. (2001). A majority of the models in the database correspond to rigid, man-made objects for which a notion of a centroid applies. The natural objects include a variety of animals (including humans), trees, plants and body parts. However, only a limited number of these have articulated or deformed parts. When such models are present, the precise nature of part articulation typically defines a unique base level category. For example, *animal-biped-human* contains human models which are upright, *animal-biped-human-arms-out* contains similar models with outstretched hands and *animal-biped-human-walking* contains those in a walking pose. Results reported in Shilane et al. (2004) indicate that a number of global shape descriptors perform suitably at such base levels of classification, but degrade rapidly at coarser levels, e.g., the classification *human*. In the context of *generic* 3-D model retrieval, such coarser levels in fact correspond to the notion of a *basic level* or *entry level* categorization Rosch (1978); Biederman (1987), whose exemplars might reflect a variety of complex poses and articulations, such as those seen in Fig. 1.1. Our matching and indexing algorithms have the potential to work at this more challenging level, because they use intuitive part-based representations.

To demonstrate this, we have constructed our own database adopting some of the models in the Princeton repository, but adding several of our own. Our database includes a total of 320 exemplars taken from several *basic level* object classes (hands, humans, teddy bears, glasses, pliers, tables, chairs, cups, airplanes, birds, dolphins, dinosaurs, four-legged animals, fish). A large number of these

models are shown in Fig. 5.1. We divide these classes into two categories, those with significant part articulation, and those with moderate or no part articulation. In our experiments we merge the categories “four-legged” and “dinosaurs”, treating them as a single category “four-limbs”

5.2 Indexing Experiments

In order to test our indexing algorithm, which utilizes only the topological structure of medial surface-based DAGs, we carried out two types of experiments. In the first we evaluated percentage recall. For a number of rank thresholds the percentage of models in the database in the same category as a query (not including the query itself) with higher indexing rank, are shown in Fig. 5.2. The results indicate that on average 70% of the desired models are in the top 80 (25% of 320) ranks. In the second experiment we examine the average ranks according to object classes. For all queries in a class the rank of all other objects in that class is computed. The ranks averaged across that class are shown in Fig. 5.3. The results indicate that for 9 of the 13 object classes the average rank is in the top 80 (25% of 320). The higher average ranks for the remaining classes are due to the fact that certain categories have similar part decompositions. In such cases topological structure on its own is not discriminating enough, and part shapes also have to be taken into account.

It should be emphasized that the indexer is a fast screener which can quickly prune the database down to a much smaller set of candidates to which the matcher can be applied. Furthermore, the eigen characterization used to compute the index is also used at matching time, so the same eigen structure calculation is exploited for both steps. The systems against which we evaluate the matcher in the following section run a linear search on the entire database for each query. This approach does not scale well, since the indexing problem is essentially ignored.

5.3 Matching Experiments

On a large database we envision running the indexing strategy first to obtain a smaller subset of candidate 3-D models and to match the query only against these. However, given the moderate size of our database we were able to generate the $320 \times 320 = 102,400$ pairs of matches in a matter of 15-20 minutes on a 3.0 GHz desktop PC. We compare the results using medial surfaces (MS) with those obtained using harmonic spheres (HS) Kazhdan et al. (2003c) and shape distributions (SD) Osada et al. (2002). The pair-wise distances between models using harmonic spheres were obtained using Michael Kazhdan’s executable code (<http://www.cs.jhu.edu/~misha>) and those using shape distributions were based on our own implementation of the algorithm described in Osada et al. (2002). For both HS and SD we used as input a mesh representation of the bounding voxels of the voxelized model used for MS. The comparisons are performed using the standard information retrieval notion of *recall versus precision*, where curves shifted upwards and to the right indicate superior performance.

The results for objects with articulating parts are presented in Fig. 5.4. For the category “teddy” both MS and HS give excellent results and for all other categories MS outperforms the other two techniques. Fig. 5.5 shows the results for objects with moderate or no part articulation. For categories in the top row MS gives superior results. For categories in the middle row HS gives slightly better results than MS, but both are significantly superior to SD. For categories in the third row the results are comparable for birds, but for four-limbs and fishes, both HS and SD out perform MS. In Table 5.1 we show the average similarity scores using MS, organized by object class. Red and blue boxes are drawn, respectively, around the two highest similarity scores. In all cases the highest score coincides with the correct object class. Overall these results demonstrate the significant potential of medial surface based representations and their graph spectra for *generic* level 3-D model retrieval, despite substantial articulation of parts.



























