

PART DECOMPOSITION AND SHAPE DESCRIPTION

by

Hillel Shaul Rom

A Dissertation Presented to the
FACULTY OF THE GRADUATE SCHOOL
UNIVERSITY OF SOUTHERN CALIFORNIA

In Partial Fulfillment of the
Requirements for the Degree
DOCTOR OF PHILOSOPHY
(Computer Science)

December 1993

Copyright 1993 Hillel Shaul Rom

For Nava,

I accounted to your favor
The devotion of your youth,
Your love as a bride –
How you followed me in the wilderness,
In a land not sown.

Jeremiah 2.2

Acknowledgments

First and foremost, I thank my advisor, Dr. Gérard Medioni for his advice, friendship, and continuous support. Many of the ideas presented here, and probably the better ones, were actually his. I thank him also for our joint ride on the roller-coaster of business entrepreneurship.

I would like to thank Dr. Ram Nevatia for his most useful advice and help, and for serving on my guidance committee. The many long “corridor conversations” we had throughout my years at USC, have helped shape my understanding of the field of computer vision. Many of the underlying philosophical ideas expressed in this work, I have undoubtedly “inherited” from him.

I thank Dr. Irv Biederman for serving on my guidance committee. He has opened a window for me into the world of human vision. His theory of Recognition By Components is an important foundation of this work. His enthusiasm towards his own work, and towards mine, was always a source of encouragement.

I also thank Dr. Ari Requicha and Dr. Ken Goldberg for serving on my qualifying guidance committee. They both gave invaluable suggestions and were always there for advice.

I thank Delsa Tan and Dorothy Steele for their help with all my administrative concerns throughout these years. Special thanks go to Dr. Keith Price, my guru for all software problems, and to Andres Huertas for his friendship, technical support and 4th of July parties. I am also indebted to my fellow students in the IRIS group, past and present. I will always cherish the great experience of working and socializing with such a diverse set of people from, literally, all over the world. I especially thank Dr. Fridtjof Stein for his friendship and support, Mourad Zerroug for his help with differential geometry, Yang Chen for his help with range images and software, and Gideon Guy for his friendship, hours of carpooling, and joint ventures.

My deepest thanks go to my parents, to whom I owe so much, but have always returned so little.

Finally, I am short of words to express my thanks to my beloved wife, Nava. Throughout all these long, and often frustrating, years of being an eternal student’s wife, she has always been my best friend. This thesis is dedicated to her and to our son, Ofek.

Contents

Acknowledgments	iii
List Of Figures	vi
Abstract	viii
1 Introduction	1
2 Issues and Motivation	4
2.1 Model Based vs. Data Driven	5
2.2 Segmented vs. Non-segmented Descriptions	6
2.3 Dimensionality	7
2.4 Imperfect Data	9
2.5 The Figure-Ground Problem	10
2.6 Stability and Scale	10
2.7 Local Features vs. Global Features	11
2.8 Summary	12
3 Survey of Related Work	13
3.1 2-D Shapes	13
3.2 3-D Shapes	16
3.2.1 3-D Shape from 3-D Data	16
3.2.2 3-D Shape from 2-D Data	18
3.3 Other Classifications	20
4 Description of Two-Dimensional Shapes	22
4.1 Introduction	22
4.2 Overview of Approach	23
4.3 Issues and Implementation	25
4.3.1 B-Spline Approximation	25
4.3.2 Smooth Local Symmetries	28
4.3.3 Parts	30
4.3.4 Representing Global Relationships	34

4.3.5	Shape Decomposition	38
4.4	Experimental Results	39
4.5	Concluding Remarks	44
5	Description of Three-Dimensional Shapes	47
5.1	Introduction	47
5.1.1	Background	47
5.1.2	Issues and Approach	48
5.1.3	Outline of Chapter	49
5.2	Preliminaries	51
5.2.1	The Input	51
5.2.2	Differential Geometry	51
5.2.3	Basic Shapes and Parabolic Curves	54
5.3	Segmented Description of Compound Objects	56
5.3.1	From Image to Segmented Surface Description	57
5.3.2	From Surface to Volumetric Description	58
5.4	Axial Description of Parts	61
5.4.1	Straight Homogeneous Generalized Cylinders	62
5.4.2	Planar Right Constant Generalized Cylinders	68
5.5	Concluding Remarks	75
6	Conclusion	78
Appendix A		
	Recovering SLS of B-splines	80
Appendix B		
	Estimating Gaussian Curvature of Triangular Patches	82
Reference List		84

List Of Figures

2.1	General framework for recognition.	5
2.2	Non-segmented and segmented approaches to recognition.	6
2.3	Two possible paths for recognition.	8
2.4	Multiple levels of description.	9
2.5	Edges from real scenes are less than perfect.	10
2.6	Human shape at three different scales of resolution.	11
4.1	Outline of the data flow in the system.	24
4.2	The decomposition of a man shape.	25
4.3	Example of a B-spline approximation of a planar curve.	27
4.4	The geometry of a local symmetry.	28
4.5	A point may form more than one local symmetry.	29
4.6	All the SLS axes of a shape.	30
4.7	Segmenting at curvature minima results in 4 parts in both cases.	31
4.8	Curvature extrema.	32
4.9	A shape and the axes of its potential parts.	33
4.10	Special cases of multiple extrema: Termination, Bend, and T.	34
4.11	Removing parts: The shape and its guiding polygon.	35
4.12	Global relationships are important for natural descriptions.	35
4.13	Parallel symmetries, before and after filtering.	36
4.14	Parallel ribbons which are filtered out.	37
4.15	Relationship between global and local ribbons.	38
4.16	Decomposition of a man shape.	40
4.17	Two decompositions of a snake shape.	40
4.18	Decomposition and skeleton of a “bean” shape.	41
4.19	Decomposition and skeleton of a seven shape.	41
4.20	The decomposition of airplane shapes.	42
4.21	Example of the stability of the process.	42
4.22	The decomposition of airplane shapes obtained from real data.	43
5.1	Outline of our approach	50
5.2	Two types of 3-D input.	52
5.3	Examples of basic generic shapes.	55
5.4	Two types of parabolic curves: “part-like” and “hump-like.”	56

5.5	An anticlastic region (the glue between parts).	56
5.6	Teapot and segmented regions.	58
5.7	Shaded range images of teapot (left) and torus (right).	59
5.8	Sign of the Gaussian curvature of teapot and torus.	60
5.9	Connected regions of equal curvature sign.	60
5.10	Region graph of teapot (left) and torus (right).	60
5.11	The classification of the parabolic curves.	61
5.12	Final graph description of teapot (left) and torus (right).	61
5.13	Example of a Generalized Cylinder.	62
5.14	Examples of Straight Homogeneous Generalized Cylinders.	63
5.15	An SHGC.	64
5.16	Illustration of Property 5.4.1	64
5.17	Illustration of Property 5.4.3	65
5.18	Shaded range image of a teapot.	67
5.19	Cross sections, axis and shaded image of reconstructed teapot body.	67
5.20	Examples of Planar Right Constant Generalized Cylinders.	68
5.21	A PRCGC	69
5.22	A PRCGC and the recovered cross section.	73
5.23	A PRCGC and the recovered cross section.	73
5.24	Shaded range images of teapot (left) and torus (right).	74
5.25	Cross sections and shaded image of reconstructed torus.	74
5.26	Cross sections and shaded image of reconstructed teapot.	75
B.1	Geometrical interpretation of the Gaussian curvature.	82
B.2	Estimating Gaussian curvature on triangular patches.	83

Abstract

We address the problem of obtaining natural (intuitive) descriptions of shapes. Shape description is a major problem in machine perception, and is the basis for recognition. The requirements from good descriptions, which facilitate recognition, lead to descriptions which are segmented, that is, given in terms of parts and the parts arrangement. In this work we address both two dimensional and three dimensional shapes. In both cases we suggest methods for decomposing the shapes into their parts and for deriving natural descriptions of these parts.

Given a planar shape, we suggest a method for producing a segmented axial description of the shape, together with a hierarchical decomposition of the shape into its parts. The novelty of our approach lies in the combination of several competing approaches and tools, into a unified scheme and an efficient implementation producing natural descriptions. We use Smooth Local Symmetries for the axial description of parts, which are suggested by curvature sign changes. We also use parallel symmetries to provide information on global relationships within the shape. This information is used for parsing the shape into a hierarchy of parts. Currently we assume that the shape is a closed planar curve. Our method is computationally efficient, robust, stable, and we present results which show that it provides an intuitive shape description.

We present one of the first attempts to address the description of three dimensional compound objects, where the parts are connected smoothly. The input we consider is either complete 3-D data or range data from a single view. We suggest a *volumetric* graph representation of the object, where the nodes represent individual parts and the edges represent connectivity information. In an analogous way to our 2-D approach, we suggest the use of properties of the parabolic curves for performing the part decomposition. We currently consider objects with convex blobs and “tube like” parts. The graph description presents a structural description of the shape in terms of parts and their arrangement. We are also interested in the internal description of the parts. We study two well defined classes of shapes, namely Straight Homogeneous Generalized Cylinders, and Planar Right Constant GCs. We suggest the use of properties of the parabolic curves for recovering natural descriptions of these classes in terms of their cross sections and axes.

Chapter 1

Introduction

Shape description is one of the major problems in machine perception. Solving this problem (i.e. obtaining good shape descriptions), is the key to solving the recognition problem and some of the other fundamental problems in this field. Shape description is also important in related fields, such as robotics, and human vision understanding. Various methods for shape description have been suggested through the years of research in machine and human perception, but none provide a complete and natural solution to the problem. Furthermore, this problem seems to be one of the most challenging problems in the field, and is perhaps equivalent to *the vision problem* itself.

The term *shape* in this work refers to the outer form of the objects, or more specifically, to the geometry of the surfaces of an object in three dimensions, or to the geometry of the bounding contour of an “object” in two dimensions. We consider compact objects only, where the surfaces and contours are regular and, in general, “well behaved.” For example, we do not consider fractals, clouds of points, statistical objects, etc. By the words *shape description* we may refer to both the actual description or representation of the shape in terms of data structures and primitives, and to the *process* of describing the shape or reconstructing the description from the input shape.

Many researchers have discussed the requirements from a good shape description. Examples are Binford [11], Marr and Nishihara [62], Nevatia [64], Marr [61], Hoffman [39], Pentland [69], and Rao [73]. A good description should be rich (i.e. data preserving), to allow discrimination between similar shapes. It should be stable so that small changes in the shape result in small changes in the description. It also should be invariant to changes in the viewing conditions (e.g. lighting, rotation, scale, etc.). In addition, it should be capable of describing partially occluded parts. As noted by Rao [73], these requirements lead to descriptions which are segmented (given in terms of their parts) and hierarchical.

In addition to the above requirements, we are interested in producing shape descriptions which are *natural* (intuitive). Our ultimate goal is in achieving human like performance, that is, given a shape, we would like our algorithm to produce a description which is as close as possible to the description given by an “average” human to the same shape. Since the study of human perception also suggests that objects are naturally described in terms of their parts and part arrangement [10, 69], this supports our quest for part based descriptions.

The necessity for part segmentation and part based shape descriptions, in order to facilitate *general* object recognition, is the recurring theme throughout this work. It is one of our fundamental beliefs and is the foundation for the methods developed herein.

In this thesis, we first discuss some of the issues involved in designing a shape description mechanism, followed by a review of the related literature. We then suggest a method for producing an axial description of *planar* shapes, along with a hierarchical decomposition of the shapes into their parts. The novelty of our approach lies in the combination of several, often competing approaches and tools, into a *unified scheme* and an *implementation* producing natural descriptions. Combining different approaches as tools allows us to overcome the drawbacks of one approach by taking advantage of the merits of another. This is especially true since, as noted by Kimia *et al.* [49], previously suggested approaches have often taken extreme stands on the issues of shape description (e.g. contour vs. region and local vs. global). Our hierarchical approach results in descriptions which are rich and stable. The description is also invariant to rotation, translation, and uniform scaling. The description is segmented into parts and addresses the issue of scale. The issue of local vs. global information is also addressed. Our method is computationally efficient, robust, and we present results which show that it provides intuitive shape descriptions.

We then turn our attention to 3-D shapes. Our main interest is in producing *natural* descriptions of three dimensional compound shapes. Most of the work so far has concentrated on the recovery of descriptions of single components. When compound objects are addressed, the segmentation is assumed to be an easy step. We attempt to handle compound objects where the parts are connected smoothly. We suggest a graph representation of the object, where the nodes represent the individual natural parts and the edges represent connectivity information. We suggest the use of properties of the parabolic curves for performing the part decomposition. We currently consider shapes with convex blobs and “tube like” parts. The graph description presents a structural description of the shape in terms of parts and their arrangement. We are also interested in the internal description of the parts. This is especially important when performing tasks, such as grasping. We study two well defined classes of shapes. These are Straight Homogeneous Generalized

Cylinders (SHGCs [88]), and Planar Right Constant GCs (PRCGCs [98], planar axis and constant cross section). Based on known and novel mathematical properties, we suggest the use of the parabolic curves (curves of inflection of the Gaussian curvature) for recovering natural descriptions of these classes of objects in terms of their cross sections and axes.

Our methods require the computation of differential geometric properties of the surface of the object (especially the sign of the Gaussian curvature). We, therefore, require 3-D surface data which we currently obtain from range images or from CAD models.

The outline of this thesis is as follows: in the next chapter we present and discuss some of the major issues one has to consider when designing and evaluating mechanisms for shape description. The importance of part segmentation is also discussed here. In Chapter 3 we review some of the research done in this area, emphasizing work which is closely related to our own. In Chapter 4 we present our work on planar shape description. We start with an overview of our approach, we then present the details of our approach and implementation, followed by experimental results. In Chapter 5 we discuss three dimensional shapes. Following some preliminaries, we present our method for producing the segmented graph description of generic compound shapes. We then discuss the properties of SHGCs and PRCGCs and show how they are used for recovering their axes and cross sections. Finally, in Chapter 6, we present some concluding remarks and a brief discussion of some related research topics and possible extensions of our work.

Chapter 2

Issues and Motivation

The ultimate task of a complete vision system is to be able to recognize and classify general objects from intensity images of general scenes. By *general* we require that the system should not assume any limitations on the types of objects in the scene, on their location, nor on their relative positions (occlusions etc.). Although a substantial fraction of the research work in the field of computer vision is concentrated on object recognition, the state of the art systems are still far from achieving the above task. A summary of the recent work on object recognition is presented by Stein and Medioni [92]. As they note there, most object recognition systems to date either rely on exact models of the objects handled, or they make restrictive assumptions on the possible shapes of the visible surfaces. Current recognition systems in industrial environments are capable of recognizing objects from a limited set. Some occlusion can be handled, but the conditions are controlled and the problem is two dimensional in nature. Many of the object recognition systems under research assume the object to be known, and what they perform is actually pose estimation (i.e. estimating the exact position and orientation of the object). Many of the approaches have restricted the objects handled to be polygonal objects, or specific types of general cylinders. In addition, none of the systems has demonstrated the capability of handling a large database of possible objects. For comparison, humans are capable of recognizing thousands of objects with no difficulty from a single two dimensional image. Multiple objects in the scene and occlusion are overcome in most cases. Instances of object classes which were never seen before could also be recognized. Furthermore, new objects which were never seen before could also be classified correctly, and compared to other instances. As noted, it is our ultimate objective to reach recognition systems capable of similar performance. It is our belief that, in order to be able to perform recognition in this general sense, it is necessary to first derive meaningful segmented shape descriptions.

In the following we discuss several issues which have to be addressed when evaluating or designing a shape description mechanism. Keep in mind that our evaluation

Figure 2.1: General framework for recognition.

of the “goodness” of a shape description is in the context of how well it facilitates recognition. Some of these issues and others have also been discussed by several other researchers (e.g. [62, 64, 61, 39, 73]).

2.1 Model Based vs. Data Driven

Figure 2.1 presents the general framework for recognition: on the right side, predictions of 3-D and 2-D descriptions are made from objects models. These descriptions are matched with the equivalent descriptions, constructed from the scene. Unfortunately, this approach is only possible if the object in the scene is known a priori. Indeed, as mentioned above, most current systems assume either one or a very limited set of known objects. On the other hand, we believe that in order to be able to handle large databases of objects, it is necessary to use the descriptions derived from the scene to retrieve the recognized object, or a limited set of candidates. This approach (data driven) is depicted on the left side of Figure 2.1. Unfortunately, recovering descriptions, which facilitate efficient retrieval from large databases of objects, is an extremely difficult problem and is not feasible in current systems. Our work, presented in this thesis, is an attempt in this direction.

Figure 2.2: Non-segmented and segmented approaches to recognition.

2.2 Segmented vs. Non-segmented Descriptions

Figure 2.2 contrasts two approaches to address the problem of object recognition. The first approach, which is common to most current data driven object recognition systems, is based on non-segmented descriptions. By this approach, features are first extracted from the image. These features are used to index into the database of exact models of objects. The discriminatory power of the features used is very small. Therefore, once the database gets large, the indexing cannot produce a single object, but only a list of many candidates. This in turn requires a costly verification step. Due to this reason the current object recognition systems are model based, as discussed above, and are limited to very small databases (many of them even only handle one object, performing only the verification step also known as *pose estimation*). We believe that these are inherent problems to all recognition schemes based on non-segmented descriptions.

In the second approach, depicted in Figure 2.2, the object is first segmented into its parts. The parts and their arrangement are used for indexing into the database of models. It has been suggested (e.g. [10]) that even a small number of parts and their arrangement are sufficient to discriminate between thousands of objects, in a similar way that 26 letters and their arrangement are sufficient to recognize thousands of words. Due to this discrimination power, the indexing results in a unique object

(or perhaps a small number of candidates), and therefore, verification is easy. In addition, by this approach it is no longer necessary to have exact models in the database. It is possible to index into a database of generic models.

The main problem, however, with this second approach is that it requires the segmentation step which is a very difficult and unsolved problem. Since we believe that the segmentation into parts is the key to successful general recognition, the shape descriptions we suggest in this work are in terms of parts and their arrangement.

2.3 Dimensionality

The next issue we address is the dimensionality considered. There are three aspects to the problem: 1) The dimension of the objects handled, 2) The dimension of the input, and 3) The dimension of the description used for recognition:

- **Objects dimension** The objects handled could be either three dimensional or two dimensional. It is more realistic, of course, to consider three dimensional objects, for these are the objects we encounter in our environment. Unfortunately, the study of 3-D objects is much more difficult than that of 2-D shapes. One reason for that is the ambiguity which results from the projection of the 3-D object onto the 2-D image plane. Although easier than the 3-D case, the analysis of 2-D shape is still a very challenging and interesting problem. It is also very important since 2-D shapes are encountered in many cases, such as typed characters, or when the objects are viewed from a large distance. In addition, the ideas and methodology developed from analyzing the 2-D case could help in addressing the more general 3-D case. We address 2-D shapes in Chapter 4, and some classes of 3-D shapes in Chapter 5.
- **Input dimension** Assuming 3-D objects, the input could be either 3-D complete input, such as in a CAD system, or $2\frac{1}{2}$ -D range information (visible surfaces), or it could be 2-D input. The higher the dimension of the input, the easier the description process. However, higher dimension input is more restrictive and not as readily available as 2-D input. Another related issue is the number of available views of the object. It is usually possible for humans to describe objects based on a single view. The problem, however, could be made easier with the use of information from multiple views to assist the description process. In our work on 3-D objects we restrict ourselves to 3-D complete input or to range input from a single view. We do not attempt to derive 3-D descriptions from 2-D input.
- **Description dimension** In the process of data driven recognition there

Figure 2.3: Two possible paths for recognition (highlight of detail from Figure 2.1).

are two basic stages: it is first necessary to reconstruct a description of the object in the scene, and then to use the description to retrieve the recognized object. Assuming we consider the problem of recognizing 3-D objects from 2-D input (scene), there are two possible paths to be taken with trade-offs in the relative difficulty of the two stages, the description reconstruction vs. the retrieval. As depicted in Figure 2.3, which shows a highlight of the data driven recognition approach of Figure 2.1, one path is to reconstruct a 2-D description of the object in the scene, which is relatively easy, and then use the 2-D description for recognition, which is very difficult due to many ambiguities. The second path is to reconstruct a 3-D description of the object in the scene, which is very difficult, and then use the 3-D description for recognition, which is relatively easy. We emphasize that the term “easy” is used for comparison to the problems we consider extremely hard, but these “easy” problems are difficult themselves. In addition, in the context of the above discussion, the descriptions, both 2-D and 3-D, are assumed to be *good* descriptions as discussed earlier. For example, there are many levels of possible 2-D descriptions, starting from edgels, through connected contours, through symmetries, and so on (see Figure 2.4). The higher the level of the

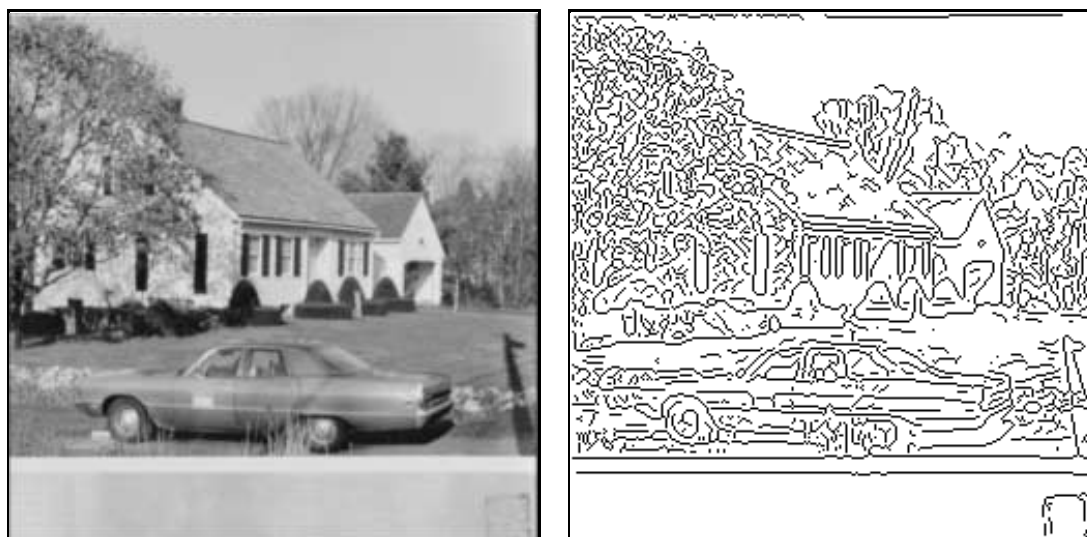
Figure 2.4: Multiple levels of description.

description the more it is useful for reconstructing 3-D descriptions and for recognition, and therefore, as discussed above, we consider it better.