Query													
	.61	.37	.00	.45	.23	.20	.02	.02	.10	.00	.09	.16	.26
	.37	.38	.00	.21	.25	.18	.12	.10	.18	.02	.18	.23	.25
	.00	.00	.51	.29	.17	.15	.07	.03	.00	.15	.02	.00	.07
	.45	.21	.29	.64	.34	.23	.04	.04	.00	.01	.05	.04	.28
	.23	.25	.17	.34	.43	.24	.16	.15	.12	.04	.22	.06	.19
	.20	.18	.15	.23	.24	.28	.20	.22	.14	.07	.26	.05	.14
	.02	.12	.07	.04	.16	.20	.51	.46	.37	.08	.45	.00	.09
	.02	.10	.03	.04	.15	.22	.46	.53	.29	.03	.47	.02	.06
	.10	.18	.00	.00	.12	.14	.37	.29	.58	.04	.31	.02	.23
	.00	.02	.15	.01	.04	.07	.08	.03	.04	.48	.08	.15	.07
	.09	.18	.02	.05	.22	.26	.45	.47	.31	.08	.56	.02	.12
	.16	.23	.00	.04	.06	.05	.00	.02	.02	.15	.02	.71	.21
	.26	.25	.07	.28	.19	.14	.09	.06	.23	.07	.12	.21	.40

TABLE 5.1: Average Matching Results Using MS. Each object in the database is matched against all the other objects in the database. Each cell shows the average similarity between objects selected from two fixed object classes. In each row red and blue boxes are drawn, respectively, around the two highest average similarity scores. In all cases the highest score coincides with the correct object class. In most cases there is also a very significant difference between the top two average similarity scores.

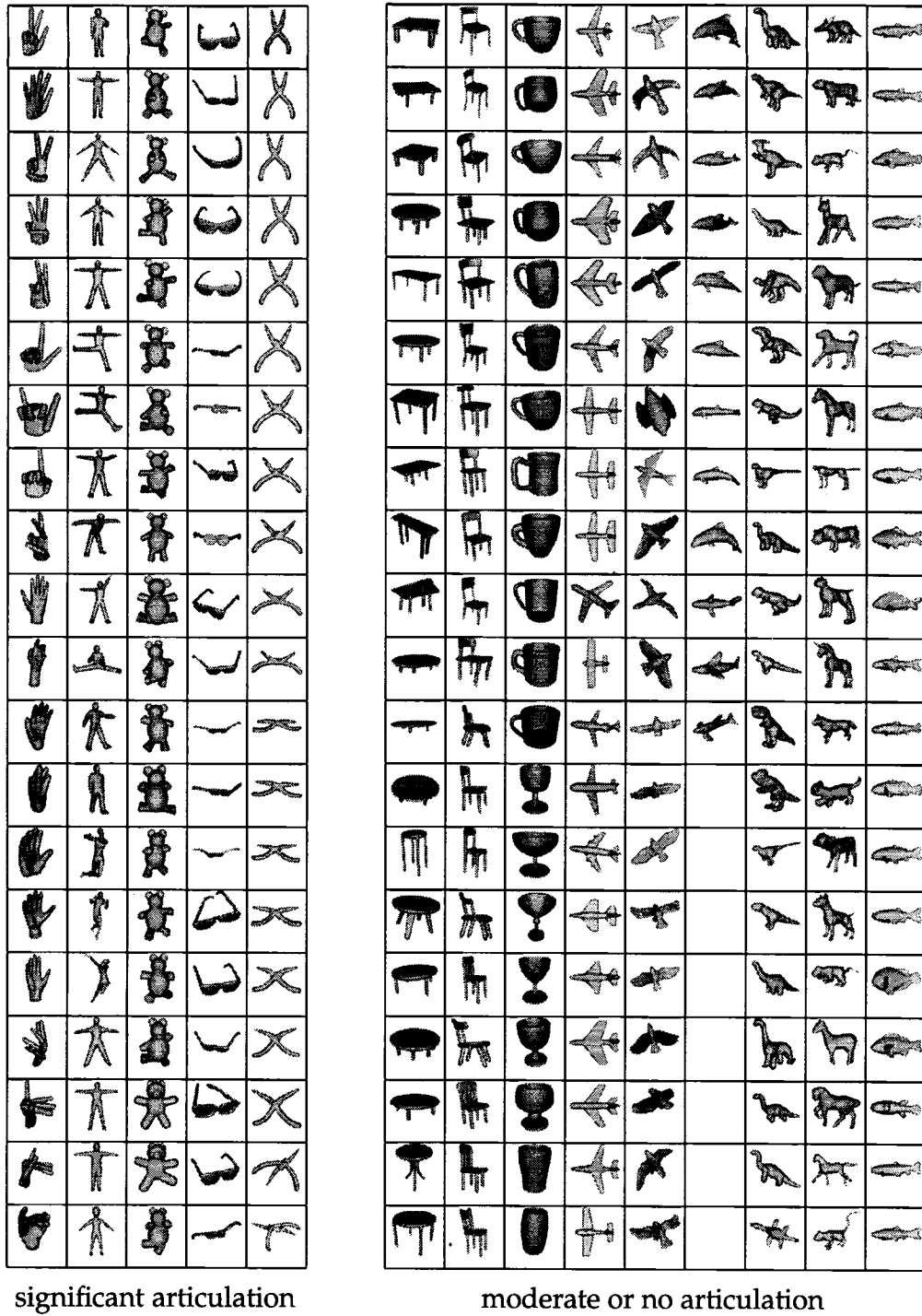


FIGURE 5.1: Database Exemplars. 20 members are shown from each of the object classes (with the exception of the class dolphins which has fewer exemplars). Exemplars from classes on the left have significant part articulation of a complexity not seen in the Princeton Shape Benchmark. Note that we treat the dinosaurs and the four-legged animals as members of a single object class “four-limbs”.

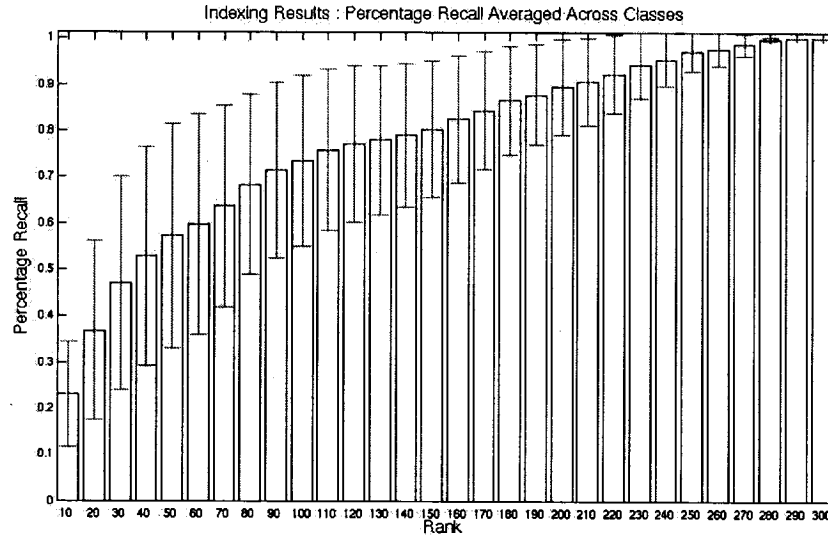


FIGURE 5.2: Indexing Results: Percentage Recall. For several rank thresholds, $N = 10, 20, \dots$, we plot the percentage of models in the database in the same category as the query (not including the query itself) with indexing rank $\leq N$. The results averaged across all classes are shown along with error bars depicting ± 1 standard deviation.

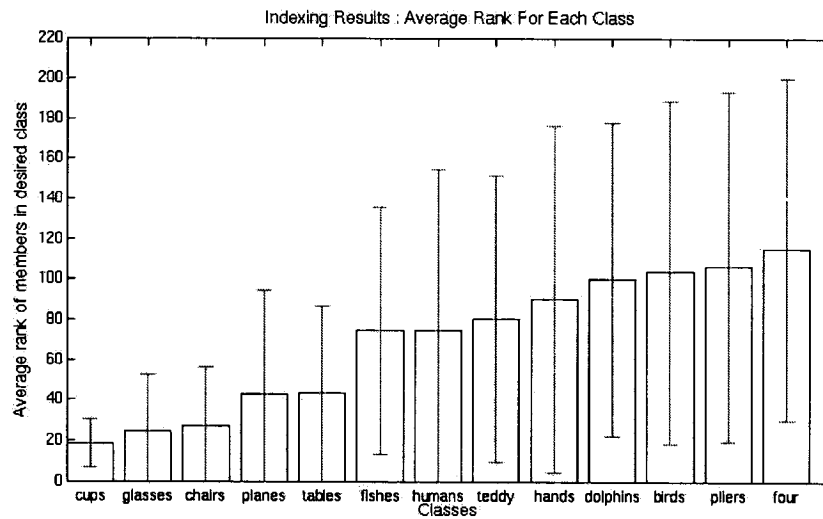


FIGURE 5.3: Indexing Results: Average Ranks. For all queries in a class the rank of all other objects in that class are computed. The ranks averaged across that class are shown, along with error bars depicting ± 1 standard deviation.