Another consideration which follows is what we may call the dimensionality of the description. If the description is three dimensional, it could be either volumetric or surface based ($2\frac{1}{2}$ -D). If it is two dimensional, it could be either region based or contour based. There are trade-offs between the merits and disadvantages of each. A hybrid approach should be considered favorably.

2.4 Imperfect Data

Any system which tries to interpret images in a realistic environment has to be capable of handling imperfect data due to noise, quantization errors and other effects which are inherent in the imaging process. In particular, we cannot expect to get perfect contours from the current edge detectors. We should expect many missing edge segments, many fragmented edge segments, and many edges which do not exist in the scene. Figure 2.5 is an example of a real outdoor scene and the edges obtained using a state of the art edge detector (Canny [21]). The problem of image segmentation has not been sufficiently solved yet, and it could not be solved by low level methods alone. One of the attempts to obtain descriptions from noisy imperfect data has been performed by Rao and Nevatia [73, 75]. They suggest that the description is needed for the scene segmentation, as well as the fact that the segmentation is needed for the description. The issue of imperfect data has to be considered in both the high levels and the low levels of the description process. Any applicable description scheme has to be stable and robust.



(a) Outdoor scene.

(b) Results from the Canny edge detector.

Figure 2.5: Edges from real scenes are less than perfect.

2.5 The Figure-Ground Problem

One of the major problems in obtaining shape descriptions is the segmentation of the image into the object vs. the background. Due to texture, surface markings, noise, and other factors, we cannot always determine the exact closed contours of the object. Allowing multiple objects in the scene aggravates this problem even further. Given open contours, it is impossible to decide locally on which side of the contour is the object. Some inference process, such as *perceptual grouping* [59], is necessary for obtaining the shape or hypotheses for the shape, prior to the description process. Although we recognize that this issue is crucial for any real application, we do not attempt to solve it in our current work. The assumptions we make regarding the input, for the 2-D and 3-D cases, are specified and discussed in the respective chapters.

2.6 Stability and Scale

An important issue, which is somewhat connected to the problem of noise, is the stability of the description in terms of scale. Our requirement from a description is for it to be stable. This implies that slight changes in the shape, maybe due to noise, will not have a major effect on the description. However, it should also be rich

Figure 2.6: Human shape at three different scales of resolution.

so that even small differences between objects could be detected in the description. This raises the issue of scale at which the shape is observed. Any object could have different shapes depending on the scale at which it is interpreted. For example, as demonstrated in Figure 2.6, a shape of a human could be perceived from a long distance (low resolution) as just a long and narrow rectangle. From a closer look the arms, legs and head are seen. Yet, from even a closer look the hand and fingers are seen, and so on. To facilitate recognition, all these scale levels should be represented in the description. To be stable, and to be capable of handling partial occlusions, it is necessary for the description to preserve the relative significance of the shape descriptions at the different scale levels. This leads us to the natural conclusion that hierarchical descriptions are desired for stable and general object recognition systems.

2.7 Local Features vs. Global Features

When using non-segmented descriptions (see Section 2.2), an issue which should be considered is whether to base the description on local or on global features. Global features, such as area, perimeter squared divided by area, eccentricity, moments, and Fourier descriptors, are stable and can convey the general shape of the object, but

they fail in the case of occlusion. On the other hand, local features, such as curvature and local symmetries, are sensitive to noise and, by definition, they cannot capture the global shape. For a general shape description it is not sufficient to have local features only. The global shape should also be captured. This could be achieved by the use of global features, or by an hierarchical description process, as discussed in later chapters.

2.8 Summary

To summarize, we outline briefly how our work, presented in this thesis, relates to the above issues.

In both the 2-D and 3-D cases, we suggest data driven approaches to obtain segmented shape descriptions. The segmentation is based on the natural parts of the shape. For 2-D shapes we suggest an hierarchical segmented description. We assume perfect input of a closed planar curve. The description is based on local and global features. For 3-D shapes we suggest a segmented description in terms of parts and their arrangement. We suggest a part description for some classes of shapes. We consider 3-D complete data and $2\frac{1}{2}$ -D input from range data from a single view. We show results on real range images. We assume, however, that the figure-ground is given.

Chapter 3

Survey of Related Work

In this chapter we discuss some of the previous work done on shape description. We concentrate on work which we find to be closely related to our own. In Section 3.1 we discuss methods suggested for the description of 2-D shapes, while in Section 3.2 we discuss the methods suggested for 3-D shapes. In Section 3.3 we briefly discuss an alternative classification of the previous work according to the assumptions made on the model types. As noted, our requirements for a good shape description call for segmented descriptions, which describe objects in terms of their parts and the relations between their parts. Examples of non-segmented descriptions, to name but a few, are: Fourier descriptors [101], moments [30], Extended Gaussian Images [42, 46], Generalized Hough Transform [3, 4], and cell decompositions (e.g. oct-tree [47, 86]). Since these methods do not meet the above requirements, we do not consider them further here. Several description mechanisms were developed for solid modeling in the areas of computer graphics and CAD/CAM. Requicha [76] presents a survey and comparison of these methods and systems. Some of these descriptions are segmented and rich (e.g. Constructive Solid Geometry (CSG) [99]). However, they fall short on the reconstruction aspect, since it is very hard to derive these descriptions from images. For this reason, these descriptions are not used in the computer vision context and we do not discuss them here.

3.1 2-D Shapes

In this section we discuss research on the description of planar shapes. As noted above, we concentrate on work which is closely related to our own.

Many authors have suggested a representation scheme based on a skeleton or an axis. One approach is to look for a single (usually straight) axis based on a global function of all the contour points (e.g. axis of inertia). A recent example of this

approach is a scheme for finding a curved axis of inertia suggested by Subirana-Vilanova [94]. Another approach is to compute a curved axis, in which every point is determined by a small local area of the contour. A subset of this approach are the methods of describing “ribbon like” shapes with an axis around which the shape is locally symmetric. Examples for this approach are Blum’s *Symmetry Axis Transform (SAT)* [14, 15], Brooks’s *Generalized Ribbons* [19], which are a two-dimensional version of Binford’s *Generalized Cylinders* [11], and Brady’s *Smooth Local Symmetries (SLS)* [18]. These methods capture contour and region information. An in depth comparison of these three methods was performed by Rosenfeld [82], and extended by Ponce [70]. Rosenfeld [82], has compared the advantages and disadvantages of these methods from the standpoints of both generation (of the shape from the axis) and recovery (of the axis from the shape). SAT and SLS are both description (recovery) oriented, developed for recovering the axial representation given the shape. In general, these methods are local in nature, they are very sensitive to noise, and they are expensive to compute, even when using analytical approximations to the shape. These methods do not address the issue of segmentation of shape into parts and the issue of scale.

Several methods have been suggested for recovering SLS, SAT or Brooks ribbons. One approach is to compute the axis analytically. Unfortunately, analytical solutions exist only for limited cases, such as for straight lines or circular arcs. An approach, known as the method of projections, for computing Brooks ribbons was suggested in [65]. This method consists of “projecting” all contour points into buckets defined by all possible discrete orientations of local ribbon axes. The local symmetry condition is then verified for points in each bucket. Finally the local ribbons are grouped together into ribbons. The complexity of this method is $O(kN)$ where N is the number of points on the curve and k is the number of orientations. Another approach for computing symmetries is based on *local signatures*, which are properties that can be verified locally on the two points examined for symmetry. By this approach, every point of the curve is compared with every other point and tested for a local symmetry. The points for which the symmetry property holds are grouped into ribbons. The complexity of these methods is $O(N^2)$, where N is the number of points on the curve. By definition, local signatures exist for SAT and SLS (see Section 4.3.2). Ponce [70] has discovered a local signature for skew symmetries, which are a subset of Brooks ribbons with a straight axis. Other approaches for recovering skeletons of given shapes include methods based on thinning or morphology, methods based on the Voronoi diagram, and moments based methods.

Brady and Asada [18] use SLS for shape description. In order to compute the SLS axes analytically, they approximate the shape with circular arcs. This leads to an $O(n^2)$ algorithm for recovering the axes, where n is the number of

circular arcs. Since the number of arcs is far less than the number of points, N , this algorithm compares favorably to the $O(N^2)$ complexity of the direct approach, discussed above. They compute all of the possible SLS axes on the whole shape. As shown in Section 4.3.3, this results in many axes which are not relevant for the description of the shape. In addition, many significant axes may not exist in the initial shape. Later, **Connell and Brady** [26, 27] have suggested constraints for extracting axes which seem to be salient for the description. They use these axes to construct a semantic network for describing the shapes, which is then used for learning shape models.

In a pioneering work, **Hoffman and Richards** [39, 77] suggest that shapes are naturally segmented into parts at negative curvature minima. This was based on psychophysical observations and on the *transversality principle* [37], by which, in general, concavities arise when two convex parts are joined. They then suggest a contour based segmented representation. They segment the contour into segments, called *codons*, at negative curvature minima. Curvature maxima and inflections are used for internal part description. A shape is represented by the sequence of the codon types, forming its contour. This scheme, being contour based, does not incorporate any region information. It also does not account for any global relationships between distant sections of the curve. In Section 4.3.4 we show an example where segmenting at curvature minima, ignoring global relationships, does not produce intuitive results.

Leyton [55, 56] has suggested a process-grammar for shape. He argues that a shape can be understood as the outcome of processes (deformations) that formed it. This approach assumes that shapes (he studies natural shapes like clouds) are basically circles which have been deformed by a history of protrusions and indentations. Based on an important duality between curvature extrema and symmetry, he uses symmetry axes to infer the process history.

Kimia et al. [49] discuss the multidimensional nature of shape. They present examples where approaches, which take an extreme stand along a certain dimension (e.g. region vs. boundary or local vs. global), fail to give intuitive results (some of the examples we present in Chapter 4 were inspired by this work). They propose a theory of shape based on the tension between reaction and diffusion in the context of a conservation law. Their method produces a hierarchical decomposition of shape into parts and protrusions, defined as singularities in the continuous reaction-diffusion space. The parts descriptions are given by the dynamics of how each part shrinks to a circle. In a more recent work, **Siddiqi and Kimia** [89] suggest the decomposition of shapes based on two kinds of parts: *neck-based* parts, which are local minima of the distance between two points on the contour, and *limb-based* parts, which arise from a pair of negative curvature extrema with co-circular tangents.

3.2 3-D Shapes

The methods described in the previous section are concerned with the description and reconstruction of 2-D shapes. The description and reconstruction of 3-D shapes have also been studied by many researchers. In the following we discuss some of the previously suggested approaches, classified by the dimension of the input considered: 3-D shape from 3-D input ($2\frac{1}{2}$ -D range images) vs. 3-D shape from 2-D input (grey level images).

3.2.1 3-D Shape from 3-D Data

In an early work **Agin and Binford** [1] suggest a system for recovering descriptions of generalized cylinders with circular cross sections from dense range data. Based on an initial estimate for the axis, they fit circular cross sections to the data. A new axis is then generated from the centers of the cross sections and the process is repeated. The algorithm produces unstructured descriptions and is restricted to circular cross sections.

Nevatia and Binford [65] present a system for recovering generalized cylinders descriptions from boundary contours of objects in range data. Since they only use the 3-D range data for recovering the shape's contour, their approach could also work using 2-D closed curves as input. In their system, local cylinders with straight axes are first extracted, using the method of projections, discussed above. These are then extended while allowing the smooth curving of the axis. The objects are described with a relational graph of the cylinders and their connectivity. This structured graph description is then used for recognition.

Fan et al. [31] suggest a system for object recognition using surface descriptions. By their approach, 3-D surfaces are segmented at discontinuities, measured by zero-crossings and extremal values of surface curvature, into simpler patches. Attributed graphs are generated from these surface patches to represent the 3-D object. Multi-view models are then generated from several views. The recognition is based on graph matching with the graphs recovered from the scene.

Parvin and Medioni [66] suggest a dynamic system model for solving ambiguities in computer vision problems. They present a system of differential equations, which exploits geometric constraints for extracting boundaries of objects in range images. Feedback is introduced to ensure contour closure. Each view of an object is represented by an attributed graph with nodes as surfaces and links as relations between the surfaces. A dynamic system is also used for matching between views and reconstructing a volumetric representation of the object. Results are shown on

objects from real range images. The method relies on the existence of discontinuity edges between the different surfaces.

Flynn and Jain [34] present a system, which recognizes instances of predefined models in range images. The models are represented with relational surface graphs extracted automatically from CAD models. These graphs are matched with graphs constructed from the image based on surface segmentation and classification modules. Geometric constraints are used to prune the search space. Results are shown on real images, however, the scenes contained no more than two objects.

Several researchers have suggested superquadrics as a tool for shape description [68, 69, 90, 17, 32]. Superquadrics, which are actually a class of generalized cylinders, are a family of parametric shapes, which can have different shapes depending on the settings of five parameters. Three of the parameters control the scaling along the three axes, while the other two determine the shape of the cross section. Due to their versatility they are very useful as modeling primitives. Most of the research in this area has concentrated on the problem of finding the best superquadric to fit the given data. **Pentland** [69], discusses the importance of segmented descriptions in terms of parts. He suggests the use of basic superquadrics, which undergo deformations such as bending and tapering, for modeling the different parts. The issue of segmentation of the shape into the parts is ignored in this work, as well as in most other work in this area, save from the work of **Ferrie et al.** [32], who present some results of recovering volumetric descriptions of articulated objects. Since the superquadrics are controlled by only five parameters, their expressive power is somewhat limited (for example they can only bend in one direction so they cannot model a snake). This limitation enhances even further the need for segmentation before the fitting of the superquadric.

Recently, **Sato et al.** [87] suggested a method for recovering an axial representation of 3-D shapes using what they call *Smoothed Local GCs (SLGCs)*. These SLGCs can be recovered locally (as in the case of the 2-D SLS). Global optimization is used to group the local axes into global smooth axes. Due to the three degrees of freedom of the SLGC the complexity of the process is very high. The search space is somewhat constrained by assuming mirror symmetric objects.

Finally, a recent approach, which could be considered as a combination of the approaches which use 2-D data and those using 3-D data, is shape description and reconstruction from a stereo pair of 2-D images. We do not elaborate on these systems here. Examples are: Rao and Nevatia [74], Lim and Binford [57], and recently Chung and Nevatia [24].

3.2.2 3-D Shape from 2-D Data

ACRONYM developed by **Brooks** [19, 20] is a domain independent model based vision system. Objects are modeled as hierarchies of generalized cylinders and represented by a graph of their geometric and algebraic constraints. Two dimensional features of the object are predicted as ribbons and ellipses, based on reasoning using about 200 production rules. These predictions are represented in a *prediction* graph. From the image, the ribbons and ellipses are extracted with use of information from the prediction graph. These features are represented in the *picture* graph. Finally, recognition is based on sub-graph isomorphism between the prediction graph and the picture graph. *ACRONYM* has been applied successfully to aerial images of airports, and is capable of extracting instances of specific airplanes. Its main limitations are due to it being model based and not general. In addition, it cannot handle curved ribbons.

Terzopoulos et al. [95] suggest deformable models of 3-D shapes based on 2-D contours. These deformable models are active, changing shape according to variational principles. They consider generalized cylinder type models, by coupling deformable tubes with deformable spines. Instances of the models are reconstructed by combining symmetry based internal constraints of the model with external constraints derived from the image. The external constraints deform the models so that their projections are consistent with the 2-D silhouettes in the image. This method captures the symmetrical character of generalized cylinders, but it is not limited to shapes possessing exact symmetries. This method is more of a reconstruction tool than a shape description scheme. It also relies on an initial guess of the model which in many cases requires an interactive connection with a human user.

Rao and Nevatia [73, 75] challenge the shape description methods which assume the *figure-ground* problem to be solved prior to description. They argue that when working with intensity images, it is not realistic to assume that the object boundaries are given. They suggest that scene segmentation should take place through the process of shape description. They use 2-D ribbons for basic part description, under the assumption that the 2-D ribbons are projections of 3-D generalized cylinders. However, although 3-D generalized cylinders project to 2-D ribbons, the converse is not necessarily true. Therefore, 2-D ribbons in the image are not always projections of parts of the 3-D shape. They present convincing results of recovering shape descriptions from noisy intensity images with fragmented boundaries and surface markings. [In view of the examples chosen, as well as, the choice of 2-D ribbons as the description tool, this work could be classified in the 2-D shape class.]

Ponce et al. [72] consider objects which are restricted generalized cylinders. They provide a mathematical study of the class of Straight Homogeneous Generalized Cylinders (SHGCs) and characterize their Gaussian curvature and occluding boundaries. They go on and prove several important invariant properties of the contours of SHGCs. These properties are used for recovering SHGCs from the image contours. The theoretic results in this paper could prove to be important in the future research on generalized cylinders based shape descriptions (we use some of these properties in our work on 3-D shapes). In a different work, Kriegman and Ponce [53] suggest a method for recognizing and positioning parameterized three dimensional curved objects from their image contours. This approach assumes the models do be known exactly and is therefore not appropriate for general object recognition.

Several researchers have addressed the problem of 3-D interpretation of 2-D line drawings. Most of the early work has concentrated on line drawings of polyhedra [44, 25, 60, 48]. Some attempts to handle curved surfaces have also been made [5, 93, 100, 41]. **Ulupinar and Nevatia** [96–98] have proposed methods for reconstructing the surfaces of several classes of objects, such as objects with zero Gaussian curvature surfaces, and some classes of generalized cylinders (SHGC and CGC), from their 2-D contours. From the study of the geometric and differential properties of various types of surfaces they derive constraints on the object’s surfaces, imposed by the 2-D images of its contours. They show that three types of symmetry, *parallel* symmetry, *line-convergent* symmetry, and *skew* symmetry, formed by the object’s contours, convey important shape information. By recovering the symmetries in the image, and based on some basic assumptions, they can produce dense shape descriptions of the objects from the above classes. Their implementation currently assumes that the contours and symmetries are given.

Zerroug and Nevatia [103, 104] address the problem of recovering the 3-D shape from the 2-D image contours in real images. They use mathematical invariant and quasi-invariant (see [12]) properties of the contours of several classes of generalized cylinders to hypothesize the contours of the GCs, while overcoming noise, surface markings, and partial occlusion. The contours are then used to recover 3-D object centered descriptions. They present promising results on complex real images. The issues of compound shapes, however, are not yet addressed.

Biederman [9, 10], suggests a theory of human image understanding, called *Recognition By Components (RBC)*. By RBC an object is represented as an arrangement of simple generalized cylinder volumes. The fundamental assumption of the theory is that a subset of these generalized cylinders, called *geons*, can be derived from invariant properties of the edges in a 2-D image. This is supported by the important observation, mentioned earlier, that a line drawing of an object is classified

by humans as fast as its full colored image. It is suggested that objects could be recognized based on the arrangement of three of their geons. Parsing of the object into parts is done at matched deep concavities, and at changes of nonaccidental properties of the edges (e.g. parallelism and colinearity). A set of 36 geons is derived from contrasts in the nonaccidental properties of four attributes of generalized cylinders. It is argued that this set suffices for describing most objects. They present many psychological experiments which support RBC. **Hummel and Biederman** [45] present a neural network model of the process of generating a structural description, in terms of geons and relations between geons, from a line drawing (the surface and depth discontinuities in the image) of an object. This description is then used for recognition. Their model uses synchronized oscillatory activity to temporarily bind together independent units representing an object's parts attributes, when these attributes occur in conjunction in the input. The seven layer network is used to derive the following description hierarchy: From image edges, to vertices, axes, and blobs, to geon attributes, to relations of geons, to geons feature assemblies, and finally to objects. The representation is invariant to scale and translation, and the performance of the model conforms with empirical findings. However, from a computer vision perspective, the most challenging task is to recover the vertices, axes, and blobs from the image edges. These are assumed by the authors to be given.

Bergevin and Levine [7] have implemented a system, called PARVO, based on geons. Given line drawings the system recovers parts and classifies them into one of the set of geons, based on the properties of the geons discussed above. The system relies on perfect line drawings and on the fact that matched concavities could always be found.

Dickinson et al. [28] address the problem of recognizing compound objects which are composed of a limited set of primitives (a subset of Biederman's geons). These primitives can be recovered based on their aspect graph. The assumption in this work, as in most cases where compound objects are handled (e.g. [7, 45]), is that the primitives are not connected smoothly, that is, there is a sharp discontinuity edge separating the primitives. This assumption therefore translates into a problem of volumetric descriptions, given that the components are already segmented.

3.3 Other Classifications

In addition to the above classification of descriptions in terms of the dimension of the objects, there are many other possible classifications. We briefly mention two possible classifications here with respect to the methods discussed above.

One possible classification is by the classes of models assumed. The three classes we consider are: exact models, parametric models, and generalized cylinder based models. Most object recognition systems to date assume that an exact model of the object exists in the systems database. Features from the model and the scene are extracted and compared to produce an optimal match. For an example of such a system and a survey of others, refer to Stein and Medioni [92]. The approaches for recognizing polygonal shapes [44, 25, 60, 48] could also be classified under this category. One of the problems with these approaches, which is often ignored, is the difficulty in obtaining the exact model in the first place. Given an existing object it is not a trivial task to recover its CAD model or some other rich enough model.

Another class of systems is based on parameterized models. These models are based on some basic modeling primitives, which are deformed to fit the input, according to some function. The representation is based on the list, which is usually small, of parameters defining the deformed primitive. Examples for such approaches and primitives used are: Snakes [95], superquadrics [69], and superellipsoids [17].

The third class of systems are generalized cylinders based descriptions. Note that although generalized cylinders are also parameterized models, their popularity warrants an independent category. These approaches model the objects as composed of generalized cylinder components. The set of allowable generalized cylinders is usually limited to specific subclasses (e.g. straight axis, constant shape cross sections, etc.). Examples for this class are: [1, 65, 20, 73, 10, 9, 72, 97, 98].