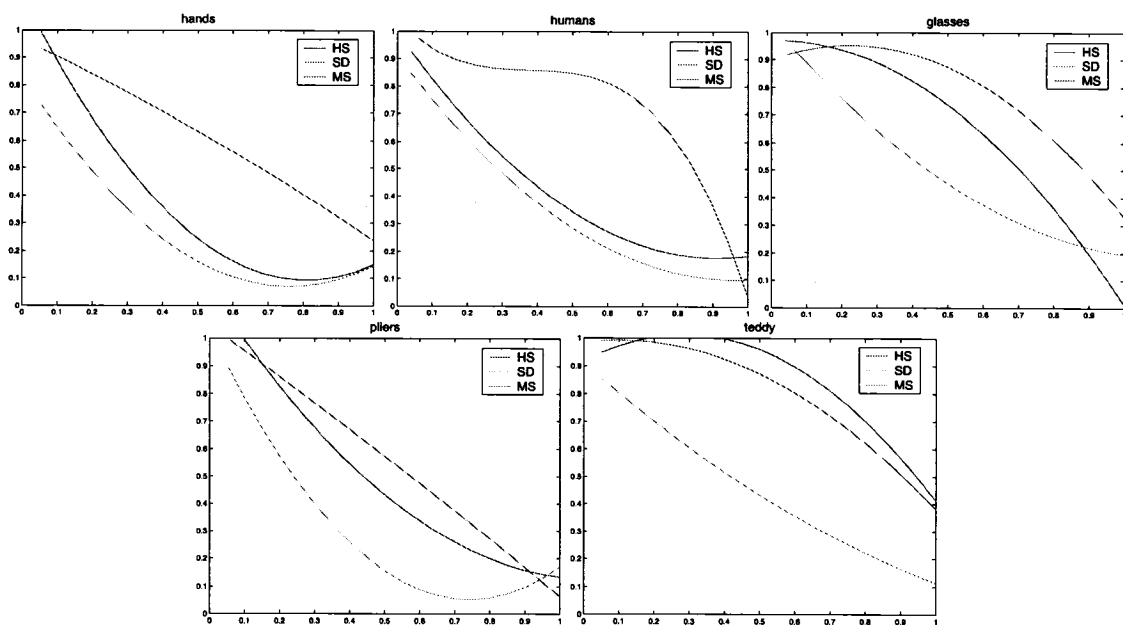


FIGURE 5.4: Precision (y axis) versus Recall (x axis): Objects with articulating parts. The results using medial surfaces (MS) are shown in red, those using harmonic spheres (HS) are shown in blue and those using shape distribution (SD) are shown in green. The results obtained using MS are superior for all categories with the exception of the category “teddy” for which both HS and MS give excellent results.

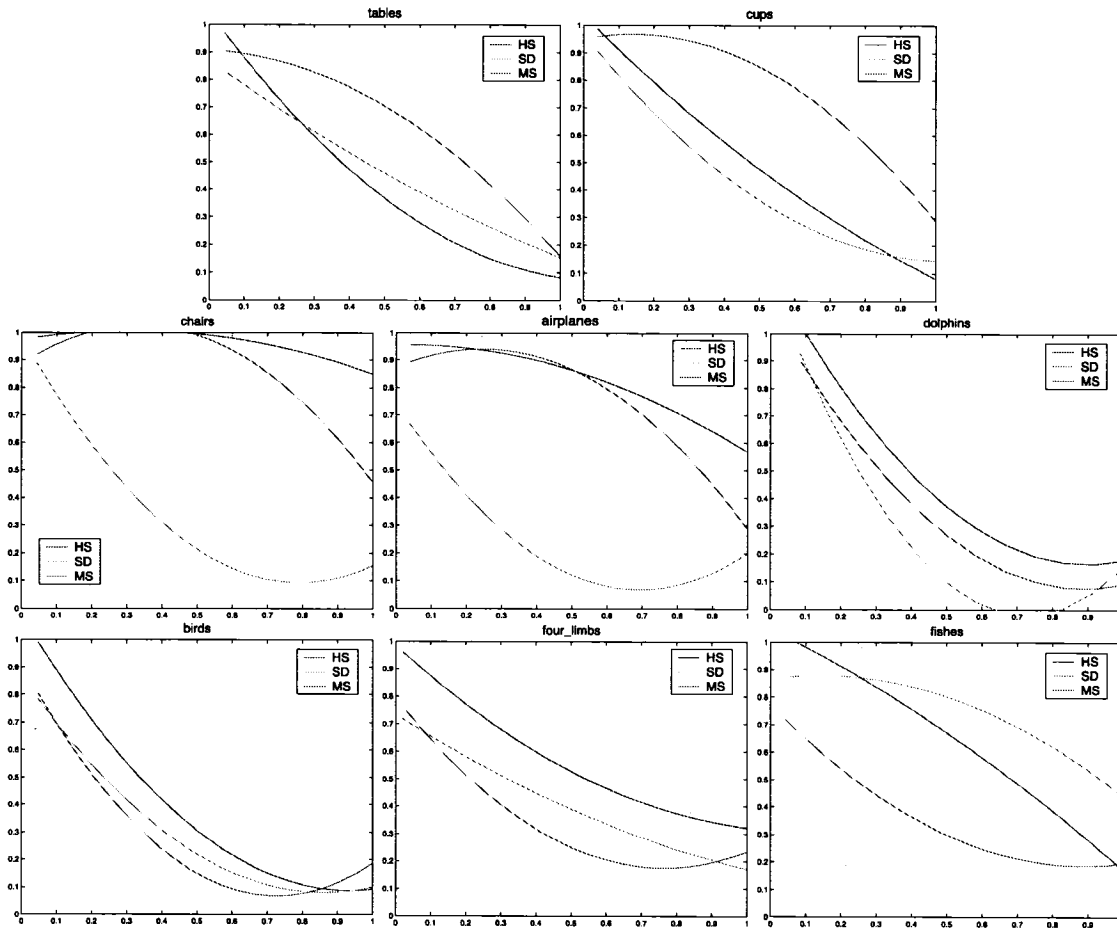


FIGURE 5.5: Precision (y axis) versus Recall (x axis): Objects with moderate or no articulation. The results using medial surfaces (MS) are shown in red, those using harmonic spheres (HS) are shown in blue and those using shape distribution (SD) are shown in green. For categories in the top row MS gives superior results. For categories in the middle row HS gives slightly better results than MS, but both are superior to SD. For categories in the third row the results are comparable for birds, but for four-limbs and fishes, both HS and SD out perform MS.

Chapter 6

Conclusions

In this thesis, we have addressed the problem of 3-D model indexing and matching based on medial surfaces and associated directed acyclic graphs. This last and final chapter emphasizes again the motivation for our work, and summarizes its main contributions. We also discuss a number of possible directions for future work.

6.1 Discussion

With the growth in availability of 3-D object models, the 3-D model retrieval problem has gained significance. We have argued that most current techniques do not use shape descriptors that are invariant to articulation and deformation of parts. In this thesis, we have proposed a matching method that is appropriate in this setting, and is based on a decomposition using medial surfaces.

We have proposed a graph representation with nodes corresponding to parts and edges corresponding to links between adjacent parts. In this way, matching based on an eigen characterization of the graphs can accommodate articulation and deformations. Moreover, indexing schemes developed in the literature can be readily employed when the database is large.

6.2 Contributions and Summary

We propose a novel method for 3-D object indexing and matching based on medial surfaces. To our knowledge, this is the first time that medial surfaces have been adopted and implemented for 3-D object indexing and matching. This thesis makes three main contributions.

First, we have employed an average outward flux based method for computing and segmenting medial surfaces (Siddiqi et al., 1999a; Bouix, 2003). Using the junction points, we decompose the medial surface into parts which in turn reflect a natural segmentation of the underlying 3-D object. We have proposed a DAG

representation of the medial surface which captures a notion of part saliency and allows for existing matching and indexing methods to be applied (Shokoufandeh et al., 2005).

Second, we have built on algorithms in the computer vision literature to address the problem of 3-D model indexing and matching in a uniform framework. Few techniques have been demonstrated in the context of 3-D model matching and indexing. The indexing step can facilitate retrieval by filtering a number of undesired 3-D objects after which the matcher is applied only to the subset of candidates that survive. Here the indexing is based on the topological structure of the DAG. The matcher considers both the topological structure of the DAG and the geometric properties stored in the nodes. The correspondences among the nodes between the DAGs are determined and consistent topology is preserved.

Finally, we have presented extensive indexing and matching results on a database of object models organized according to a generic level of categorization, but with significant articulation and deformation of parts. Few 3-D shape matching techniques have been tested on a database containing a significant number of articulated objects. We have not only collected many articulated object models, but have also modified some to create additional exemplars. The experimental results on a database of 320 objects demonstrate that our method significantly outperforms the techniques of Shape Distributions (Osada et al., 2001) and Harmonics Spheres (Kazhdan et al., 2003c) for articulated objects.

Whereas all the pieces of this system have been developed in past work, putting them together and demonstrating them in the context of 3-D model retrieval with comparative results against competing methods has been the focus of this thesis.

6.3 Future Work

There are several possible directions for future work.

First, medial surfaces can only be applied to objects with closed surfaces. Ob-

jects with incomplete surfaces and large holes, or objects that are not closed due to the articulation of parts, such as the opening a door of a car, do not define the notion of an interior and an exterior. Hence medial surface-based DAGs would not be appropriate. It is feasible to “patch” models with a few missing triangles, so that voxelization becomes possible. It might also be fruitful to explore Voronoi methods for computing medial surface-based DAGs that could in principle be applied directly to point clouds, provided that the sampling density is high enough Amenta et al. (2001) or to use the shock scaffold technique Leymarie (2003).

Second, the segmentation of medial surfaces considers only the digital (local) classification of (Malandain et al., 1993). In practice the determination of medial surface points can be sensitive to the local properties (as exemplified by the poorer results on the four-limbed animals). It is well known that certain regions of the medial locus are less stable than others, such as Blum’s ligatures Blum (1973). Thus, there is more work to be done both in the direction of developing robust techniques for segmenting 3D skeletons as well as in selecting its stable manifolds for building representations. A more robust segmentation algorithm would incorporate the orientation of the medial surfaces. In the continuum, junction points coincide with locations where the orientation of two adjacent medial surface manifolds changes significantly.

Third, the interpretation of a set of decomposed medial surfaces of a 3-D object into a directed acyclic graph is sensitive to the segmentation and saliency of the parts. The rule to construct the directed acyclic graph is to pick up the most salient part of the decomposed medial surfaces as the root. Take a lion as an example, the most salient part of the lion would be the body. However, if the body of the lion breaks into parts, the most salient part of the lion would probably be the legs. It would be worthwhile to study other ways of constructing the DAGs for objects within same category.