Another possible classification of description schemes is based on whether the input considered is from real data or assumed perfect. The above mentioned research works are classified as follows: [77, 55, 49, 98, 10, 45, 7] consider perfect input while [18, 1, 65, 31, 66, 68, 69, 90, 32, 87, 20, 95, 75, 103, 28] use real data as input, either range images or intensity images. Note however that although the latter use real data as input, only Rao and Nevatia [75] and Zerroug and Nevatia [103] do not assume that the figure-ground is given.

Chapter 4

Description of Two-Dimensional Shapes

4.1 Introduction

As discussed earlier, we are interested in producing *natural* descriptions of shapes. Our ultimate goal is in achieving human like performance. In this chapter we address the description of planar shapes. We suggest a method for producing an axial description of a shape, along with a hierarchical decomposition of the shape into its parts. The novelty of our approach lies in the combination of several, often competing approaches and tools, into a *unified scheme* and an *efficient implementation* producing natural descriptions. Combining different approaches as tools allows us to overcome the drawbacks of one approach by taking advantage of the merits of another. This is especially true since, as noted by Kimia *et al.* [49], previously suggested approaches have often taken extreme stands on the issues of shape description (e.g. contour vs. region and local vs. global). We use Smooth Local Symmetries (SLS) for the axial description of parts. We also use *parallel symmetries*, introduced by Ulupinar and Nevatia [96], to provide information on global relationships within the shape. This information is used to parse the shape into a hierarchy of parts. Currently we assume that the shape is a closed simple planar curve which we approximate using quadratic B-splines. We also make the basic assumption that the input shapes can be interpreted and parsed without any 3-D inference. Our approach is region and contour based, and addresses the issues of local and global information, the issue of scale and the notion of part. Our method is computationally efficient, robust (all experiments were performed using the same parameter settings), stable, and we present results which show that it provides intuitive shape descriptions.

This chapter is organized as follows: In the next section we outline the basic ideas of our approach. In Section 4.3 we present the details of our approach and implementation. We discuss briefly two of the tools we use in our application, namely the B-spline approximation and Smooth Local Symmetries. We introduce our notion

of what is a part, the local element of a shape. We also discuss the importance of global relationships for shape description and introduce parallel symmetries as a tool for capturing these relationships. We end this section by presenting our method for decomposing the shape into a hierarchy of parts using the local and global information. Examples of the results we obtained are presented in Section 4.4 followed by some concluding remarks in Section 4.5. Parts of the work presented here were also published in [80, 81, 85].

4.2 Overview of Approach

We suggest a hierarchical approach for shape description, combining local and global information. As noted, we try to combine competing and complementary tools for shape description, gaining from the benefits of each. We produce a decomposition of the shape into parts together with an axial description of these parts. We outline our method here, the details of our implementation are presented in the next section.

A strategy often used for obtaining axial representations, using local features such as symmetries, is to first generate all the possible axes on the whole shape, and then to select the appropriate axes according to some criteria (see for example [26, 27, 73]). The computation of the symmetries and the selection process are difficult and computationally expensive. Furthermore, the selection process is based on the assumption that *all* the important axes can be found by the symmetry computing stage. However, this assumption is often wrong since some of the significant axes may not exist in the initial shape, until some of the parts are removed. We believe that it is necessary to use a hierarchical strategy instead, in which, at each step, local and well defined parts are described and removed. Once these parts are removed, the next level parts can be analyzed. This process is efficient and produces a decomposition of the shape into its intuitive parts with a stable axial description of these parts. One remaining problem of this approach is that it ignores global relationships between different parts of the shape. We use *parallel symmetries* [96], to detect such global relationships. Once conflicts between the local description and the global relationships are detected, we create a branch in our decomposition process, thus allowing multiple plausible interpretations, which turn out to be rare.

Figure 4.1 presents an outline of the data flow of our approach. Initially, we perform some preprocessing on the image including edge detection and linking. We then approximate the input shape with approximating B-splines. We currently assume that the shape is a closed planar curve. We ignore many of the practical difficulties of this stage, mainly the figure-ground problem. Following this step, we

Figure 4.1: Outline of the data flow in the system.

have an analytical representation of the contour. This reduces the computational complexity, and increases the stability of our description.

Since we are interested in an axial representation, we find *Smooth Local Symmetries (SLS)* [18] to be a powerful descriptive tool. However, as we show later, applied globally on the entire shape, many superfluous axes, not relevant for the description, are also produced. Therefore, we first segment the contour at curvature sign changes into initial local parts. We represent these parts only using SLS. A theorem proved by Leyton [54], relates the existence and uniqueness of an SLS axis describing each part to the curvature extrema of its contour. This enables us to efficiently generate stable and noise free descriptions for these local parts. In order to obtain natural

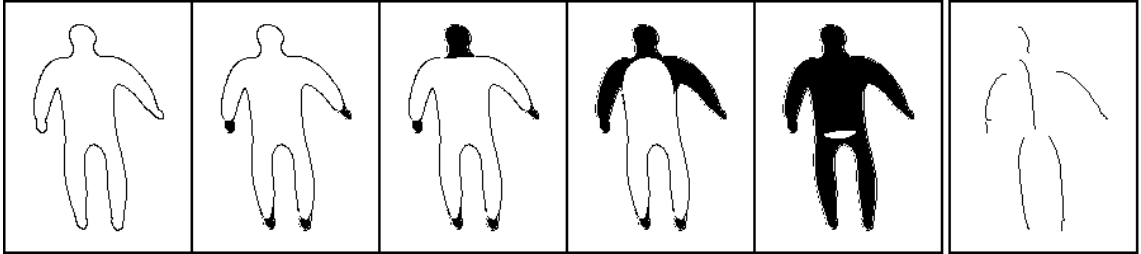


Figure 4.2: The decomposition of a man shape. The original shape is on the left. At every step the parts already explained are shaded. The skeleton description is also presented.

results in cases where there are global relationships between different sections of the shape, we compute the parallel symmetries in addition to the local description. Computing the parallel symmetries, even on the entire shape, is very efficient.

Now that we have the local parts and the information on the global relationships, we decompose the shape into parts. The decomposition is done hierarchically, first removing the small and well defined parts, and then analyzing the remaining shape. We generate the possible parsings of the shape into parts, starting from the initial shape and producing the axial representation determined by every parsing.

In order to clarify what we mean by an hierarchical decomposition, and in order to demonstrate the capability of our method, we present an example here (we present it in more detail in Section 4.4). Figure 4.2 presents the decomposition of a *man* shape (inspired by [49]). The original shape is on the left. At every step the parts which were previously removed (explained) are shaded. The current shape not yet explained is left white. In this case, our method finds a unique decomposition of the shape into a hierarchy of parts. The skeleton description obtained is also presented.

4.3 Issues and Implementation

In this section we describe the details of our approach and implementation. We present the tools we use and the motivation behind our choice of these tools. We also present how these tools are combined into the shape decomposition process producing the hierarchical shape description.

4.3.1 B-Spline Approximation

Following Saint-Marc *et al.* [84, 85], we use approximating B-splines to represent planar curves. As noted there, this representation is rich, compact, stable and local.

We review here some of the basics of the B-spline approximation, emphasizing the characteristics which are important for our application. The reader is referred to [6] or any of the other textbooks on the subject. (Note that in this section we discuss the representation of the curve only, not the representation or description of the *shape*, which is the subject of this work).

A B-spline is a piecewise polynomial which is expressed as a linear combination of polynomial basis functions. The coefficients, or *control points*, are the vertices of the B-spline *guiding polygon*. In the planar case, a B-spline $Q(u) = (X(u), Y(u))$ with m vertices is defined as follows:

$$Q(u) = \sum_{j=0}^{m-1} V_j B_j(u) = \sum_{j=0}^{m-1} (X_j B_j(u), Y_j B_j(u))$$

where $V_j = (X_j, Y_j)$ are the vertices of the guiding polygon and $B_j(u)$ are the basis functions. In our application, we use quadratic B-splines, every curve segment is a quadratic polynomial depending on three adjacent control points.

The process of obtaining the B-spline approximation of the input planar curves is as follows (see also [84, 85]). We first produce an edge map from the image using an edge detector such as Canny's [21], followed by linking of the elementary edgels. Following linking, corners, defined as significant tangent discontinuities, are found using adaptive smoothing [83]. Finally, each curve segment, either closed or between two corners, is approximated by a B-spline. This is done by performing a least-squares fit of the curve segment by a B-spline curve, minimizing the distance between the discrete data points of the curve and its approximation. Let C be the approximated curve, an ordered set of p points $P_i = (x_i, y_i)$. The approach consists of minimizing the distance

$$\begin{aligned} R &= \sum_{i=0}^{p-1} \|Q(u_i) - P_i\|^2 \\ &= \sum_{i=0}^{p-1} \left(\sum_{j=0}^{m-1} X_j B_j(u_i) - x_i \right)^2 + \left(\sum_{j=0}^{m-1} Y_j B_j(u_i) - y_i \right)^2 \end{aligned}$$

where u_i is some parameter value associated with the i^{th} data point [6]. Minimizing R is equivalent to setting all partial derivatives $\partial R / \partial X_l$ and $\partial R / \partial Y_l$ to 0, for $0 \leq l < m$, which yields the following systems of linear equations

$$\sum_{j=0}^{m-1} X_j \sum_{i=0}^{p-1} B_j(u_i) B_l(u_i) = \sum_{i=0}^{p-1} x_i B_l(u_i) \quad (4.1)$$

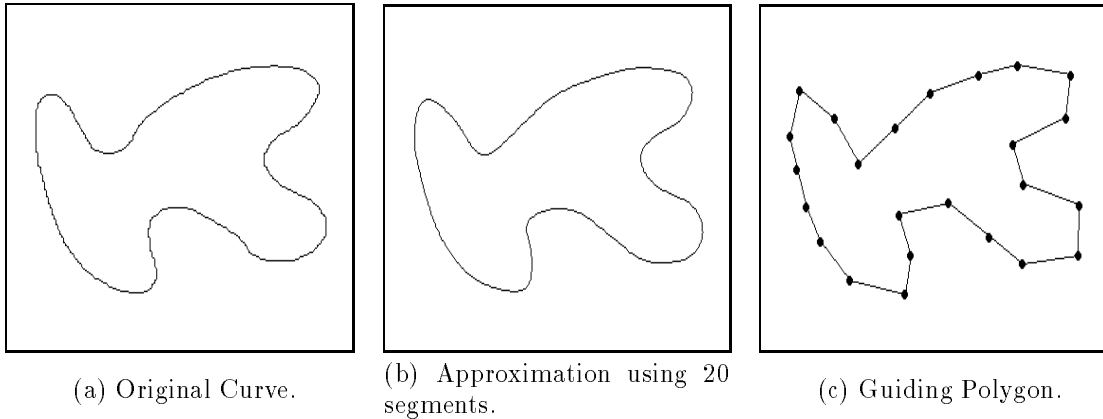


Figure 4.3: Example of a B-spline approximation of a planar curve.

$$\sum_{j=0}^{m-1} Y_j \sum_{i=0}^{p-1} B_j(u_i) B_l(u_i) = \sum_{i=0}^{p-1} y_i B_l(u_i) \quad (4.2)$$

with $0 \leq l < m$

The linear systems, (4.1) and (4.2), are easily solved for all X_j and Y_j , respectively, using standard linear algebra. This yields the guiding polygon of the B-spline which best approximates the original curve, in the least squares sense.

The choice of m (the number of vertices) determines how close is the approximation to the original data. This is measured by R (see above). The automatic selection of the number of vertices is not trivial. Our approach is to preliminary set a fitting tolerance r_0 and find the value of m which yields the normalized distance $r = R/p$ closer to r_0 using a binary search approach. Figure 4.3 shows an example of a planar curve, its B-spline approximation, and the guiding polygon.

Besides providing a simple analytical representation, the following features make this approximation attractive for our application:

- The B-spline is easily manipulated by modifying its guiding polygon. It is also defined locally, so changing the position of a vertex of the guiding polygon does not have a global effect on the representation. As shown later, this allows us to remove parts efficiently without changing the entire representation.
- Quadratic B-splines are continuous and smooth (i.e. have continuous tangents). Discontinuities of the tangent can be introduced by using multiple control points.

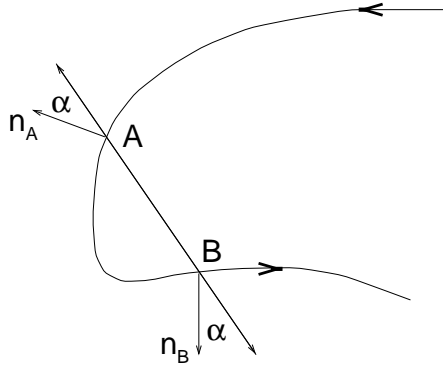


Figure 4.4: The geometry of a local symmetry.

- Each quadratic spline has *constant sign* curvature, implying that zero crossings of curvature can occur only at the *knots* (the connection points between two splines). This bounds the number of parts (as defined in Section 4.3.3) to be less than the number of control points, and therefore guarantees the termination of the decomposition process, described in Section 4.3.5.
- Quadratic splines have at most one curvature extremum. Many researchers have discussed the importance of curvature extrema for shape segmentation [2, 10, 40]. We also use curvature extrema in our definition of parts. The limits on the number and location of the curvature extrema simplify the complexity of our process significantly.

In the rest of this chapter, we ignore the problem of obtaining the B-spline representation from the real data. From now on, we assume that our input consists of a *single closed* planar curve. Furthermore, we often do not make the distinction between the input curve and its B-spline representation. We do not attempt to solve the *figure-ground* problem, although we recognize that solving this problem is essential prior to shape description in practical implementations. We discuss this issue further in the conclusion of this chapter in Section 4.5.

4.3.2 Smooth Local Symmetries

Brady and Asada [18] have introduced *Smooth Local Symmetries (SLS)* as a method for planar shape description. Two points A and B on a planar curve form a *local symmetry* if the angle between the vector in the direction BA and the normal to the curve at A is equal to the angle between the normal to the curve at B and the vector in the direction AB (see Figure 4.4). In general, a point may form several

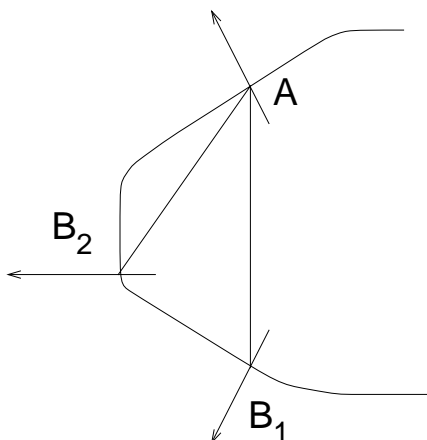


Figure 4.5: A point may form more than one local symmetry.

different local symmetries, as in Figure 4.5. The line between two points which form a local symmetry is known as the *cross section*. The loci of the mid points of the cross sections, are the *symmetry axes*. A symmetry axis together with its cross sections is sometimes called an *SLS ribbon* or a *Brady ribbon*. We say that the region of the shape covered by the ribbon is *explained* by the ribbon. It has been shown [35] that, in general, an SLS ribbon with smooth boundaries has a smooth axis.

Two of the major problems with using SLS are sensitivity to noise and the computational difficulty of recovering the SLS axis. The direct method for recovering the SLS axis by comparing every point of the curve to every other point yields an $O(N^2)$ algorithm, where N is the number of points on the curve. Brady and Asada [18] have overcome these problems by approximating the curves with circular arcs. In this case, finding the SLS reduces to the solvable case of finding the SLS between circular arcs. The complexity then is $O(n^2)$ where n is the number of circular arcs in the approximation. Since the number of circular arcs is far less than the number of points, this algorithm is computationally more attractive.

As suggested by Saint-Marc *et al.* [85], a similar approach can be taken while approximating the curve with quadratic B-splines. In this case, however, since the axis between two quadratic spline segments cannot be recovered analytically, it is approximated as follows: First the two end points of the axis are found using search, and then the axis is interpolated between the end points (refer to Appendix A or to [85] for full details). Since, in general, this method produces very close approximations to the SLS axis, and due to its efficiency, we also use this method for recovering the SLS in our application.

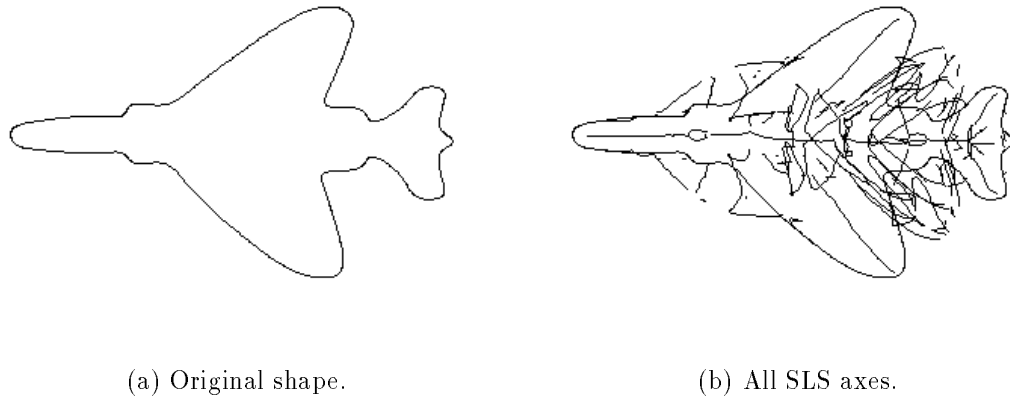


Figure 4.6: All the SLS axes of a shape.

4.3.3 Parts

We believe that the segmentation of a shape into parts is the key to any shape description used by higher tasks, such as recognition. However, a well known problem is that although a good description may help segmentation, the segmentation may be needed for obtaining a good description.

Computing the SLS within the whole shape results in many axes which are ambiguous or not relevant for a natural representation. For example, two very distant segments may also create a symmetry axis, which may even lie outside the shape. Figure 4.6 shows an example of an airplane shape and all of its SLS axes. Grouping the axes and selecting the significant ones is a difficult and computationally expensive problem (for an example of such an implementation refer to [26, 27]). Therefore, we suggest that the shape be segmented into “natural” parts prior to the computation of the SLS. The SLS are then computed within the parts only. This, of course, forces us to address the issue of defining parts.

The Definition of a Part

Hoffman and Richards [40], and more recently Biederman [9, 10] argue that a segmentation into parts occurs at the negative minima of curvature. This corresponds to the human intuition about parts, and is based on the fact, known as *transversality* [37], by which, in general, concavities arise when two convex volumes are arbitrarily joined. Kimia *et al.* [49], however, note that although parts are bounded by curvature minima, the converse does not necessarily hold. See for example Figure 4.7: we

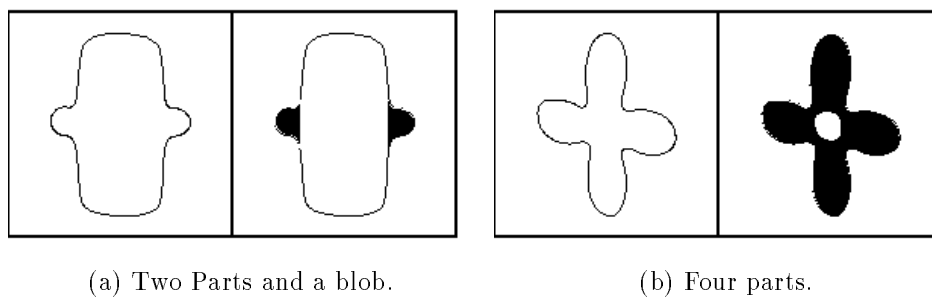


Figure 4.7: Segmenting at curvature minima results in 4 parts in both cases.

perceive the two shapes as having a different structure, one being composed of two parts and a large blob, and the other is composed of four similar parts. Segmenting at curvature minima results in four parts in both cases.

We also recognize the importance of the transversality principle for the perception of parts. Therefore, given a shape, every positive curvature curve section bounded by negative curvature curve sections implies a *potential part* in the shape. However, at the current level of the hierarchical description process, only the smaller potential parts, which are completely defined, are considered *parts*, and they are removed before the larger ones are analyzed at the next levels. By this hierarchical decomposition process (see Section 4.3.5), we overcome the above problem raised by Kimia *et al.*

Note that in contrast to Hoffman and Richards, parts are not segmented exactly at the matched negative curvature minima (we believe that this does not always correspond to the intuition of the parts termination). In our approach the curvature sign changes imply the *existence* of the parts only. The final delineation of each part (the exact termination points), is inferred from the symmetry axis describing it.

Describing Parts

As seen in Section 4.3.2, a point may belong to several different local symmetries. This may lead to several ribbons, giving different explanations to the same regions of the shape. We believe that this is a major drawback of the descriptive power of the SLS. Fortunately, when considering parts as defined above, there is a known constraint on the number and location of the axes within them. The following theorem, proposed and proved by Leyton [54], relates the curvature of a curve to the axis of symmetry describing its shape:

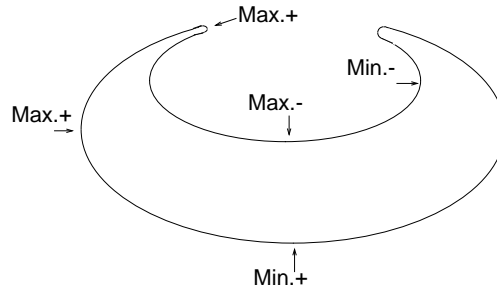


Figure 4.8: Curvature extrema. Note that it is hard to locate positive curvature minima, $\text{Min}+$, and negative curvature maxima, $\text{Max}-$.

Symmetry-Curvature Duality Theorem: Any segment of a smooth planar curve, bounded by two consecutive curvature extrema of the same type, has a unique symmetry axis (SLS, as well as, SAT [14, 15], and PISA [55]), and the axis terminates at the curvature extremum of the opposite type.

Note that Leyton includes positive curvature minima and negative curvature maxima in his definition of curvature extrema (Figure 4.8). However, since it is hard, if not impossible, to find their precise location, we consider only positive curvature maxima and negative curvature minima. Therefore, the theorem guarantees the existence and uniqueness of an SLS axis describing parts with one positive curvature extremum. For every additional positive curvature extremum within the part, there exists an additional axis going into it.