Fourth, the node similarity is based on a mean curvature histogram on the re-

constructed parts. This histogram is a translation and rotation invariant shape representation, but it is quite coarse. Indeed we expect that the performance of graph theoretic algorithms for comparing medial surface based representations will improve with more discriminating part similarity measures, and any one of a number suggested in the literature can be investigated.

Finally, the number of objects in the database can be increased for a more convincing demonstration of our technique. The majority of available models on websites are in a polygonal mesh format. Our computation of medial surfaces is applied for volumetric data and thus it takes some effort and labor to voxelize the mesh models.

Appendix A

Useful links for online 3D shape retrieval

1. <http://www.informatik.uni-leipzig.de/~vranic/ICME2002/>
2. <http://www.shapesearch.net/index.html>
3. <http://www.cs.uu.nl/centers/give/imaging/3Drecog/3Dmatching.html>
4. <http://shape.cs.princeton.edu/search.html>
5. <http://www.dbs.informatik.uni-muenchen.de/Forschung/Similarity/Demos/protein/>
6. <http://3d.csie.ntu.edu.tw/~dynamic/>
7. <http://3d-search.itl.gr/>
8. <http://www.cleopatra.nrc.ca/>
9. <http://www.nime.ac.jp/~motofumi/Ogden/>
10. <http://www.deepfx.com/meshnose>

Bibliography

- Alt, H. and L. J. Guibas: 1996, 'Discrete Geometric Shapes: Matching, Interpolation, and Approximation A Survey'. Technical Report B 96-11. [8]
- Amenta, N., S. Choi, and R. Kolluri: 2001, 'The Power Crust, Unions of Balls, and the Medial Axis Transform'. *Computational Geometry: Theory and Applications* 19(2), 127–153. [21, 53]
- Ankerst, M., G. Kastenmüller, H. Kriegel, and T. Seidl: 1999a, '3-D Shape Histograms for Similarity Search and Classification in Spatial Databases'. In: *Advances in Spatial Databases, 6th International Symposium*, Vol. 18. pp. 700–711. [3, 35, 41]
- Ankerst, M., G. Kastenmüller, H.-P. Kriegel, and T. Seidl: 1999b, '3D Shape Histograms for Similarity Search and Classification in Spatial Databases'. In: *SSD '99: Proceedings of the 6th International Symposium on Advances in Spatial Databases*. London, UK, pp. 207–226, Springer-Verlag. [9, 11]
- Arman, F. and J. Aggarwal: 1993, 'Model-Based Object Recognition in Dense-Range Images - a Review'. *ACM Computing Surveys* 25(1), 5–43. [8]
- Besl, P. J. and R. C. Jain: 1985, 'Three-dimensional object recognition'. *ACM Comput. Surv.* 17(1), 75–145. [8]
- Biederman, I.: 1987, 'Recognition-By-Components: A Theory of Human Image Understanding'. *Psychological Review* 94(2), 115–147. [3, 16, 41]
- Bierderman, I.: 1987, 'Recognition-by-Components: A Theory of Human Image Understanding'. *Psychological Review* 94(2), 115–147. [16]
- Binford, T.: 1987, 'Generalized Cylinder Representation'. In: *Encyclopedia of A. I.* John Wiley & Sons, pp. 321–323. [16]
- Binford, T. O.: 1971, 'Visual Perception by Computer'. In: *IEEE Conference on Systems and Control*. [3]
- Blum, H.: 1973, 'Biological Shape and Visual Science'. *Journal of Theoretical Biology* 38, 205–287. [4, 21, 28, 53]

- Bochuan, Z., P. Wei, Z. Yin, Y. Xiuzi, and Z. Sanyuan: 2004, 'A Survey on 3D Model Retrieval Techniques'. 16. [8]
- Borgefors, G.: 1984, 'Distance Transformations in Arbitrary Dimensions'. *Computer Vision, Graphics, and Image Processing* 27, 321–345. [3]
- Bouix, S.: 2003, 'Medial Surface'. Ph.D. thesis, McGill University. [21, 51]
- Cyr, C. M. and B. B. Kimia: 2004, 'A Similarity-Based Aspect-Graph Approach to 3D Object Recognition'. *Int. J. Comput. Vision* 57(1), 5–22. [18]
- Elad, M., A. Tal, and S. Ar: 2001, 'Content Based Retrieval of VRML Objects- An Iterative and Interactive Approach'. In: *6th Eurographics Workshop on Multimedia*. Manchester, UK, pp. 107–118. [3, 35, 41]
- Elad, M., A. Tal, and S. Ar: 2002, 'Content based retrieval of VRML objects: an iterative and interactive approach'. In: *Proceedings of the sixth Eurographics workshop on Multimedia 2001*. New York, NY, USA, pp. 107–118, Springer-Verlag New York, Inc. [9, 10]
- Funkhouser, T., P. Min, M. Kazhdan, J. Chen, A. Halderman, and D. Dobkin: 2003a, 'A Search Engine for 3D Models'. *ACM Transactions on Graphics* 22(1), 83–105. [2, 3, 41]
- Funkhouser, T., P. Min, M. Kazhdan, J. Chen, J. A. Halderman, D. Dobkin, and D. Jacobs: 2003b, 'A search engine for 3D models'. *ACM Trans. Graph.* 22(1), 83–105. [9, 12, 14]
- Giblin, P. J. and B. B. Kimia: 2004, 'A formal Classification of 3D Medial Axis Points and Their Local Geometry'. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 26(2), 238–251. [21, 23, 28]
- Hamza, A. B. and H. Krim: 2003, 'Geodesic Object Representation and Recognition'. In: *Proceedings of DGCI*, Vol. LNCS 2886. pp. 378–387. [4]
- Hilaga, M., Y. Shinagawa, T. Kohmura, and T. L. Kunii: 2001, 'Topology matching for fully automatic similarity estimation of 3D shapes'. In: *SIGGRAPH '01: Proceedings of the 28th annual conference on Computer graphics and interactive techniques*. New York, NY, USA, pp. 203–212, ACM Press. [4, 16]
- Horn, B.: 1984, 'Extended Gaussian Images'. 72, 1671–1686. [13]
- Igarashi, T., S. Matsuoka, and H. Tanaka: 1999, 'Teddy: a sketching interface for 3D freeform design'. In: *SIGGRAPH '99: Proceedings of the 26th annual conference on Computer graphics and interactive techniques*. New York, NY, USA, pp. 409–416, ACM Press/Addison-Wesley Publishing Co. [14]

- Kazhdan, M., B. Chazelle, D. Dobkin, T. Funkhouser, and S. Rusinkiewicz: 2003a, 'A Reflective Symmetry Descriptor for 3-D Models'. *Algorithmica* **38**(1), 201–225. [3, 35, 41]
- Kazhdan, M., B. Chazelle, D. Dobkin, T. Funkhouser, and S. Rusinkiewicz: 2003b, 'A Reflective Symmetry Descriptor for 3D Models'. *Algorithmica* **38**(1). [9, 10]
- Kazhdan, M., T. Funkhouser, and S. Rusinkiewicz: 2003c, 'Rotation Invariant Spherical Harmonic Representation of 3D Shape Descriptors'. In: *Symposium on Geometry Processing*. [ii, 6, 13, 43, 52]
- Körtgen, M., G.-J. Park, M. Novotni, and R. Klein: 2003, '3D Shape Matching with 3D Shape Contexts'. In: *The 7th Central European Seminar on Computer Graphics*. [12]
- Leymarie, F. F.: 2003, '3D Shape Representation via Shock Flows'. Ph.D. thesis, Division of Engineering, Brown University, Providence, RI, USA. [21, 53]
- Loncaric, S.: 1998, 'A survey of shape analysis techniques'. *Pattern Recognition* **31**(8), 983–1001. [8]
- Macrini, D.: 2003, 'Indexing and Matching for View-Based 3-D Object Recognition Using Shock Graphs'. Ph.D. thesis, University of Toronto. [4]
- Macrini, D., A. Shokoufandeh, S. J. Dickinson, K. Siddiqi, and S. W. Zucker: 2002, 'View-Based 3-D Object Recognition using Shock Graphs.'. In: *ICPR* (3). pp. 24–. [19]
- Malandain, G., G. Bertrand, and N. Ayache: 1993, 'Topological Segmentation of Discrete Surfaces'. *International Journal of Computer Vision* **10**(2), 183–197. [xiii, 21, 22, 24, 26, 53]
- Marr, D. and K. H. Nishihara: 1978, 'Representation and Recognition of the Spatial Organization of Three Dimensional Structure'. *Proceedings of the Royal Society of London B* **200**, 269–294. [3, 16]
- Novotni, M. and R. Klein: 2003, '3D zernike descriptors for content based shape retrieval'. In: *SM '03: Proceedings of the eighth ACM symposium on Solid modeling and applications*. New York, NY, USA, pp. 216–225, ACM Press. [14]
- Ohbuchi, R., T. Otagiri, M. Ibato, and T. Takei: 2002, 'Shape-Similarity Search of Three-Dimensional Models Using Parameterized Statistics'. In: *PG '02: Proceedings of the 10th Pacific Conference on Computer Graphics and Applications*. Washington, DC, USA, p. 265, IEEE Computer Society. [9, 11]