Given a shape, we first segment the contour into sections bounded by consecutive negative curvature curve sections. These sections are the potential parts. We compute the axes of symmetry of all parts. From the stated theorem, we now know the origins of the axes, and we also know on which side of the ribbon every quadratic segment may be. Therefore, we have very few comparisons to make and we can recover the SLS axes explaining each part very efficiently. Figure 4.9 shows an example of a shape, and its initial segmentation into potential parts with their axial description.

Special Cases: Terminations, Bends, and T-shapes

As described above, parts which have more than one positive curvature extremum within them, are initially described with more than one SLS axis. We have identified three special cases, which occur frequently, where a more intuitive single axis can be

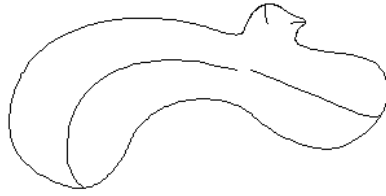


Figure 4.9: A shape and the axes of its potential parts.

found to describe these parts. Refer to Figure 4.10a for examples of these cases which we have intuitively named: *Termination*, *Bend*, and *T (or Mushroom)*. Figure 4.10b presents the SLS axial description of the corresponding shapes. Figure 4.10c presents our preferred description.

Once a part with two maxima of curvature is detected, the three possible interpretations are tested and the best is chosen as the parts description. In our current implementation the interpretations are compared in terms of the variance of their cross section function. The results in Figure 4.10c and in Section 4.4 demonstrate that the correct decision is made by using this simple criterion.

Removing Parts

To produce a hierarchical decomposition, we need a mechanism for removing parts from the shape (see Section 4.3.5). Once a part is identified, it is not clear how one removes it “gracefully.” The direct method of simply connecting the two minima does not always produce the intuitively correct result. In addition this is not an easy procedure using B-splines.

By removing control points of the guiding polygon we are able to remove parts efficiently and locally (i.e. not affecting other parts). Given a part, it is represented by a B-spline $\mathcal{S} = \{C_i(u)\}$, for $i = 0, \dots, n$, where C_0 and C_n contain the part’s end points (\mathcal{S} is a subset of the spline describing the whole shape). Let $p_0, p_1, p_2, \dots, p_n, p_{n+1}, p_{n+2}$ be the control points generating \mathcal{S} . We remove all points p_2, \dots, p_n . Consequently, all the quadratic curve segments between C_0 and C_n , are removed. Since we are left with 4 control points, the gap is automatically closed smoothly with two splines. Figure 4.11 demonstrates the performance of the above removing procedure on a shape and on its guiding polygon.

(c) Selected axis for description

Figure 4.10: Special cases of multiple extrema: Termination, Bend, and T.

4.3.4 Representing Global Relationships

In Sections 4.3.2 and 4.3.3, we have presented SLS as a method for describing local parts. Leyton's theorem provides the basis for the efficient computation of the axes in these parts. Unfortunately, this description provides only local information. It is clear that this cannot suffice for an intuitive description. Consider for example the "snake" in Figure 4.12a. Using only the local information the snake would be

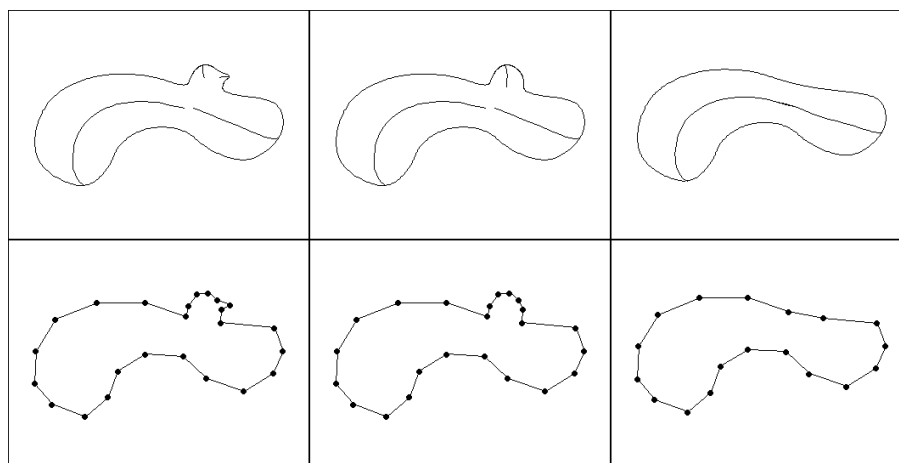


Figure 4.11: Removing parts: The shape and its guiding polygon.

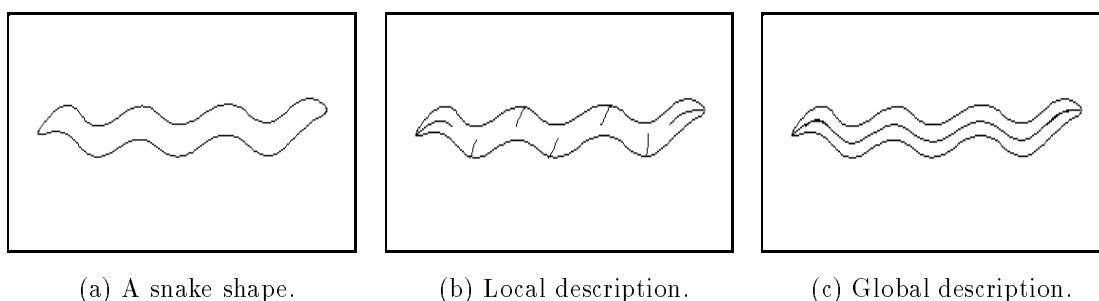


Figure 4.12: Global relationships are important for natural descriptions.

described as a composition of seven different parts (Figure 4.12b). We, however, would prefer to capture the global relationship and describe the snake as a single entity using the more intuitive symmetry axis, as in Figure 4.12c.

Parallel Symmetries

We have found *parallel symmetries*, suggested by Ulupinar and Nevatia [96], to be very useful in capturing global relationships within planar shapes. Let $C_i(s) = (x_i(s), y_i(s))$, for $i = 1, 2$ be two parametric planar curves, and let $\theta_i(s)$ be their tangent orientation. $C_1(s)$ and $C_2(s)$ are said to be *parallel symmetric* if there exists a continuous monotonic function $f(s)$ such that $\theta_1(s) = \theta_2(f(s))$. The symmetry axis is the loci of the mid points of the cross sections between $C_1(s)$ and $C_2(f(s))$ for all s where $f(s)$ exists. When $C_1(u)$ and $C_2(v)$ are only defined on the interval

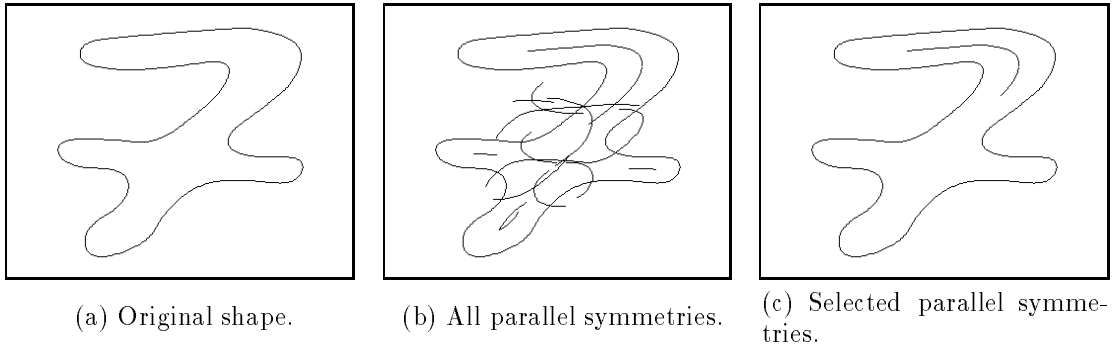


Figure 4.13: Parallel symmetries, before and after filtering.

$[0, 1]$, we say that the two curve segments are parallel symmetric *on* $[u_0, u_1] \subseteq [0, 1]$ iff $[f(u_0), f(u_1)] \subseteq [0, 1]$. The symmetry axis together with the cross sections is a *parallel ribbon*. We say that the part of the shape covered by the ribbon is *explained* by the ribbon. Figure 4.12c is an example of a parallel symmetry axis, in this case the two curves are sections of the contour of the same object.

As shown by Saint-Marc *et al.* [84, 85], the detection of parallel symmetries between two quadratic B-spline segments is computationally efficient. Using this algorithm we compute all the elementary parallel ribbons (ribbons between two spline segments) within the shape. This process produces many elementary parallel ribbons. Further filtering and grouping are needed to produce the parallel ribbons which are significant for the description of the shape. We shall often refer to these significant parallel ribbons as *global ribbons*. (Note that these ribbons once found could also be described with SLS).

Selecting Significant Parallel Ribbons

We have found several simple intuitive criteria for deciding that a ribbon has no (or very little) importance for the shape perception. These ribbons are filtered out as described below. Figure 4.13 shows an example of a shape, all elementary parallel symmetry axes found, and the significant parallel axes selected after applying the following filtering. The order in which the different filters are applied was decided according to efficiency considerations. The results are the same in any order.

First, we filter out all *skewed* ribbons. Let $C_1(u)$ and $C_2(f(u))$ be parallel symmetric on $[u_0, u_1] \subseteq [0, 1]$. We define the ribbon to be skewed if for all $u \in [u_0, u_1]$ the angle between the normal to C_1 at u and the cross section from $C_1(u)$ to $C_2(f(u))$ is very large (larger than 45° in our implementation). Figure 4.14a shows examples of skewed parallel ribbons.

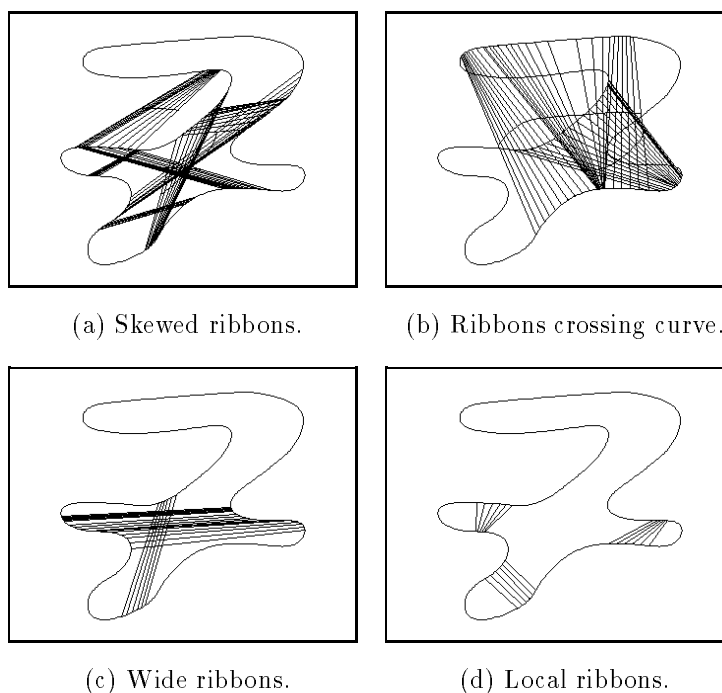


Figure 4.14: Parallel ribbons which are filtered out.

Next, we examine every ribbon to see if for all u , as above, the cross section from $C_1(u)$ to $C_2(f(u))$ does not intersect with the curve, excluding of course the cross section's end points. We filter out ribbons for which this condition does not hold (Figure 4.14b).

The ribbons found so far consist of symmetries between single spline segments only, they may, however, be part of larger parallel ribbons. Therefore, we group all connected ribbons into maximal parallel ribbons. Now each ribbon may cover several spline segments on both sides.

After linking, we remove ribbons in which the axis is very short compared to the “thickness” (i.e. the length of the cross sections) of the ribbon (Figure 4.14c). Intuitively, we consider these thick ribbons to be non “ribbon like” and we believe that their global information is less significant. (Consider for example a long and narrow rectangle: Given the two parallel axes, each parallel to two sides and bisecting the other two, we as humans prefer the longer axis as a representative of the rectangle).

Finally, we filter out parallel ribbons which are already totally explained by a local ribbon. If the curve sections of the parallel ribbon (the sides of the ribbon) are totally included in the curve sections of some local ribbon found earlier, then

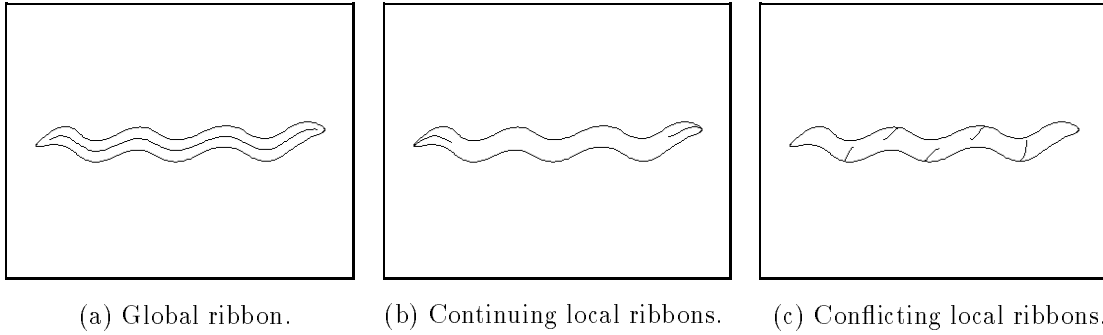


Figure 4.15: Relationship between global and local ribbons.

this parallel ribbon does not add any global information and is therefore removed (Figure 4.14d).

Relating Global Ribbons and Local Ribbons

Once a global (significant parallel) ribbon is found, it is important to record its relationships with the local parts. A local ribbon (part) may continue the global ribbon, thus the two ribbons support each other in the parsing. A local ribbon may also conflict with the global ribbon, when both ribbons give a different explanation of the same region of the shape. In this case they cannot be used together in a description of the shape. Figure 4.15 shows an example of a global ribbon and its continuing and conflicting local ribbons. For every global ribbon, we keep pointers to the continuing parts and to the conflicting parts. In the next section we describe how the global ribbons are used for deciding on branching when decomposing the shape.

4.3.5 Shape Decomposition

We now describe the decomposition of the shape, a planar closed curve, into a hierarchy of parts, based on the size of the parts and on global relationships.

The shape is decomposed using a recursive procedure: given the current shape, actually the B-spline approximation of the shape, we compute all local parts, as defined in Section 4.3.3, and all the global ribbons, as defined in Section 4.3.4. Every local part is represented by an SLS ribbon (axis and cross sections), and every global ribbon is represented by a parallel ribbon and pointers to its local continuations and local conflicts (see Section 4.3.4). We now create a new shape, generated from the current shape by removing its *smallest* parts in parallel. In our practical implementation we define a *small* part to be a part with an axis not longer

than 1.5 times the shortest axis. In case of a conflict with a global ribbon we generate the two possible interpretations, the first ignoring the global relationships and the second considering the global relationships. (We do not attempt to decide which is the correct interpretation. We believe that this decision usually requires higher level knowledge). Under the second interpretation a new shape is generated from the current shape by removing its smallest parts, this time using the information on the global relationships. The global information is used by assigning high size values to parts which are related to the global ribbon, either continuing or conflicting with the global axis. Therefore, these parts are removed only if there are *no* local parts which are not related to the global ribbon. Note that the second interpretation is generated only if the removed parts are different than the removed parts in the “all local” case. The number of interpretations is, therefore, bounded by $n + 1$ where n is the number of global ribbons in the shape. However, in most cases there are no significant global relationships within the shape, and the parsing of the shape is unique. The process is continued recursively on each branch.

Since we only remove parts, the complexity of the shape, in terms of the number of control points of the guiding polygon, reduces from level to level. Therefore, the process is bound to terminate rapidly. The recursive decomposition stops (an atomic state) in one of the following cases:

- The shape is empty. This could happen if during the previous step, all parts had about the same size.
- The shape is a positive curvature blob (no zero crossings of curvature).
- The shape has exactly two zero crossings of curvature (a bean shape).

The axes of the parts removed are accumulated down the decomposition hierarchy. Therefore, at every atomic level we have a skeleton representation of the shape, given by the single parsing of the shape, defined by the path from the original shape to that atomic state.

4.4 Experimental Results

We have applied our method to several shapes. For each, we present the decomposition, which is the output of our description process. The original shape is on the left. At every node the parts which were previously removed are shaded. The current shape, not yet explained, is left white. As explained earlier, in Section 4.3.5, the axial representations of the parts removed are accumulated. Therefore, every interpretation generates a skeleton description of the shape. That skeleton description is presented at the end of the decomposition on the right.

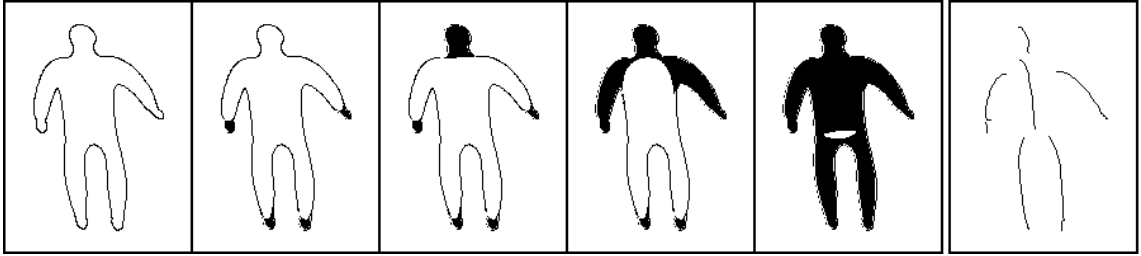


Figure 4.16: Decomposition of a man shape (inspired by [49]). The original shape is on the left. The skeleton description is on the right.

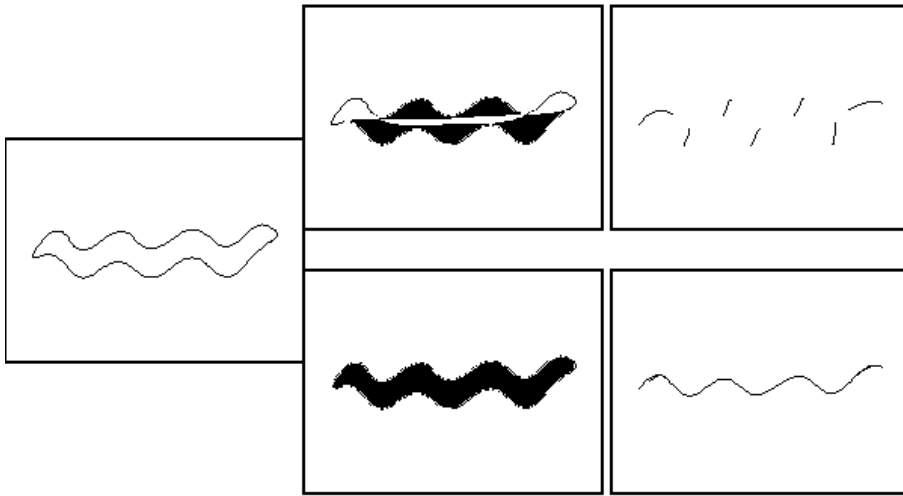


Figure 4.17: Two decompositions of a snake shape.

The first example is the man shape (Figure 4.16). Since there are no significant global ribbons in this shape and in its sub shapes, there are no branches in the interpretation process. There is a unique decomposition of the shape into a hierarchy of parts. Note that, since in our current implementation we define the size of the part to be the length of the axis, at the last step the legs and the torso are removed in parallel, leaving the final blob (one of the atomic cases discussed in Section 4.3.5) at the last stage. If the criterion was different (e.g. combination of area and length), then there may have been an additional stage in the decomposition process (first the legs would have been removed and then the torso). It is important to note, however, that the final skeleton description would still be the same.

The next example is the snake shape (Figure 4.17). As discussed in Section 4.3.4, this is an example of the importance of global relationships. Segmenting at minima of curvature (e.g. Richards and Hoffman [77]), or interpreting the shape as a result of a history of protrusions and indentations (e.g. Leyton [55]), results in the 7 parts

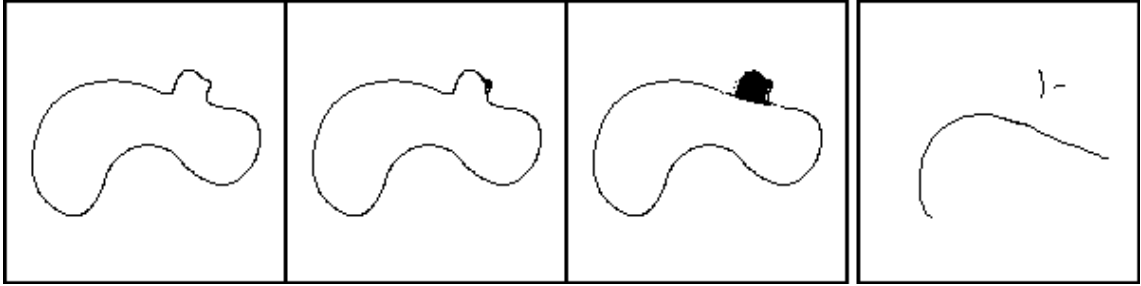


Figure 4.18: Decomposition and skeleton of a “bean” shape.



Figure 4.19: Decomposition and skeleton of a seven shape.

description, as produced by the top interpretation path, which is the result of our local interpretation. Using the global ribbon, we are able to produce the other, perhaps more intuitive, description on the bottom path.

In the next four examples, the bean shape, the seven shape, and two sketches of airplanes, figures 4.18 through 4.20 respectively, there is once again no significant global information conflicting with the local one. Therefore, they each have a unique decomposition leading to the skeleton description.

The next example illustrates the inherent stability of our hierarchical approach. In Figure 4.21a we present the decomposition of an unspecified shape. We have taken this shape and added several bumps to its curve. A good shape description mechanism should be stable enough to produce a description which is similar to the original and yet rich enough to show the differences. As Figure 4.21b shows, our algorithm performs well on this example, by detecting and removing all the small bumps in the first step, after which the process is identical to the original case. A higher level task interpreting these results could easily examine the similarities and differences between the two shapes.

All the examples we have presented so far were based on synthetic data. We have also applied our method to shapes obtained from real images of model airplanes taken on a light table. The results are presented in Figure 4.22. Note also that the “*F16*”

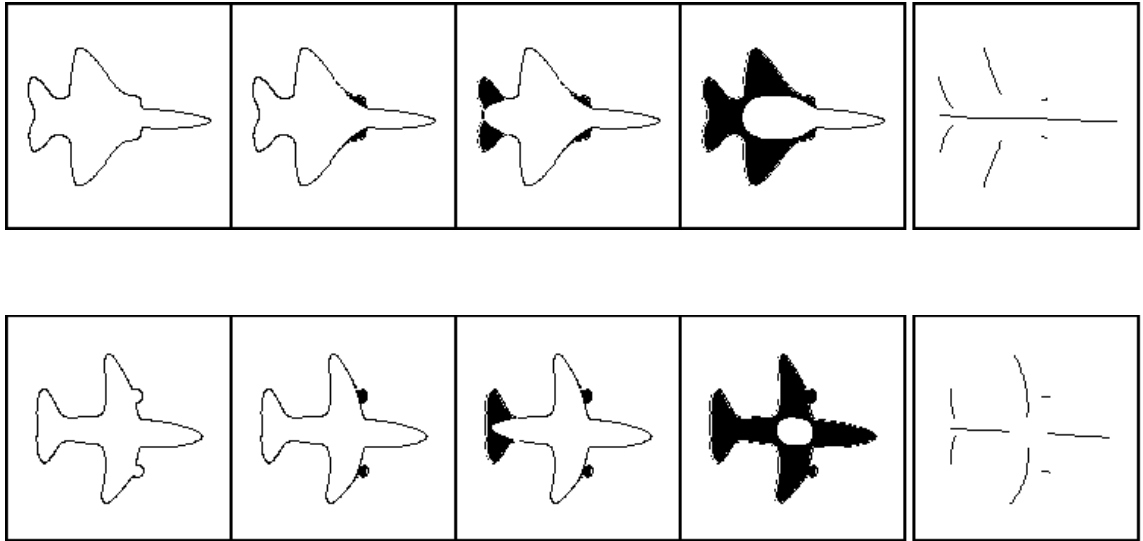


Figure 4.20: The decomposition of airplane shapes.

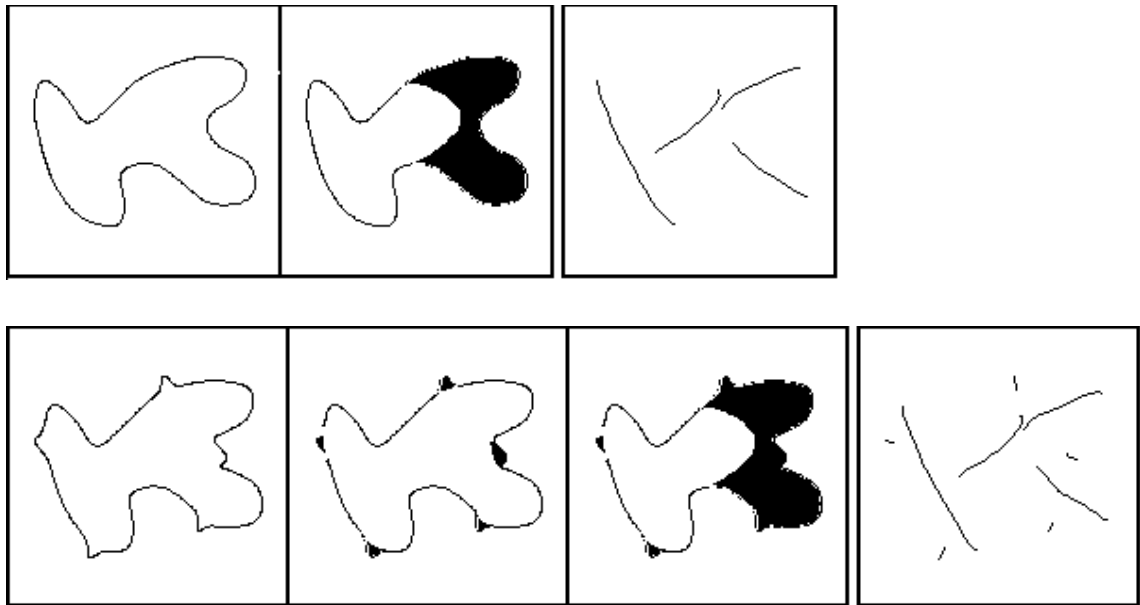
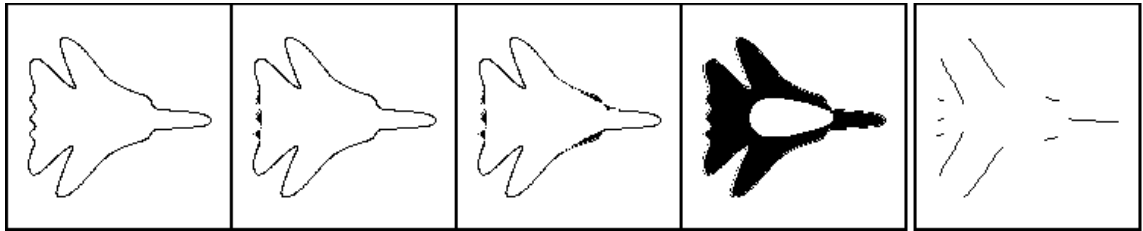
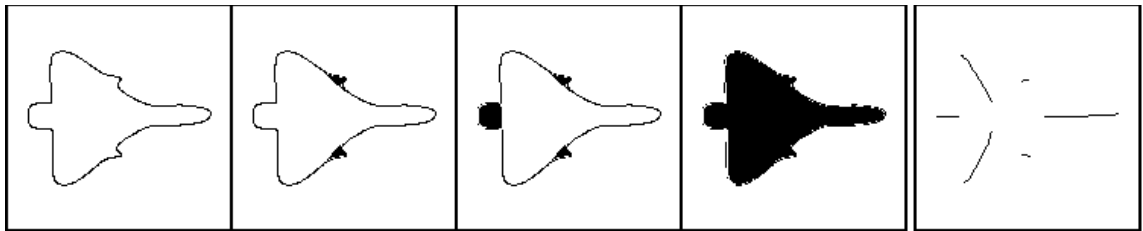


Figure 4.21: Example of the stability of the process (see text)

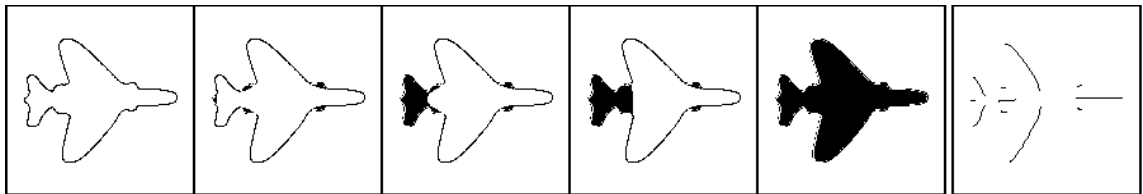
shape of Figure 4.22d has sharp corners (i.e. discontinuities of the tangents). These corners are simply represented using multiple control points in the guiding polygon. They are considered as any other curvature extrema, and do not require any special treatment in the decomposition process.



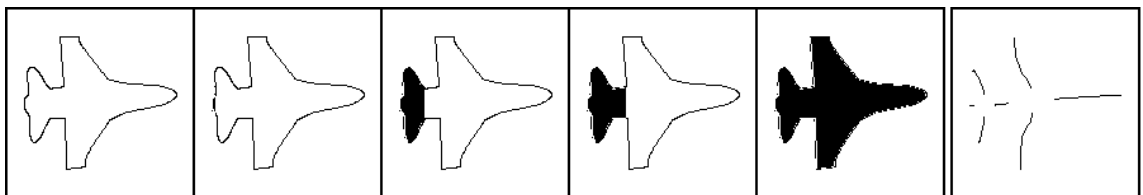
(a) Decomposition and skeleton of an “F14” shape.



(b) Decomposition and skeleton of an “F106” shape.



(c) Decomposition and skeleton of an “F4” shape.



(d) Decomposition and skeleton of an “F16” shape.

Figure 4.22: The decomposition of airplane shapes obtained from real data.

4.5 Concluding Remarks

We have presented a powerful method for obtaining natural descriptions of planar shapes. Our method produces an axial description of a shape, and a discrete hierarchical decomposition of the shape into its parts. Our approach is both region and contour based, combines local and global information, can handle corners, and addresses the issue of scale and the notion of part. Our method is computationally efficient, robust to changes in parameter settings (all experiments are performed using the same settings), and stable. We have presented results which show that it provides intuitive shape descriptions for various shapes.

In the following we discuss some limitations of our work and issues which require additional investigation.

The Figure-Ground problem As noted in Chapter 2, one of the major problems in obtaining shape descriptions is the segmentation of the image into the object vs. the background. This is known as *the figure-ground problem*. Due to texture, surface markings, noise, and other factors, we cannot always determine the exact closed contours of the object. It is clear that for many realistic applications, this problem has to be solved prior or together with the process of recovering the shape description. In the presented work, however, we have assumed that the given shape is a closed curve. We did not solve or address the figure-ground problem. The questions which arise are, therefore, whether there is a place for the ideas suggested here in an eventual complete vision system? Or, is there a contribution here, beyond the pure academic interest, towards the solution of the vision problem?

We believe that the answer to both of these questions is *yes*. First, for many applications the requirement for an input of a closed simple curve is quite feasible. These applications include scenarios, such as robotic applications in controlled industrial environments (e.g. part sorting or grasping [36, 13]). For these applications our methods are directly applicable.

In addition, there is currently a lot of research effort in areas, such as *perceptual grouping*. These efforts attempt to solve the figure-ground problem. Although it is unlikely that a complete general solution will be found in the near future, it is reasonable to assume that in many restricted domains a solution could be found. In these cases, since the shape is finally given, once again our methods will be applicable.

Finally, even in the most general cases, where the figure-ground problem is not completely solved, our methods could still be applicable, perhaps in a combination with a perceptual grouping process. Recall that our methods are completely local in nature. Therefore, even in the presence of occlusions, surface markings or noise,

meaningful descriptions of local areas could still be recovered. Grouping and filtering could be applied both on the curve segments, in order to overcome breaks for example, and also on the recovered parts and symmetries, in order to select the significant and meaningful ones. It is of course likely that an image will give rise to several hypothesized interpretations, but hopefully it will be easy to select one or a few significant ones. Although these general cases will require modifications to our methods, we believe that our approach and methods could be of use, both for aiding the perceptual grouping, and for recovering the description. An example of such an approach which combines the recovery of the shape description with the perceptual grouping, was suggested by Zerroug and Nevatia [103, 104]. In their case the symmetries and groupings are used for recovering 3-D shape descriptions.

Stability on low curvature curves A problem we have encountered with our current implementation, is that due to artifacts of the spline representation, the results may not be stable on very low curvature curves. This is because when approximating such a curve, the spline approximation might fluctuate between positive and negative curvature segments. This may be solved by ignoring such changes of the sign of the curvature if the absolute values are very small. Another possible solution is to detect these low curvature curves and replace them by straight lines. This could be done easily by introducing multiple knots in the spline representation. For straight lines the symmetry algorithms are straightforward and our processes do not exhibit in problems.

Negative curvature “parts” In the presented work, we consider only positive curvature parts. The negative curvature curve sections were used as indications for the presence of the positive curvature parts. They were not considered as entities by themselves. However, for intuitive descriptions it is sometimes needed to capture negative curvature parts, as in the famous example of a rectangle with a sharp indentation. One approach to address this problem is to reverse the figure-ground roles and treat the concavities as the dual of the convexities. Analysis is then needed to decide on the correct interpretation.

Correctness and completeness The goal we have set for our system was to produce an intuitive description of a given shape. The issue of what is a “correct” interpretation is not well defined and is not quantifiable. We have currently judged the results by subjectively evaluating how reasonable they are, that is, we asked the question: Is it reasonable that a human would give a similar description to the shape as our system did? Scientific psychological tests on humans and comparison to our systems performance were not performed, but could turn out to be useful.

In addition, a shape may have several acceptable interpretations. It is not clear how to determine which is the most reasonable or intuitive interpretation. Currently we require our system to produce one intuitive description. In order to facilitate consistency and stability of the interpretation under deformations of the shape, and especially if the description is used for recognition, it would be useful for the system to be able to generate several, if not all, of the intuitive interpretations. Groupings of axes and parts based on criteria, such as co-linearity, parallelism, and orthogonality, could also be of great importance.

Chapter 5

Description of Three-Dimensional Shapes

5.1 Introduction

In this chapter we address the problem of recovering *segmented volumetric* descriptions of three dimensional shapes. We suggest a *volumetric* graph representation of the shape, where the nodes represent individual parts and the edges represent connectivity information. In an analogous way to our 2-D approach (see Chapter 4), we suggest the use of properties of the parabolic curves (curves of inflection of the Gaussian curvature) for performing the part decomposition. The graph description presents a structural description of the shape in terms of parts and their arrangement. We handle components with tubular structure with a straight or curved axis. We are also interested in the internal description of the parts. This is especially important when performing tasks, such as grasping. We study two well defined classes of shapes, namely Straight Homogeneous Generalized Cylinders (SHGCs [88]), and Planar Right Constant GCs (PRCGCs [98], planar axis and constant cross section). We suggest the use of properties of the parabolic curves for recovering natural descriptions of these classes in terms of their cross sections and axes.

5.1.1 Background

In the previous chapter, we have suggested a method for obtaining hierarchical axial descriptions of *planar* shapes, together with a decomposition of the shapes into their parts. Unfortunately, it is not straightforward to extend this method, and other axial based planar methods suggested by several researchers, to handle three dimensional shapes. This is because in the three dimensional space the SAT and SLS axes are, in general, not curves, but surfaces, leading to unnatural descriptions [63].

The problem of obtaining descriptions of three dimensional shapes has received considerable attention. Several researchers have addressed the problem of recovering three dimensional shape descriptions from their two dimensional imaged contours

(e.g. [50, 19, 72, 97, 7, 102]), while others have used range data as input (e.g. [1, 69, 95, 66, 28, 87]). Due to the difficulty of the problem, assumptions are usually made, restricting the type of objects viewed to be either a restricted class of Generalized Cylinders (GCs) or other basic components (e.g. geons), or to possess certain properties such as symmetries.

Recent theoretical work has produced several important results regarding the properties of certain classes of GCs and their projections [72, 97, 104]. These properties can be used to recover the 3-D information of a simple observed GC. The issues of compound objects which are composed of several of these GCs are not addressed. Compound objects are far more complex than their GC components, since even though the individual components have certain mathematical properties, these properties do not hold for the shape as a whole, and it is necessary to segment the shape into its components. This is the problem we address here.

5.1.2 Issues and Approach

We are interested in producing *natural* descriptions of three dimensional *compound* shapes. It is clear that this problem involves two tasks, namely, the decomposition of the shape into parts, and the description of these parts. As in the 2-D case, we strongly believe that these two tasks are interconnected and cannot be performed independently. It is through the description that parts could be segmented, and on the other hand, based on the decomposition the different parts can be identified and a description attempted.

Decomposition As discussed in Chapter 3, most of the work on 3-D shapes so far has concentrated on the recovery of descriptions of single components. When compound shapes are addressed, the segmentation is assumed to be an easy step. We attempt to handle compound shapes where the parts are connected smoothly. We do not assume the existence of any discontinuity edges between parts. We believe that the case of parts joined discontinuously are the limiting case of the more general continuous case which we address. Although our method could handle such cases, they could also be handled using other methods and they will not be discussed here.

We suggest a *volumetric* graph representation of the shape, where the nodes represent individual parts and the edges represent connectivity information. Graph representations have been suggested previously (e.g. [8, 31, 66, 67, 34]), but those are surface representations (i.e. the nodes represent different surfaces) and not volumetric in terms of parts. In an analogous way to our 2-D approach (see Chapter 4), we suggest the use of properties of the parabolic curves for performing the part decomposition. We currently consider shapes with convex blobs and “tube like” parts.

The parabolic curves could be either on the surface of the individual parts, or on the border of the “glue” between parts. Note that, due to the transversality principle [37], there is almost always an anticlastic (negative Gaussian curvature) region between convex parts where they are joined. Using simple tests we can hypothesize the role of each parabolic curve and generate a segmented graph description.

Part description The graph description presents a structural description of the shape in terms of parts and their arrangement. We are also interested in the internal description of the parts. This is especially important when performing tasks such as grasping. We study two well defined classes of shapes, namely Straight Homogeneous Generalized Cylinders (SHGCs [88]), and Planar Right Constant GCs (PRCGCs [98], planar axis and constant cross section). We suggest the use of properties of the parabolic curves (curves of inflection of the Gaussian curvature) for recovering natural descriptions of these classes in terms of their cross sections and axes. As we show, the parabolic curves on the surfaces of SHGCs and PRCGCs are either the meridians or the cross sections. Using simple tests we can hypothesize (or in many cases determine) the role of each parabolic curve. The axis is easily recovered from the meridians and cross sections. We advocate the use of the parabolic curves over the, often used, occluding contours, since the latter are unstable in range data. In addition, due to the use of the parabolic curves, we do not need to assume, as several authors do, that the objects are cut along a cross section or that a cross section is visible.

Unfortunately, not all parts are perfect SHGCs and PRCGCs. It is clear that these techniques could not be used to describe all possible parts. However, in practice, many parts are only slight deformations of the above classes and, therefore, these methods could still result in meaningful descriptions.

Figure 5.1 shows the outline of our approach. Starting from the input, which is either complete or from one view, we first recover the surface properties, namely, compute the sign of the Gaussian curvature, connect regions of the same sign, and extract the parabolic curves. This results in a surface based description. Based on the classification of the parabolic curves, we generate the volumetric descriptions in terms of parts, parts arrangement, and axial descriptions of some classes of parts.

5.1.3 Outline of Chapter

The outline of this chapter is as follows: In the next section we discuss some preliminary issues. We first discuss the types of input we work with, we then present a brief overview of the basics of differential geometry, and finally we discuss some

Figure 5.1: Outline of our approach

generic 3-D shapes and parabolic curves. In Section 5.3, we present our method for generating segmented volumetric graph descriptions of compound shapes based on the classification of their parabolic curves. In Section 5.4, we discuss the recovery of axial descriptions of well defined parts. We first present some known properties of SHGCs, and then discuss how these properties are used for recovering their description in terms of their axis and cross sections. We then discuss PRCGCs. We derive the formulation for the Gaussian curvature of PRCGCs. We then state and prove two important properties of PRCGCs. As in the case of SHGCs, we show how these properties can be used for recovering their description in terms of their axis and cross sections. Finally, we present some concluding remarks in Section 5.5.

5.2 Preliminaries

5.2.1 The Input

We address the problem of recovering volumetric descriptions of three dimensional shapes. Humans can generally recover volumetric descriptions based on 2-D information alone. Unfortunately, we are far from understanding these processes, and even further away from imitating them in machine vision systems. As often is the case when addressing difficult problems in scientific research in general, and in computer vision research in particular, progress is made by making simplifying assumptions and by solving less general problems. The way we choose to progress in understanding the problem is to limit ourselves to 3-D surface input. We believe that the solutions we derive for this limited case will be of help in understanding the more general case. In addition, for many applications of machine vision systems it could be reasonable to assume the existence of 3-D data from either stereo or from a range finder. Furthermore, progress in some other fields of computer vision, such as shape from shading [42, 43] and shape from contour [98, 104], may result in methods for recovering reliable surface information from 2-D data.

In order to achieve complete volumetric descriptions, complete 3-D surface data of the input shape is needed. This input could be given by a CAD model or from registered range images, in terms of surface patches. We have developed mechanisms for handling triangular patches. Unfortunately, such complete data is only available for very simple objects. Research on the problem of producing complete surface data from multiple registered range images of more complex objects is currently underway (e.g. [91, 22, 23]).

Due to the problems of availability of complete data, we also consider input from a range image from a single view. We have developed tools and methods to handle such data. Due to the incompleteness of the data, the resulting descriptions will also be incomplete. We do not attempt to explain what is not seen in the given view.

Figure 5.2 shows examples of the above two types of data: the triangulation of a teapot obtained from a CAD model, and a shaded range image of a teapot.

5.2.2 Differential Geometry

In this section we give a brief overview of basic surface differential geometry concepts, used in the following sections. Further information is available in many standard textbooks (e.g. [29, 58]).

A surface is given in parametric form by $S = S(u, v)$, where the variables u and v are the parameters. In the following we will assume that the surface is regular and

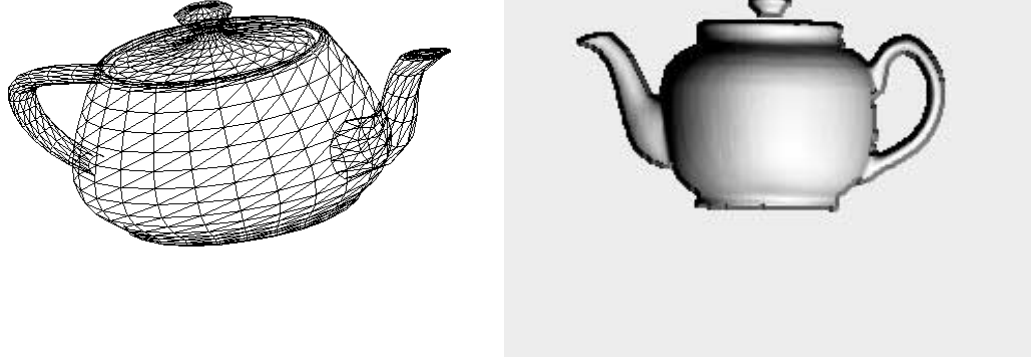


Figure 5.2: Two types of 3-D input. Left: Complete triangulated patches (generated using the SPD package by Eric Haines, 3D/Eye, Inc.). Right: A range image from one view (from the Institute of Information and Technology, The National Research Council of Canada).

of class of at least C^2 (refer to [29, 58] for exact definitions). The partial derivatives of the surface are denoted by

$$S_u = \frac{\partial S}{\partial u}, \quad S_v = \frac{\partial S}{\partial v}, \quad S_{uu} = \frac{\partial^2 S}{\partial u^2}, \quad S_{uv} = \frac{\partial^2 S}{\partial v \partial u}, \quad \text{etc.}$$

The derivatives depend on the specific parameterization chosen. However, from the derivatives we can derive properties of the surface which are independent of the parameterization.