- Osada, R., T. Funkhouser, B. Chazelle, and D. Dobkin: 2001, 'Matching 3D Models with Shape Distributions'. In: *SMI '01: Proceedings of the International Conference on Shape Modeling & Applications*. Washington, DC, USA, p. 154, IEEE Computer Society. [6, 15, 52]
- Osada, R., T. Funkhouser, B. Chazelle, and D. Dobkin: 2002, 'Shape Distributions'. *ACM Transactions on Graphics* 21(4), 807–832. [ii, 2, 3, 35, 41, 43]
- Paquet, E. and M. Rioux: 1997, 'Nefertiti: a query by content software for three-dimensional models databases management'. In: *NRC '97: Proceedings of the International Conference on Recent Advances in 3-D Digital Imaging and Modeling*. Washington, DC, USA, p. 345, IEEE Computer Society. [9]
- Pentland, A.: 1986a, 'Perceptual Organization and the Representation of Natural Form'. *Artificial Intelligence* 28, 293–331. [3]
- Pentland, A.: 1986b, 'Perceptual Organization and the Representation of Natural Form'. *Artificial Intelligence* 28(2), 293–331. [16]
- Pope, A. R.: 1994, 'Model-Based Object Recognition - A Survey of Recent Research'. Technical Report TR-94-04. [8]
- Reyner, S. W.: 1977, 'An Analysis of a Good Algorithm for the Subtree Problem'. *SIAM J. Comput.* 6, 730–732. [37]
- Rosch, E.: 1978, 'Principles of Categorization'. In: *Cognition and Categorization*. L. Erlbaum Associates. [41]
- Shilane, P., P. Min, M. Kazhdan, and T. Funkhouser: 2004, 'The Princeton Shape Benchmark'. In: *Shape Modeling International*. Genova, Italy. [41]
- Shokoufandeh, A., S. J. Dickinson, C. Jonsson, L. Bretzner, and T. Lindeberg: 2002, 'The Representation and Matching of Qualitative Shape at Multiple Scales'. In: *Proceedings, ECCV*. Copenhagen, pp. 759–775. [37]
- Shokoufandeh, A., S. J. Dickinson, K. Siddiqi, and S. W. Zucker: 1999, 'Indexing Using a Spectral Encoding of Topological Structure'. In: *IEEE Conference on Computer Vision and Pattern Recognition*. Fort Collins, CO, pp. 491–497. [3]
- Shokoufandeh, A., D. Macrini, S. Dickinson, K. Siddiqi, and S. Zucker: 2005, 'Indexing Hierarchical Structures Using Graph Spectra'. *IEEE Transactions On Pattern Analysis and Machine Intelligence* 27(7). [4, 31, 52]
- Siddiqi, K., S. Bouix, A. Tannenbaum, and S. W. Zucker: 1999a, 'The Hamilton-Jacobi Skeleton'. In: *iccv*. Kerkyra, Greece, pp. 828–834. [4, 31, 51]

- Siddiqi, K., S. Bouix, A. Tannenbaum, and S. W. Zucker: 2002, 'Hamilton-Jacobi Skeletons'. *International Journal of Computer Vision* 48(3), 215–231. [2, 3, 21, 22]
- Siddiqi, K. and S. M. Pizer: 2005, *Medial Representations: Mathematics, Algorithms and Applications*. Kluwer. [21]
- Siddiqi, K., A. Shokoufandeh, S. J. Dickinson, and S. W. Zucker: 1998, 'Shock Graphs and Shape Matching'. In: *ICCV '98: Proceedings of the Sixth International Conference on Computer Vision*. Washington, DC, USA, p. 222, IEEE Computer Society. [19]
- Siddiqi, K., A. Shokoufandeh, S. J. Dickinson, and S. W. Zucker: 1999b, 'Shock Graphs and Shape Matching'. *International Journal of Computer Vision* 35(1), 13–32. [3, 22, 32, 33, 37]
- Sundar, H., D. Silver, N. Gagvani, and S. Dickinson: 2003, 'Skeleton Based Shape Matching and Retrieval'. In: *SMI '03: Proceedings of the Shape Modeling International 2003*. Washington, DC, USA, p. 130, IEEE Computer Society. [17]
- Tam, G., R. Lau, and C. Ngo: 2004, 'Deformable Geometry Model Matching by Topological and Geometric Signatures'. In: *Proceedings of the 2004 Computer Graphics International (CGI 2004)*. Greece, pp. 335–342, IEEE Computer Society Press. [18]
- Veltkamp, R. and M. Hagedoorn: 1999, 'State-of-the-art in shape matching'. Technical Report UU-CS-1999-27, Utrecht University, the Netherlands. [8]
- Vranic, D. and D. Saupe: 2001, '3-D Model Retrieval With Spherical Harmonics and Moments'. In: *Proceedings of the DAGM*. pp. 392–397. [3, 35, 41]
- Vranic, D. V., D. Saupe, and J. Richter: 2001, 'Tools for 3D-object retrieval: Karhunen-Loeve Transform and spherical harmonics'. In: *Proceedings of the IEEE 2001 Workshop Multimedia Signal Processing*. Cannes, France, pp. 293–298, Springer-Verlag. [9, 12]
- Zhang, J., K. Siddiqi, D. Macrini, A. Shokoufandeh, and S. Dickinson: 2005, 'Retrieving Articulated 3-D Models Using Medial Surfaces and their Graph Spectra'. In: *EMMCVPR '05: Proceedings of the Fifth International Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition*, Vol. to appear. Florida, USA, Springer. [6]