From S_u and S_v we can derive the unit normal, n , to the surface as

$$n = \frac{S_u \times S_v}{\|S_u \times S_v\|}$$

A surface is uniquely determined by two local invariant properties called the first and second fundamental forms. The *first fundamental form* denoted by I is given by

$$I = dS \cdot dS = Edu^2 + 2Fdudv + Gdv^2$$

where

$$E = S_u \cdot S_u, \quad F = S_u \cdot S_v, \quad G = S_v \cdot S_v$$

are the first fundamental coefficients. The *second fundamental form*, denoted by II is given by

$$II = d^2S \cdot n = Ldu^2 + 2Mdudv + Ndv^2$$

where

$$L = S_{uu} \cdot n, \quad M = S_{ut} \cdot n, \quad N = S_{tt} \cdot n$$

are the second fundamental coefficients.

The first fundamental form gives the distance between neighboring points, (u, v) and $(u + du, v + dv)$, on the surface, to first order. This provides the basis for computing distances and areas on surfaces, but it does not reveal any information on how the surface “curves” away from the tangent plane. The second fundamental form gives the distance to second order. Together they are used to derive surface curvature information, as discussed below. As noted above, they uniquely define the surface.

Given a point P on a surface, the discriminant of II , $LN - M^2$, determines qualitatively the nature of the surface around P . There are four cases:

Elliptic case A point is called *elliptic* if $LN - M^2 > 0$. In the neighborhood of an elliptic point the surface lies on one side of the tangent plane (as in a hill or a dimple).

Hyperbolic case A point is called *hyperbolic* if $LN - M^2 < 0$. In the neighborhood of a hyperbolic point the surface lies on both sides of the tangent plane (as for saddle points).

Parabolic case A point is called *parabolic* if $LN - M^2 = 0$ and L, M, N are not all zero. The surface at a neighborhood of a parabolic curve may lie on both sides of the tangent plane.

Planar case A point is called *planar* if $L = M = N = 0$.

Given a point P on the surface, the planes containing the normal to the surface, n , at P , cut the surface into a family of curves called *normal sections*. Each direction, (du, dv) , in the tangent plane, at P , defines a normal section. The *normal curvature* of the surface, denoted by k_n , in a given direction, (du, dv) , is the curvature of the respective normal section and is given by

$$k_n = \frac{Ldu^2 + 2Mdudv + Ndv^2}{Edu^2 + 2Fdudv + Gdv^2} = \frac{II}{I}$$

There are two perpendicular directions for which k_n takes on maximum and minimum values. These directions are called *principal directions* and the corresponding normal curvatures, k_1 and k_2 , are the principal curvatures.

The *Gaussian curvature*, K , and the *mean curvature*, H , can now be defined as

$$K = k_1 k_2 = \frac{LN - M^2}{EG - F^2}$$

$$H = \frac{1}{2}(k_1 + k_2) = \frac{EN + GL - 2FM}{2(EG - F^2)}$$

The Gaussian curvature is an invariant property of the surface. Also, the sign of K agrees with the sign of $LN - M^2$, therefore, a point is elliptic if and only if $K > 0$, hyperbolic if and only if $K < 0$, parabolic or planar if and only if $K = 0$.

A curve on a surface whose tangent at each point is along a principal direction is called a *line of curvature*. On a surface of class of at least C^3 , there exists two orthogonal families of lines of curvature. A *parabolic curve* is a curve on which the Gaussian curvature vanishes. The parabolic curves of surfaces are the analog of the inflection points of the planar curves.

5.2.3 Basic Shapes and Parabolic Curves

In the following we introduce some basic generic shape components, and the different types of parabolic curves, which we later use for constructing the description. These shape components exhaust many, if not most, of the types of shape one is likely to encounter. We follow Koenderink [51] in the selection of these examples and in the terminology used.

Bell The Bell is a convex “wart,” or part, on a globally ovoidal shape. An example is a pear shape (Figure 5.3a). The Bell is surrounded by an annular hyperbolic (negative Gaussian curvature) region.

Dimple The Dimple is a concavity inside an ovoidal shape. For example, if you take a lump of clay and push your thumb in (Figure 5.3b). The concavity is also surrounded by an annular hyperbolic region.

Furrow The Furrow is a hyperbolic region, or dent, in an overall ovoidal shape (Figure 5.3c). Note that this is not a concavity, as the Dimple (you cannot hold water in it). An example for this type is a “kidney bean.” This region is surrounded by a single closed parabolic curve.

Hump The Hump is a convex region within a larger hyperbolic region (Figure 5.3d). This could be viewed as the dual of the Furrow which is a hyperbolic region within a convex shape.

These regions are surrounded by parabolic curves, which are the curves of inflection of the Gaussian curvature. There are two cases of such parabolic curves. In the “normal” case, as in the Bell, the lines of curvature, which inflect on the parabolic curve (the lines which generate the parabolic curve), are close to being

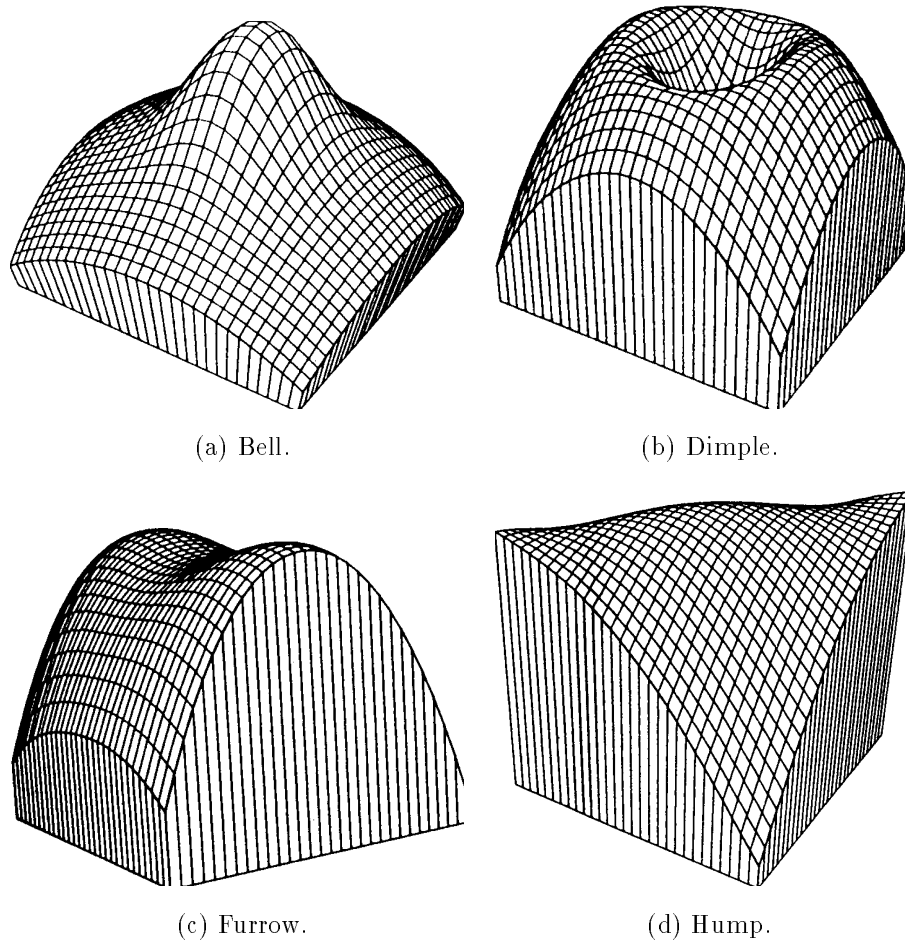


Figure 5.3: Examples of basic generic shapes.

orthogonal to the parabolic curve. We will refer to the parabolic curves of this type as being “part-like.” In the second case, for sections of the parabolic curves, the lines of curvature, which inflect on the parabolic curve, are close to being parallel to the parabolic curve. This is the case for the Hump, the Furrow, and also for the outer parabolic curve surrounding the dimple. We will refer to this type of parabolic curve as being “hump-like.” Figure 5.4 shows the two types of parabolic curves on the Bell and the Furrow. As an illustration, consider an “infinitesimal ant” walking in an orthogonal direction to the parabolic curve on the surface of the Bell. This ant will “feel” an inflection of the curvature when it crosses the parabolic curve (Figure 5.4 left). On the other hand, a sister ant walking in the same direction on the Furrow or on the Hump, will not feel any inflection when crossing the parabolic curve (Figure 5.4 right).

Figure 5.4: Two types of parabolic curves: “part-like” (left) and “hump-like” (right).

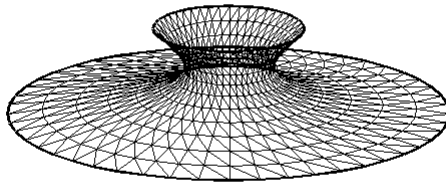


Figure 5.5: An anticlastic region (the glue between parts).

5.3 Segmented Description of Compound Objects

In this section we address the problem of recovering segmented volumetric descriptions of generic compound shapes (i.e. shapes which are composed of several interconnected parts). As discussed earlier, previous approaches have either addressed special classes of mathematically well defined shapes or they have assumed the segmentation to be given.

When convex parts are combined smoothly, unless there is an accidental alignment, there is an anticlastic (negative Gaussian curvature) annular region between them (see Figure 5.5). This is known as the transversality principle [37]. This region gives rise to the matched concavities, which we used as indication for the part

decomposition in our 2-D analysis in Chapter 4. In an analogous way to our 2-D approach (see Chapter 4), we suggest the use of the parabolic curves (inflections of the Gaussian curvature) for performing the part decomposition. When considering more complex parts, such as tori, tubes or “snake like” shapes, there could be parabolic curves within the parts, and not only on the “glue” between them. Based on the classification of the parabolic curves we can still perform the decomposition.

5.3.1 From Image to Segmented Surface Description

Given the 3-D surface data, either complete triangulation data or range data from a single view, we perform the following steps to recover the graph description:

- First, we recover the differential geometric properties of the surface. We recover the derivatives and normals to the surface, the principal directions, and the sign of the Gaussian curvature. Note that we assume that the shapes are generic (we ignore singularities) and C^2 continuous (i.e. the curvature is defined everywhere). It has been shown that this process is stable and reliable [8, 71, 31, 33]. However, in the case of range data input, considerable amounts of smoothing are necessary as an initial step. For range images as input the differential properties are computed directly on the smoothed images. In Appendix B we discuss a method for estimating the Gaussian curvature in the case of triangular patches as input. In our current work, we ignore surfaces with zero Gaussian curvature. These are in general simpler surfaces which are quite well studied (e.g. [98]).
- Next, connected regions with the same Gaussian curvature sign are found, using a simple connected components algorithm [38]. Information about each region’s neighboring regions is also recorded. A region based graph is generated, where the nodes are the different connected regions, and the edges connect neighboring regions and represent the parabolic curves.

The graph recovered at this step is similar to graphs suggested by other researchers (e.g. [31, 66, 67, 34]). This graph is a surface based description and not a volumetric description. It is not at all clear how to make the leap from this surface based description to the desired volumetric description. Consider for example the case of the teapot shown in Figure 5.6. The original shaded range image is shown on the left, and the segmented regions, obtained as discussed above, are shown on the right. The problem of this description is that some regions have no “meaning” by themselves. For example regions 3, 4, and 5 are not independent entities, but belong to the same part, the spout. In a similar manner regions 1 and 2 belong to the handle. How does one make this distinction?

Figure 5.6: A shaded range image of a teapot (left) and the segmented regions (right).

5.3.2 From Surface to Volumetric Description

In the following steps we construct the part based volumetric graph from the surface graph obtained above.

- First, the parabolic curves are recovered from the borders between the positive and negative regions.
- We now classify the parabolic curves. As discussed in Section 5.2.3, we distinguish between two cases (Figure 5.4). In the first case, the lines of curvature, which inflect on the parabolic curve, are close to being orthogonal to the parabolic curve. This case, which is illustrated by the Bell shape in Section 5.2.3, suggests the existence of a part. In the second case, along a section of the parabolic curve, the inflecting lines of curvature, are close to being *parallel* to the parabolic curve. This case, which is illustrated by the Hump and Furrow shapes in Section 5.2.3, suggests that the bordering regions belong to the same part. The actual classification process is implemented as follows. First, the tangent direction to each point of the parabolic curve is computed. The difference between the tangent direction and the direction of the inflecting line of curvature at each point is computed. The point is labelled according to whether this difference is closer to 0 or to 90 degrees. We then filter the labels using a majority filter of width 5 to eliminate random noise effects. Finally, if after the filtering there is a section which is labelled as close to parallel,

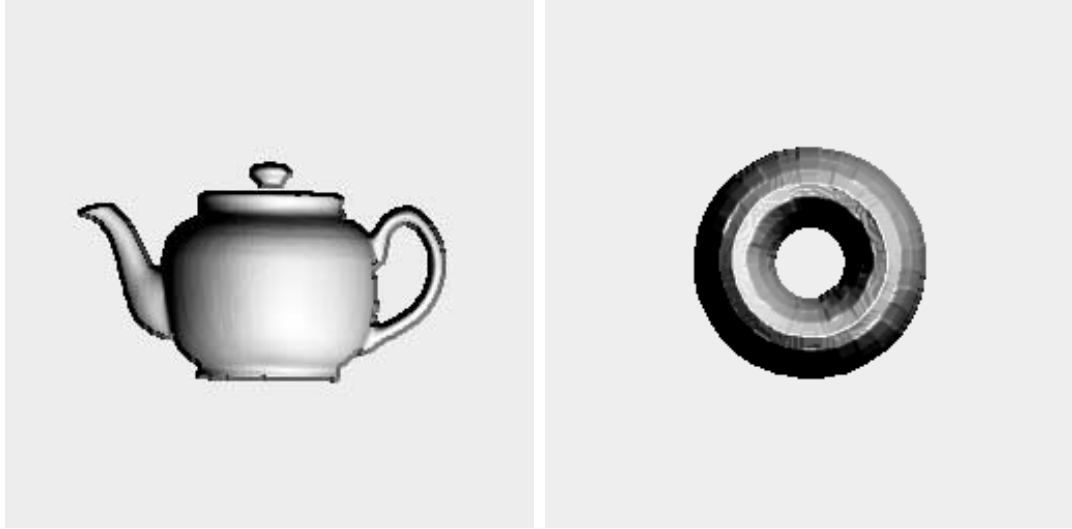


Figure 5.7: Shaded range images of teapot (left) and torus (right).

then the parabolic curve is labelled as “hump-like,” otherwise it is labelled as “part-like.”

- Finally, we generate the part based graph description from the region based graph, obtained earlier, in the following way: We combine regions which share a “hump-like” parabolic curve into a single node. These nodes are hypothesized as being “tube like.” Each negative curvature region connected to two nodes is hypothesized as being “glue.” The positive regions are hypothesized as being convex parts. Negative regions surrounded by a single “hump-like” parabolic curve are labelled as “furrows.”

Figures 5.7 through 5.12 follow the whole process on two examples, a teapot, shown on the left side, and a torus shown on the right. An important subtlety to note is that, although we use the term “torus” to name this object throughout this chapter, the input we have is not actually a torus (a mathematically defined shape), but a range image of a torus. Figure 5.7 shows the shaded range images of the teapot (shown earlier) and torus. This is the input to our system. Note that the knob on the lid of the teapot is actually disconnected making it an independent object. Figure 5.8 shows the images of the sign of the Gaussian curvature of the respective shapes. The borders between the regions are the parabolic curves. Based on the sign of the Gaussian curvature, connected regions are found as shown in Figure 5.9. The regions graphs obtained at this stage are depicted in Figure 5.10. Next, the parabolic curves are classified into “hump-like” and “part-like”, as discussed above. The parabolic curves and their classification are shown in Figure 5.11. Finally, the volumetric part based graphs of the teapot and torus are shown in Figure 5.12.

Figure 5.10: Region graph of teapot (left) and torus (right).

Figure 5.12: Final graph description of teapot (left) and torus (right).

5.4 Axial Description of Parts

In the previous section we have presented a method for recovering a segmented graph description of compound objects. This graph represents the different parts and their arrangement, and it can be used for indexing into a database of objects. It is clear, however, that for many applications this information may not be sufficient. For example, when trying to grasp an object it is important to have a description of the parts themselves. In this section we discuss the recovery of descriptions of some basic components.

The most simple components are convex blobs (or Ovoids, borrowing the terminology from Koenderink [51]). These could be represented in numerous ways (e.g. superquadrics [90]), and we do not address them further. The other type of components we are interested in are tubular structures with a straight or a curved axis. In

Figure 5.13: Example of a Generalized Cylinder.

the following, we discuss two classes of well defined shapes, SHGCs and PRCGCs. Both of these classes could be described naturally by their axis and cross sections. We present some known and novel properties of these classes of shapes, and show how based on these properties we can recover their axes and cross sections.

Unfortunately, not all parts are perfect SHGCs and PRCGCs. It is clear that these techniques could not be used to describe all possible parts. However, in practice, many parts are only slight deformations of the above classes and, therefore, these methods could still result in meaningful descriptions. In this context it is interesting to note the work on quasi-invariant properties (properties which hold for “most” viewpoints) of some classes of GCs [12, 104].

5.4.1 Straight Homogeneous Generalized Cylinders

A *Generalized Cylinder (GC)* [11, 65] is the solid obtained by sweeping a planar region (the cross section) along an arbitrary space curve (the axis) while scaling the cross section by an arbitrary scaling function (see Figure 5.13). A *Straight Homogeneous Generalized Cylinder (SHGC)* [88] is a GC with a straight axis and the scaling of the cross section is homogeneous. The cross section is usually, but not necessarily, assumed to be orthogonal to the axis. Figure 5.14 shows examples of SHGCs. Note that surfaces of revolution, which are well studied, are a special case of SHGCs where the cross sections are circular.

The properties of SHGCs have been studied by several researchers (e.g. [88, 72, 97, 104]). In the following, we first present some of the known properties of SHGCs, and

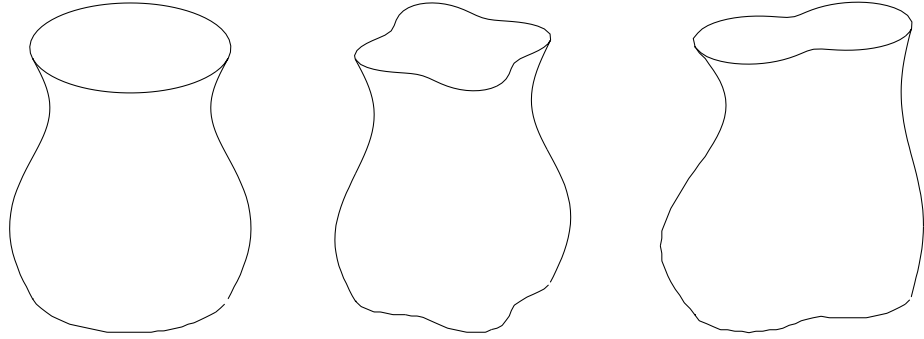


Figure 5.14: Examples of Straight Homogeneous Generalized Cylinders (adapted from [97]).

then show how these properties are used for recovering the axis and cross sections from the 3-D data.

Properties of SHGCs

Given a planar cross section $C(u) = (c_x(u), c_y(u), 0)^t$, and a scaling function $r(t)$, and assuming the axis is along the z axis of the coordinate system, the surface of an SHGC, S , can be parameterized as:

$$S(u, t) = (r(t)c_x(u), r(t)c_y(u), t) \quad (5.1)$$

The curves generated by fixing u and varying t are known as the *meridians* of the surface. The curves generated by fixing t and varying u are the cross sections, and are some time called the *parallels*. Refer to Figure 5.15 for an example. Note that both the meridians and the cross sections are planar curves. Note also that we assume that both $C(u)$ and $r(t)$, and therefore $S(u, t)$, are C^2 (continuous curvature).

Following are some known properties of SHGCs which we use for recovering the description. We present the properties without proof (the proofs can be found in the noted references):

Shafer [88] proves the following property, which relates the axis of an SHGC with the tangents to its surface in the direction of the axis.

Property 5.4.1 *For any two points on the same cross section of an SHGC, the tangent lines to the surface of the SHGC, in the direction of the axis, intersect at a common point on the axis (see Figure 5.16).*

Recall the definition of parallel symmetry from Chapter 4. Two curves are said to be *parallel symmetric* [96] if there exists a pointwise correspondence function between

Figure 5.16: The tangent lines on the same cross section in the direction of the axis intersect at a common point on the axis.

the two curves, such that the tangents to the curves at every two corresponding points are equal, and the function is continuous and monotonic. Ulupinar and Nevatia [97] have proved the following relationship between the cross sections of SHGCs.

Figure 5.17: The lines of correspondence between any pair of cross sections intersect at a common point on the axis.

Property 5.4.2 *The cross sections of an SHGC are parallel symmetric with each other with a linear correspondence function. The meridians of the SHGC join the corresponding parallel symmetric points.*

A property similar to Property 5.4.1, but often more useful, was suggested and proved by Zerroug and Nevatia [104].

Property 5.4.3 *The lines of correspondence (the lines connecting corresponding points) between any pair of cross sections are either parallel to the axis, or intersect at a common point on the axis (see Figure 5.17).*

Finally, the following important property, proved by Ponce *et al.* [72], relates the parabolic curves on the surface of an SHGC with the inflections of the cross section and scaling functions.

Property 5.4.4 *Parabolic curves on the surface of an SHGC are meridians, where the curvature of the cross section is zero, or parallels, where the curvature of the scaling function is zero.*

Recovering the axis and cross sections of an SHGC

We now discuss how the parabolic curves can be used for recovering the axis and cross sections of an SHGC. Given an SHGC, we first recover the parabolic curves on its surface. If it does not have parabolic curves then it is either simply a convex

blob, or it has zero Gaussian curvature everywhere (a linear scaling function). Both cases are not of interest here.

Based on Property 5.4.4, the parabolic curves are either cross sections or meridians of the SHGC. There are a few simple tests which can be done to classify the parabolic curves. For example, the cross section curves have to be closed (when considering complete data), where as the meridians cannot be closed. Also, from Property 5.4.2, the cross sections are parallel symmetric to each other, a property which is very easy to verify [85].

From the cross sections found we can recover the axis, using Property 5.4.3, by computing the intersection of their correspondence lines. If the SHGC may be oblique (the cross section not orthogonal to the axis), then an additional point on the axis is necessary, and it can be found in the same way using an additional cross section. Note, that if the SHGC is assumed to be orthogonal, and therefore the direction of the axis known, then Property 5.4.1 could also be used for recovering the axis.

In the case where the cross section sweeping function has no inflections, there are no parabolic curves which are cross sections. However, since the meridians are planar and the axis is in all the meridians planes, the axis can be recovered from the intersection of the meridian planes. If the SHGC is assumed to be orthogonal then the cross sections could also be recovered.

Once the axis, cross sections, and meridians are found, it is possible to verify the results by generating the SHGC and comparing with the original data. This may be important for determining the exact ends of the axis, and for eliminating possible hypotheses.

We demonstrate the reconstruction on the real range image of the teapot in Figure 5.18. The body of the teapot is an SHGC. The parabolic curves, which are also cross sections, were recovered and classified, as discussed in Section 5.3. A meridian was recovered by traversing the data in an orthogonal direction to a point on one of the cross sections (we assume an orthogonal SHGC, therefore, one point is sufficient). This meridian conveys the scaling of the sweeping function. We reconstruct a 3-D model of the SHGC by sweeping the cross section along the axis, while scaling it according to the scaling of the sweeping function. Figure 5.19 shows the cross sections and axis and the shaded image of the reconstructed 3-D teapot body. To illustrate the 3-D reconstruction, the results are shown at two novel orientations. Note that since the input is from only one view, we have no information about the “back” of the object. By making assumptions on the symmetrical nature of objects, one could infer a complete (and in this case correct) reconstruction.

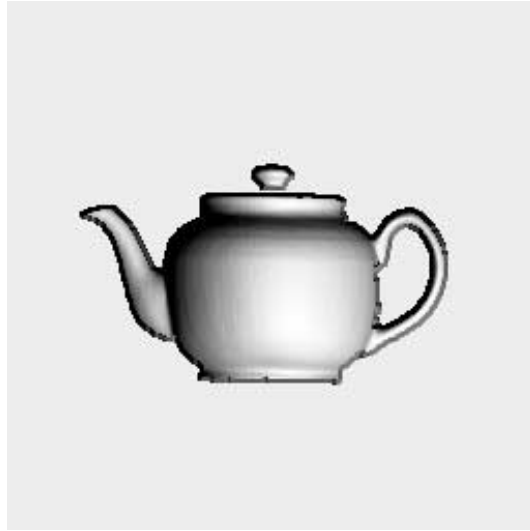


Figure 5.18: Shaded range image of a teapot.

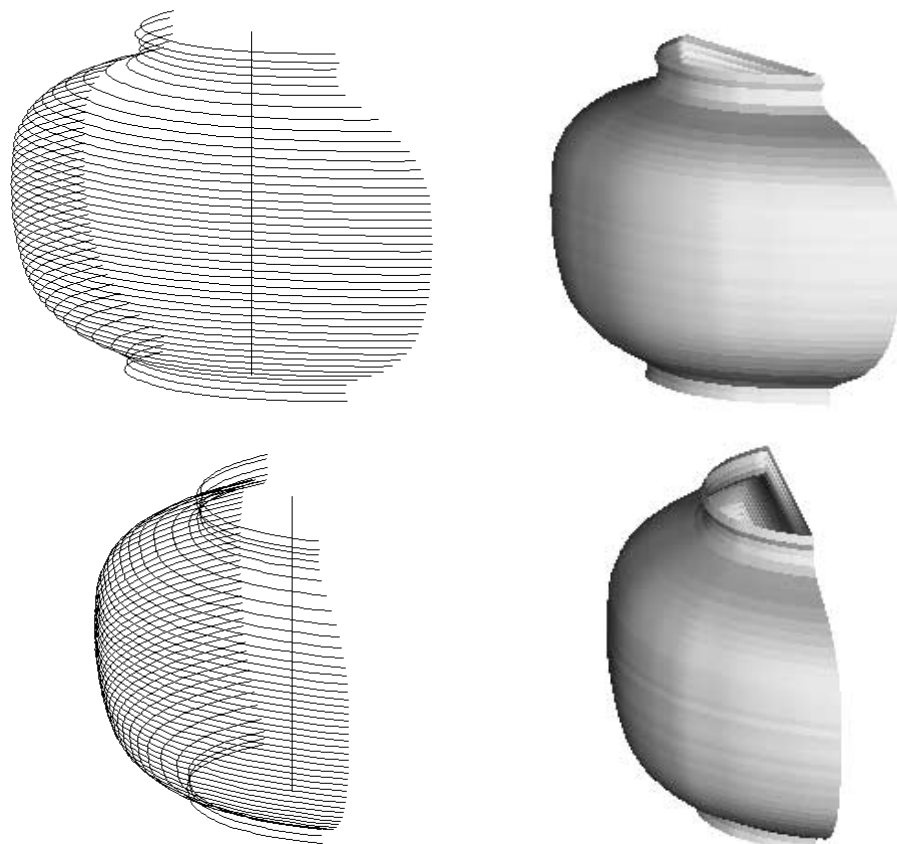


Figure 5.19: Cross sections and axis (left) and the shaded image (right) of reconstructed teapot body at two different orientations.

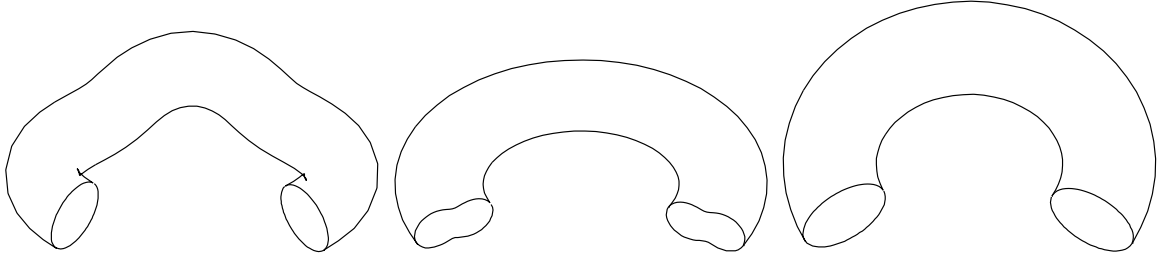


Figure 5.20: Examples of Planar Right Constant Generalized Cylinders (adapted from [97]).

5.4.2 Planar Right Constant Generalized Cylinders

A *Planar Right Constant Generalized Cylinder (PRCGC)* [98] is a GC with a planar axis, a constant cross section, and the cross section is orthogonal to the axis. Figure 5.20 shows examples of PRCGCs.

Although very useful in modeling “tube like” objects, PRCGCs (and CGCs) have received relatively little attention [88, 73, 98]. In the following, we first derive the formulation for the Gaussian curvature of PRCGCs. This will be used to derive two important properties of PRCGCs. As in the previous section, we then show how these properties are used for recovering the axis and cross sections from the 3-D data.

The Gaussian curvature of PRCGCs

Consider a PRCGC. Following the notation used in [98] (refer to Figure 5.21), we choose a coordinate system, such that the axis of the PRCGC lies in the $x - z$ plane, and one of the cross sections, $C(u) = (c_x(u), c_y(u), 0)^t$, is in the $x - y$ plane. We assume that $C(u)$ is parameterized by arc length, that is $\|C'(u)\|^2 = c_x'^2(u) + c_y'^2(u) = 1$ for all u . Let $A(t) = (a_x(t), 0, a_z(t))^t$ denote the axis, which is also parameterized by its arc length, that is, $\|A'(t)\|^2 = a_x'^2(t) + a_z'^2(t) = 1$ for all t . We also let $A(0) = (0, 0, 0)^t$, and since the cross section is orthogonal to the axis, $A'(0) = (0, 0, 1)^t$. The surface of the PRCGC, $S(u, t)$, is then given by:

$$S(u, t) = R(A'(0), A'(t)) \cdot C(u) + A(t) \quad (5.2)$$

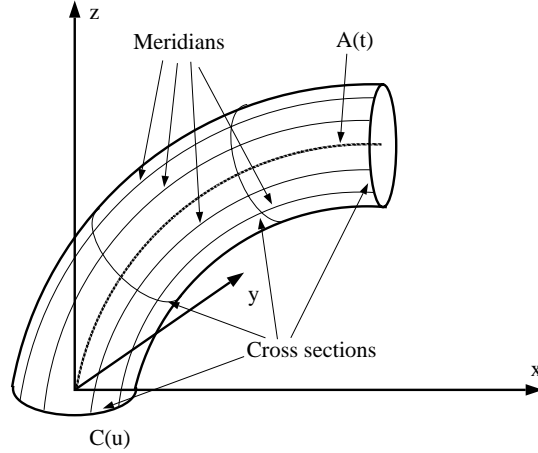


Figure 5.21: A PRCGC with some meridians and cross sections marked (adapted from [97]).

where $R(V_1, V_2)$ is the rotation matrix that transforms the vector V_1 into vector V_2 . For $A'(0) = (0, 0, 1)$ and $A'(t) = (a'_x(t), 0, a'_z(t))$ the rotation matrix R becomes:

$$R = \begin{bmatrix} a'_z(t) & 0 & a'_x(t) \\ 0 & 1 & 0 \\ -a'_x(t) & 0 & a'_z(t) \end{bmatrix}$$

Once again, the curves generated by fixing u and varying t are the *meridians* of the surface. The curves generated by fixing t and varying u are the *cross sections*.

We now turn to computing the differential properties of PRCGCs. For sake of brevity, we will omit some of the details. We also often omit the parameters (e.g. $c_x = c_x(u)$ and $a_x = a_x(t)$). In the following, let $K_A = K_A(t)$ denote the curvature of the axis, and let K_C denote the curvature of the cross section curve. From basic differential geometry [29, 58] we have the following identities for K_A and K_C

$$K_A = a''_x a'_z - a'_x a''_z \quad \text{and} \quad K_C = c''_x c'_y - c'_x c''_y \quad (5.3)$$

Expanding Equation 5.2 the surface, S , is given by

$$S = S(u, t) = \begin{pmatrix} a_x + a'_z c_x \\ c_y \\ a_z - a'_x c_x \end{pmatrix}$$

The partial derivatives of the surface with respect to u and t are then given by

$$S_u = \frac{\partial S}{\partial u} = \begin{pmatrix} a'_z c'_x \\ c'_y \\ -a'_x c'_x \end{pmatrix} \quad \text{and} \quad S_t = \frac{\partial S}{\partial t} = \begin{pmatrix} a'_x + a''_z c_x \\ 0 \\ a'_z - a''_x c_x \end{pmatrix}$$

and the second derivatives (we assume the surface is C^2) with respect to u and t are given by

$$S_{uu} = \frac{\partial^2 S}{\partial u^2} = \begin{pmatrix} a'_z c''_x \\ c''_y \\ -a'_x c''_x \end{pmatrix}, S_{ut} = \frac{\partial^2 S}{\partial t \partial u} = \begin{pmatrix} a''_z c'_x \\ 0 \\ -a''_x c'_x \end{pmatrix} \text{ and } S_{tt} = \frac{\partial^2 S}{\partial t^2} = \begin{pmatrix} a''_x + a'''_z c_x \\ 0 \\ a''_z - a'''_x c_x \end{pmatrix}$$

From S_u and S_t we can derive the normal, n , to the surface as

$$n = \frac{S_u \times S_t}{\|S_u \times S_t\|} = \frac{1}{|1 - K_A c_x|} \begin{pmatrix} a'_z c'_y - a''_x c_x c'_y \\ -(1 - K_A c_x) c'_x \\ -a'_x c'_y - a''_z c_x c'_y \end{pmatrix}$$

We can now calculate the Gaussian curvature of the surface. From standard differential geometry [58], the Gaussian curvature, K , is given by

$$K = \frac{LN - M^2}{EG - F^2} \quad (5.4)$$

where E, F, G and L, M, N are the coefficients of the first and second fundamental forms respectively and they are given by the following equations (omitting the details):

$$\begin{aligned} E &= S_u \cdot S_u = 1 \\ F &= S_u \cdot S_t = 0 \\ G &= S_t \cdot S_t = (1 - K_A c_x)^2 \\ L &= S_{uu} \cdot n = -K_C \\ M &= S_{ut} \cdot n = 0 \\ N &= S_{tt} \cdot n = -(1 - K_A c_x) K_A c'_y \end{aligned} \quad (5.5)$$

[These results were obtained by expanding the dot products and reorganizing the terms using basic differential geometric identities and properties of the geometry. The identities used are $K_A^2 = a''_x{}^2 + a''_z{}^2$, $K'_A = a'_z a'''_x - a'_x a'''_z$, $a'_x a''_x + a'_z a''_z = 0$, $K_A^2 = a'_x a'''_x + a'_z a'''_z$, and $K_A^3 = a'''_x a''_z - a''_x a'''_z$.]

Finally, from equations 5.4 and 5.5 the Gaussian curvature is given by

$$K = K(u, t) = -\frac{1}{\gamma(u, t)} K_A(t) K_C(u) c'_y(u) \quad (5.6)$$

where $\gamma(u, t) = 1 - K_A c_x$. Note that since we assume that the PRCGC does not intersect itself, γ is strictly positive.

Properties of PRCGCs

We can now prove a property, which relates the parabolic curves of a PRCGC to its axis, meridians, and cross section. This is a similar property to Property 5.4.4 for SHGCs. As in the case of SHGCs, and as shown later, it is very useful in recovering the axis and cross sections of a PRCGC.

Property 5.4.5 *Parabolic curves on the surface of a PRCGC are the cross sections where the curvature of the axis is zero, and they are meridians where either the curvature of the cross section is zero, or when the tangent to the cross section is parallel to the axis plane.*

Proof From Equation 5.6 it is clear that the Gaussian curvature of the surface, $K(u, t)$, of a PRCGC is zero if and only if

1. the curvature of the axis, $K_A(t)$, is zero for some t_0 , in which case the corresponding cross section, $S(u, t_0)$, is a parabolic curve, or,
2. the curvature of the cross section, $K_C(u)$, is zero for some u_0 , in which case the corresponding meridian $S(u_0, t)$ is a parabolic curve, or,
3. the tangent to the cross section is parallel to the axis plane, $c'_y(u_0) = 0$ for some u_0 , in which case the corresponding meridian $S(u_0, t)$ is a parabolic curve. \square

Based on the coefficients of the first and second fundamental forms in Equation 5.5 we can derive the following important property of PRCGCs. This property is well known for the specific case of the torus, but, to the best of our knowledge, was never stated for more general PRCGCs.

Property 5.4.6 *The meridians and cross sections of a PRCGC are its lines of curvature.*

Proof From standard differential geometry [29, 58] we know that the parameter curves in the neighborhood of a non-umbilical point are lines of curvature if and only if $F = M = 0$. Based on Equation 5.5 this condition holds for the above u and t parameterization of the PRCGC where, as noted, the u parameter curves are the cross sections and the t parameter curves are the meridians. \square

An additional useful property was proved by Ulupinar and Nevatia [98]. This property relates the meridians and cross sections in 3-D and we present it here without proof.

Property 5.4.7 *The meridians of a PRCGC are parallel symmetric, and the curves joining parallel symmetric points are the cross sections of the surface.*

[Note that Ulupinar and Nevatia have used this property to infer an important property of the projection of the limb edges of PRCGCs. Since we, however, are interested in the 3-D shape, we find this property to be useful in itself.]

It is interesting to note that a PRCGC with a circular axis is actually also a Surface of Revolution (SOR), reversing the role of the meridians and cross sections. This implies that the SORs are a subclass of PRCGCs, and therefore the above properties hold for them too (the fact that for SORs the lines of curvature are the meridians and cross sections is already known independently). In the previous section we have already mentioned that SORs are a subclass of SHGCs too.

Recovering the meridians and cross section of a PRCGC

The above properties suggest two methods for recovering the cross sections and the meridians of a PRCGC. One possibility, based on Property 5.4.6, is to compute the lines of curvature on the surface, which are the cross sections and meridians. To recover the cross section, it is actually sufficient to compute the principal directions at any point on the surface and cut the object along them. The other possibility, based on Property 5.4.5, is to recover the parabolic curves which are also either meridians or cross sections. Note that a PRCGC must always have at least two parabolic curves (unless it is the degenerate case of a straight axis and convex cross section). Currently we use the latter method. A comparison of the robustness of the two methods is needed.

As in the case of the SHGC, there are a few simple tests which can be done to classify the parabolic curves and the lines of curvature. For example, the cross section curves have to be closed (when considering complete data), where as the meridians are usually not closed (they are closed only if the axis is closed). Also, from Property 5.4.7, the meridians are parallel symmetric to each other. We note once again that this property is very easy to verify [85]. The parabolic curves of the third type from Property 5.4.5, which are also meridians, are easily distinguishable since the normals to the surface along this type of parabolic curve are parallel.

The cross section and meridians of a PRCGC give a complete description of it, since any curve parallel to the meridians could be the axis. We define the most intuitive axis to be the one passing through the center of the cross section (or center of gravity for a non-circular cross section).

Once the cross sections and meridians are found, it is possible to verify the results by generating the PRCGC and comparing with the original data. This may be important for determining the exact termination of the part, and for eliminating possible hypotheses.

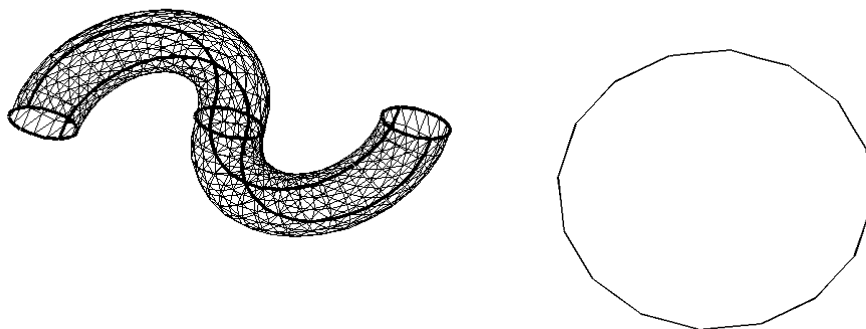


Figure 5.22: A PRCGC (left) and the recovered cross section (right). The parabolic curves (and the terminations) are shown emphasized.

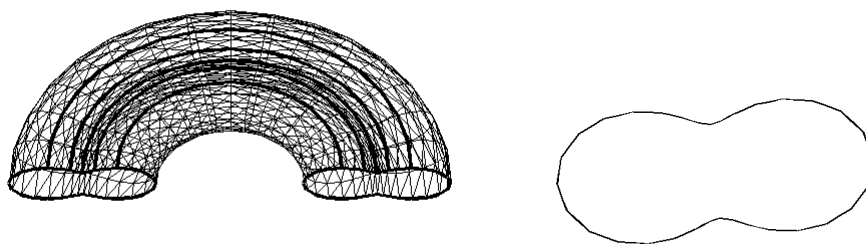


Figure 5.23: A PRCGC (left) and the recovered cross section (right). The top parabolic curves (and the terminations) are shown emphasized.

Figures 5.22 and 5.23 show results on two complete, synthetic PRCGCs. In both images the PRCGC is shown on the left. The parabolic curves and the terminations are emphasized with thicker lines. In the case of the PRCGC of Figure 5.23 only the top five parabolic curves are emphasized to avoid cluttering of the image. The remaining five parabolic curves are symmetrical to the ones shown. In both instances, the recovered cross-sections are shown on the right. Since these two shapes are synthetic, we were able to verify the results by comparing the recovered cross section and meridians with the ones used to generate the shapes originally. The results were perfect for both shapes.

Finally, we turn back to our examples of real range images of the teapot and torus (Figure 5.24). The torus and the handle of the teapot are nearly perfect PRCGCs.

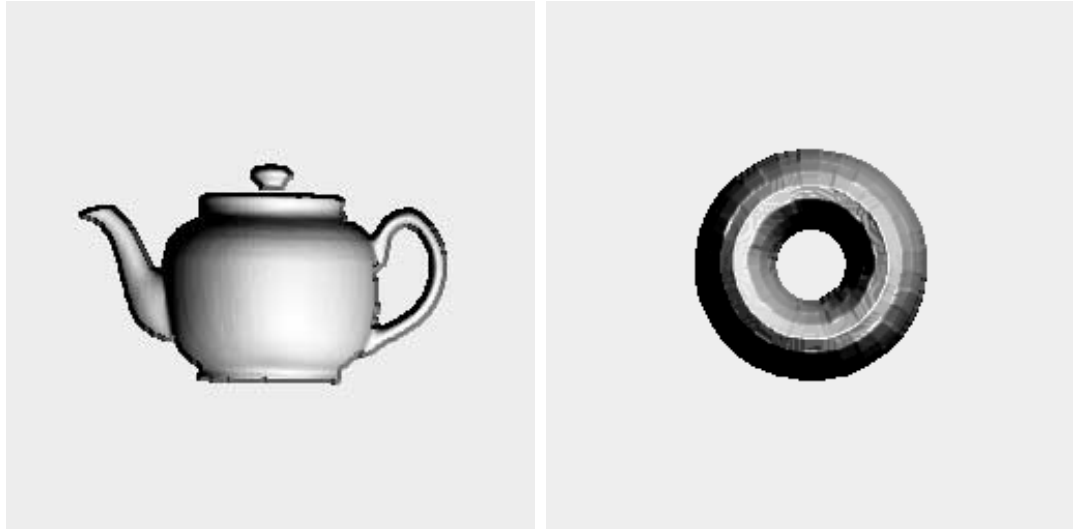


Figure 5.24: Shaded range images of teapot (left) and torus (right).

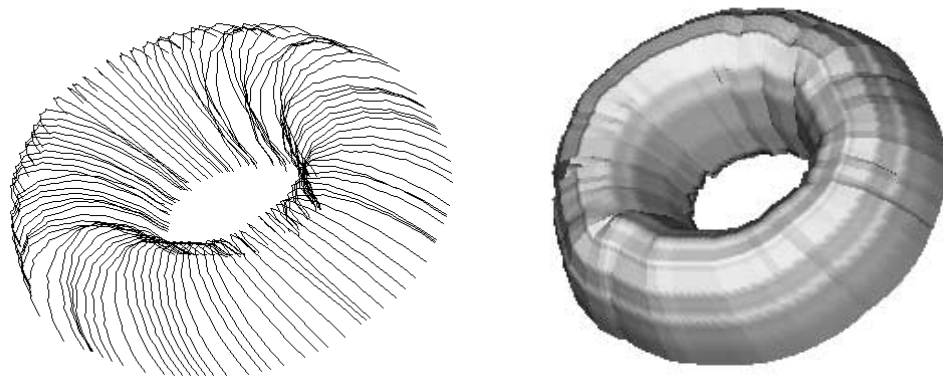


Figure 5.25: Cross sections and shaded image of reconstructed torus.

The parabolic curves were recovered and classified, as discussed in Section 5.3 (see Figure 5.11). A cross section is recovered by traversing the data in an orthogonal direction from a point on the meridian (the tangent orientations of the meridian are initially smoothed for robustness). Since, theoretically, all cross sections are identical, we currently recover only one cross section. For more robustness it is also possible to take several cross sections along the meridian. Finally, we reconstruct a 3-D model of the PRCGC by sweeping the cross section along the meridian. To illustrate the 3-D reconstruction we show the results at novel orientations. Figure 5.25 shows the cross sections and shaded image of the reconstructed 3-D torus, while Figure 5.26 shows the cross sections and shaded images of the reconstructed teapot handle at two different orientations. Note once again that since the input is from only one view, we have no information about the “back” of the objects. By making

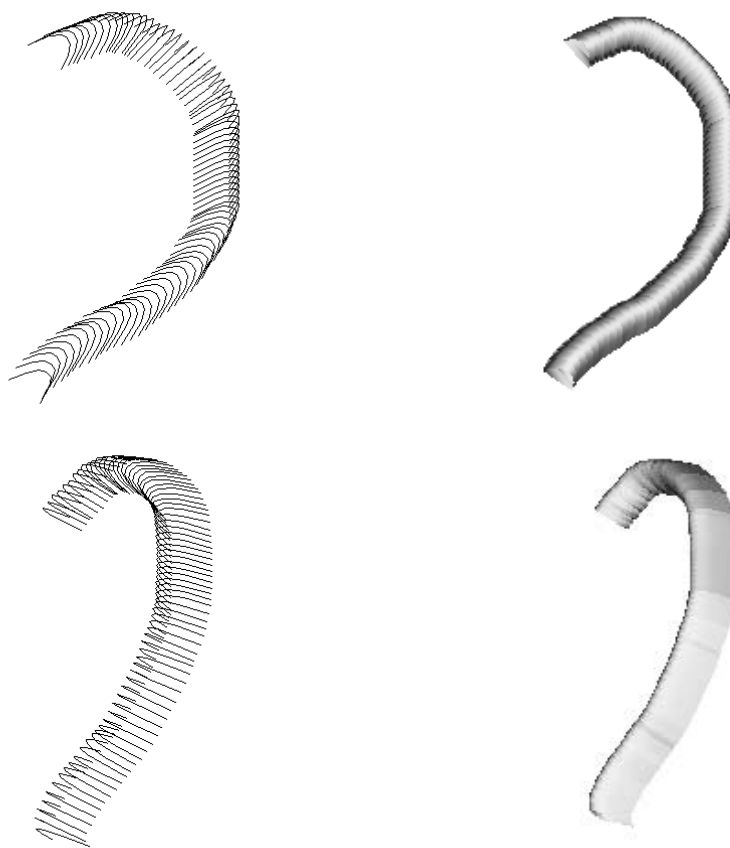


Figure 5.26: Cross sections and shaded image of reconstructed teapot handle at two different orientations.

assumptions on the symmetrical nature of the objects, one could infer a complete (and in these examples correct) reconstruction.

5.5 Concluding Remarks

We have addressed the problem of obtaining natural volumetric descriptions of three dimensional shapes. We address compound shapes where the components are connected smoothly. Parabolic curves are used for segmenting the shape into its components. We currently handle parts with tubular structure. A graph is generated in which the nodes represent the individual parts. This graph presents a structural description of the shape in terms of parts and their arrangement. We have also presented methods for recovering the axis, meridians and cross sections of SHGC and PRCGC components, based on known and novel mathematical properties.

We have presented some preliminary, but promising, results. In the following we discuss some limitations of our work and issues which require additional investigation.

The Figure-Ground problem As noted in Chapter 2 and in the conclusion to Chapter 4, one of the major problems in obtaining shape descriptions is the segmentation of the image into the object vs. the background. It is clear that for many realistic applications, this problem has to be solved prior or together with the process of recovering the shape description. In this work, however, as in the case of our 2-D approach, we have assumed that the figure-ground is given. In the concluding remarks to Chapter 4 we gave some arguments for the validity and importance of our approach even under the restrictive assumptions of given figure-ground. Most of these arguments hold even stronger in the 3-D case. It is clear, however that this problem has to be solved for achieving general purpose vision systems.

Further testing and data availability As noted above, we have presented some preliminary, but promising, results. Further testing is needed in order to validate the robustness of the approach. One of the problems we have encountered is the lack of complete 3-D data of complex objects. Most available data is either from single view range data or complete data of simple objects with no part structure. Research with promising results is currently underway by several researchers on the problem of recovering complete 3-D data from multiple range images. Once such systems are fully developed there could be an endless supply of testbeds for testing work on the description and recognition of 3-D objects.

Robustness An important requirement from a shape description is for it to be stable and robust. The robustness of our description relies on the reliable recovery of the parabolic curves. These can be recovered reliably from complete CAD models and high resolution range images. Unfortunately, the recovery of parabolic curves from noisy and low resolution range images is not as reliable. Further work is needed in this area.

Choosing between hypotheses The goal we have set for our system was to produce an intuitive description of a given shape. As in the 2-D case, a shape may have several acceptable interpretations. It is not clear how to determine which is the most reasonable or intuitive interpretation. Currently we produce one intuitive description. In order to facilitate consistency and stability of the interpretation under

deformations of the shape, and especially if the description is used for recognition, it would be useful for the system to be able to generate several, if not all, of the intuitive interpretations. This also brings up the issue of choosing an interpretation from several competing hypotheses for a single shape or part.

Verification of recovered description Another issue which we have not fully addressed is the verification of the recovered part description. Once the axis, cross sections, and meridians of SHGC and PRCGC parts are found, it is possible to verify the results by generating the SHGC or the PRCGC and comparing with the original data. This could be important for eliminating erroneous hypotheses. The most direct verification is to sum the distance between the points of the generated GC and their closest points of the original data. Higher level verifications, such as checking the coherence between different parts, could also be done.

Additional classes of shapes and surfaces Due to the enormity of the problem of recovering descriptions of 3-D shapes, we have limited the scope of this research. We handle generic smooth non-zero Gaussian curvature surfaces. We have also recovered axial descriptions of SHGCs and PRCGCs. There are many more classes of shapes and surfaces which still have to be addressed. These include planar surfaces, zero Gaussian curvature surfaces, discontinuities between surfaces, and additional classes of components. Some of these extension, such as allowing for discontinuities between parts, are relatively direct, while others probably require considerable effort.

Single view vs. complete data Our approach was targeted at describing complete 3-D objects. Due to the better availability of range images from single views, we have also developed tools to handle such data. However, when given a single view only partial descriptions can be recovered. We have not attempted to describe what is not visible in the given view, though assumptions, such as symmetry could be made. Obviously, the lack of complete data could result in incorrect descriptions in some cases. As discussed above, research is currently underway for recovering complete data from a set of registered range images.

Chapter 6

Conclusion

We have addressed the problem of recovering shape descriptions of planar and three dimensional shapes. Our underlying motivation was recovering descriptions, which will facilitate object recognition in the most general sense. The study of the requirements from such shape descriptions, lead to segmented descriptions, given in terms of the object's parts and their arrangement. This is also supported by research on human perception.

We started by studying planar shapes. We have suggested a method for producing axial descriptions of shapes together with hierarchical decompositions of the shapes into their parts. The approach was based on combining local and global information. The analysis of contour curvature combined with local symmetries allowed for reliable part segmentation and axial description.

Armed with a better understanding of the notion of a part, we set out to study three dimensional shapes. We have suggested a method for generating a volumetric segmented graph description of compound objects. The graph presents a structural description of the shape of the object in terms of its parts and their arrangement. This is one of the first attempts to address compound shapes, where the parts are connected smoothly. As input, we have considered 3-D complete data and range information from a single view. The segmentation is based on the analysis of the parabolic curves on the surface of the object. We have also studied some specific classes of parts and suggested natural descriptions of theses parts in terms of axes and cross sections.

Unfortunately, many issues were left unresolved. In the concluding remarks to Chapter 4 and Chapter 5, we have discussed some of the limitations of our approaches for 2-D and 3-D, respectively. We have also suggested some possible directions for future research. These issues include addressing the figure-ground problem, choosing between competing feasible hypotheses, further testing of our 3-D method on more objects, and extending the ideas to handle additional classes of shapes and surfaces. It is clear that many more years and effort are needed to reach the ultimate goal of

rivalling the capabilities of the human vision system. We hope, and believe, that we have taken a step, though very small, in the right direction.

Appendix A

Recovering SLS of B-splines

We describe here a method for recovering the Smooth Local Symmetries of B-splines. This method and some more details are also presented in [85]. Let $c_i(s) = (x_i(s), y_i(s))$, for $i = 1, 2$, be two parametric planar curves. Let $\vec{t}_i(s)$, or equivalently $\vec{t}_i(c_i(s))$, denote the *unit* vector in the direction of the tangent to the curve c_i at s . Finally, given two points, p_1 and p_2 , let $\vec{r}(p_1, p_2)$ denote the vector from p_1 to p_2 . If $p_1 = c_1(s_1)$ and $p_2 = c_2(s_2)$, two points on curves c_1 and c_2 respectively, form a local symmetry, then it follows from the definition of the local symmetry that

$$\vec{r}(p_1, p_2) \cdot (\vec{t}_2(s_2) - \vec{t}_1(s_1)) = 0 \quad (\text{A.1})$$

Note that we assume that the curves are parameterized (at least locally), such that the direction of both is the same (anti-clockwise) with respect to the circle, which is bitangent to the curves at the two symmetric points. This implies that $\vec{t}_2(s_2) - \vec{t}_1(s_1)$ is never zero.

Given two B-splines \mathcal{S}_1 and \mathcal{S}_2 ($\mathcal{S}_1 = \mathcal{S}_2$ for a closed contour), the SLS could be found by computing the SLS between every parabolic curve segment of \mathcal{S}_1 and every segment of \mathcal{S}_2 and then linking connected axes.

Unfortunately, we are unable to solve analytically for the SLS of two parabolic curve segments (solving Equation (A.1) for this case requires finding the roots of a polynomial of degree six). Therefore, given the two curve segments, rather than testing all possible pairs of points for local symmetry, we approximate the axis as follows:

Let $c_1(u) = (x_1(u), y_1(u))$ and $c_2(v) = (x_2(v), y_2(v))$ be the two parabolic curve segments, both defined on $[0, 1]$. Setting $u = 0$ we try to find $v \in [0, 1]$ such that $c_1(0)$ and $c_2(v)$ are locally symmetric. Using the notation of Equation (A.1) we define

$$f(v) = \vec{r}(c_1(0), c_2(v)) \cdot (\vec{t}_2(v) - \vec{t}_1(0))$$

By equation (A.1) local symmetries are also zero-crossings of f . If $f(a) \cdot f(b) < 0$, there exists at least one zero crossing of f between a and b and it can be found by

binary search. Since there may be multiple solutions, if $f(a) \cdot f(b) > 0$ we cannot be certain that there are no zero crossings. Even though, in practice these cases are rare, to reduce the chances of missing a solution, we first divide the interval into 10 subsections and then test the subsections for zero crossings.

Now we set $u = 1$ and repeat the search for a v such that $c_1(1)$ and $c_2(v)$ are locally symmetric. Finally, the whole process is repeated exchanging the roles of c_1 and c_2 . If there is an SLS ribbon describing c_1 and c_2 then, since in general the axis is smooth for smooth curves, the search produces two pairs of points, $(c_1(u_0), c_2(v_0))$ and $(c_1(u_1), c_2(v_1))$, which are locally symmetric and which are at the ends of the ribbon, that is either $u_0 = 0$ or $v_0 = 1$, and either $u_1 = 1$ or $v_1 = 0$.

Now that the ends of the ribbon are found, for every $u \in [u_0, u_1]$ the point locally symmetric to $c_1(u)$ is approximated by the point $c_2(g(u))$ where

$$g(u) = \frac{v_1 - v_0}{u_1 - u_0} \cdot u + \frac{v_0 u_1 - u_0 v_1}{u_1 - u_0}$$

Note that g is simply a linear interpolation in $u - v$ space. Our experiments show that this is a very good approximation to the true axis especially since the spline segments are very short. An alternative to this approach could be to trace the axis between the end points using a curve tracing algorithm (e.g. [52]).

Special care has to be taken not to overlook the case where the ribbon “crosses over”, that is when $u_0 < u_1$ and $v_0 < v_1$, using the above notation and the above assumption on the parameterization. This occurs whenever the radius of the circle bitangent to the points forming the local symmetry is larger than the radius of curvature at one of the points but smaller than the radius of curvature at the other point (see [35] for exact details and proof). Note that Brady and Asada [18], based on [16], have assumed that this cannot occur. However, the work in [16] was based on the *Symmetric Axis Transform (SAT)* [14] which requires the bitangent circle relating two symmetric points to be maximal within the shape, and therefore in this case the above condition cannot occur. On the other hand, with SLS there is no such limitation on the bitangent circle. Since we do not wish to limit the width of the ribbons, these “unnatural” ribbons are also produced.

The complexity of the algorithm for computing the SLS of the whole shape is $O(n_1 \cdot n_2)$, where n_i is the number of curve segments in S_i , the spline approximation of the curve c_i .

Figure B.1: Geometrical interpretation of the Gaussian curvature.

Figure B.2: Estimating Gaussian curvature on triangular patches. Top: positive K ; Bottom: negative K .

Given a triangular patch, we approximate the Gaussian curvature, K , by taking the ratio between the area of the triangle generated by the normals, and the area of the triangular patch. The sign of K is determined based on the relative orientation of the two triangles. Figure B.2 shows the case of positive K and the case of negative K , where the orientation reverses. If the solid angle (or the triangle on the Gauss map) vanishes, then K is zero.

Reference List

- [1] G. J. Agin and T. O. Binford. Computer description of curved objects. *IEEE Transactions on Computers*, 25, April 1976.
- [2] F. Attneave. Some informational aspects of visual perception. *Psychological Review*, 61:183–193, 1954.
- [3] D. H. Ballard. Generalizing the Hough transform to detect arbitrary shapes. *Pattern Recognition*, 13(2):111–122, 1981.
- [4] D. H. Ballard and C. Brown. *Computer Vision*. Prentice Hall, 1982.
- [5] H. G. Barrow and J. M. Tenenbaum. Interpreting line drawings as three dimensional surfaces. *Artificial Intelligence*, 17:75–116, 1981.
- [6] R. Bartels, J. Beatty, and B. Barsky. *An Introduction to Splines for use in Computer Graphics and Geometric Modeling*. Morgan Kaufmann Publishers, Inc., Los Altos, CA, 1987.
- [7] R. Bergevin and M. D. Levine. Part decomposition of objects from single view line drawings. *Computer Vision, Graphics and Image Processing*, 55:73–83, January 1992.
- [8] P. J. Besl and R. C. Jain. Invariant surface characteristics for 3D object recognition in range image. *Computer Vision, Graphics and Image Processing*, 33:33–80, 1986.
- [9] I. Biederman. Matching image edges to object memory. In *Proceedings of IEEE International Conference on Computer Vision*, pages 384–392, London, England, June 1987.
- [10] I. Biederman. Recognition by components. *Psychological Review*, 94:115–147, 1987.
- [11] T. O. Binford. Visual perception by computer. In *IEEE Conference on Systems and Controls*, Miami, FL, 1971.

- [12] T. O. Binford and T. S. Levitt. Quasi-invariants: Theory and exploitation. In *Proceedings of the DARPA Image Understanding Workshop*, pages 819–829, Washington, D.C., 1993.
- [13] A. Blake. Computational modelling of hand-eye coordination. *Proceedings of the Royal Society of London*, B(337):351–360, 1992.
- [14] H. Blum. *A Transformation for Extracting New Descriptors of Shape*. MIT Press, Cambridge, MA, 1967.
- [15] H. Blum and R. N. Nagel. Shape description using weighted symmetric axis features. *Pattern Recognition*, 10:167–180, 1978.
- [16] F. L. Bookstein. The Line-Skeleton. *Computer Graphics and Image Processing*, 11:123–137, 1979.
- [17] T. E. Boult and A. D. Gross. On the recovery of superellipsoids. In *Proceedings of the DARPA Image Understanding Workshop*, pages 1052–1063, Cambridge, MA, 1988.
- [18] M. Brady and H. Asada. Smoothed local symmetries and their implementation. *The International Journal of Robotics Research*, 3(3):36–61, 1984.
- [19] R. A. Brooks. Symbolic reasoning among 3-D models and 2-D images. *Artificial Intelligence*, 17:285–348, 1981.
- [20] R. A. Brooks. Model-based three dimensional interpretations of two dimensional images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 5(2):140–150, 1983.
- [21] J. F. Canny. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(6):679–698, 1986.
- [22] Y. Chen and G. Medioni. Object Modelling by Registration of Multiple Range Images. *International Journal of Image and Vision Computing*, 10(3):145–155, April 1992.
- [23] Y. Chen and G. Medioni. Surface level integration of multiple range images. In *Proceedings of the Workshop on Computer Vision in Space Applications*, Antibes, France, September 1993.
- [24] R. C. K. Chung and R. Nevatia. Use of monocular groupings and occlusion analysis in a hierarchical stereo system. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pages 50–56, Maui, Hawaii, 1991.
- [25] M. B. Clowes. On seeing things. *Artificial Intelligence*, 2(1):79–116, 1971.

- [26] J. H. Connell. Learning shape descriptions: Generating and generalizing models of visual objects. Technical Report AI-TR 853, Massachusetts Institute of Technology Artificial Intelligence Laboratory, 1985.
- [27] J. H. Connell and M. Brady. Generating and generalizing models of visual objects. *Artificial Intelligence*, 31(2):159–183, 1987.
- [28] S. J. Dickinson, A. P. Pentland, and A. Rosenfeld. 3-D shape recovery using distributed aspect matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):173–199, February 1992.
- [29] M. P. Do Carmo. *Differential Geometry of Curves and Surfaces*. Prentice-Hall, Inc., 1976.
- [30] S. A. Dudani, K. J. Breeding, and R. B. McGhee. Aircraft identification by moment invariants. *IEEE Transactions on Computers*, 26(1):39–46, 1977.
- [31] T. J. Fan, G. Medioni, and R. Nevatia. Recognizing 3-D objects using surface descriptions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(11):1140–1157, 1989.
- [32] F. P. Ferrie, J. Lagarde, and P. Whaite. Recovery of volumetric object descriptions from laser rangefinder images. In *Proceedings of European Conference on Computer Vision*, pages 387–396, Antibes, France, 1990.
- [33] P. J. Flynn and A. K. Jain. On reliable curvature estimation. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pages 110–116, San Diego, CA, June 1989.
- [34] P. J. Flynn and A. K. Jain. 3D object recognition using invariant feature indexing of interpretation tables. In *IEEE Workshop on Directions in Automated CAD-Based Vision*, pages 115–123, Maui, Hawaii, June 1991.
- [35] P. J. Giblin and S. A. Brassett. Local symmetry of plane curves. *American Mathematical Monthly*, 92:689–707, 1985.
- [36] K. Y. Goldberg. Orienting polygonal parts without sensors. *Algorithmica*, 10:201–225, 1993.
- [37] V. Guillemin and A. Pollack. *Differential Topology*. Prentice Hall, 1974.
- [38] R. M. Haralick and L. G. Shapiro. *Computer and Robot Vision*. Addison Wesley, 1993.
- [39] D. D. Hoffman. *Representing Shapes for Visual Recognition*. PhD thesis, Massachusetts Institute of Technology, May 1983.

- [40] D. D. Hoffman and W. A. Richards. Parts of recognition. *Cognition*, 18:65–96, 1985.
- [41] R. Horaud and M. Brady. On the geometric interpretation of image contours. *Artificial Intelligence*, 37:333–353, 1988.
- [42] B. K. P. Horn. Extended Gaussian images. *Proceedings of the IEEE*, 72:1656–1678, 1984.
- [43] B.K.P. Horn. *Robot Vision*. MIT Press, Cambridge, MA, 1986.
- [44] D. A. Huffman. Impossible objects as nonsense sentences. *Machine Intelligence*, 6:295–323, 1971.
- [45] J. E. Hummel and I. Biederman. Dynamic binding in a neural network for shape recognition. *Psychological Review*, 99:480–517, 1992.
- [46] K. Ikeuchi. Recognition of 3-D objects using the Extended Gaussian Image. In *Proceedings of International Joint Conference on Artificial Intelligence*, pages 595–600, August 1981.
- [47] C. L. Jackins and S. L. Tanimoto. Oct-trees and their use in representing three-dimensional objects. *Computer Graphics and Image Processing*, 14(4):249–270, November 1980.
- [48] T. Kanade. Recovery of the three-dimensional shape of an object from a single view. *Artificial Intelligence*, 17:409–460, 1981.
- [49] B. B. Kimia, A. Tannenbaum, and S. W. Zucker. Toward a computational theory of shape: An overview. In *Proceedings of European Conference on Computer Vision*, pages 402–407, Antibes, France, 1990.
- [50] J. J. Koenderink. What does the occluding contour tell us about solid shape. *Perception*, 13:321–330, 1984.
- [51] J. J. Koenderink. *Solid Shape*. MIT Press, 1990.
- [52] D. J. Kriegman and J. Ponce. A new curve tracing algorithm and some applications. In P. J. Laurent, A. Le Mehaute, and L. L. Schumaker, editors, *Curves and Surfaces*, pages 1–4. Academic Press, New York, 1990.
- [53] D. J. Kriegman and J. Ponce. On recognizing and positioning curved 3-D objects from image contours. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12:1127–1137, 1990.
- [54] M. Leyton. Symmetry-Curvature duality. *Computer Vision, Graphics and Image Processing*, 38:327–341, 1987.

- [55] M. Leyton. A process-grammar for shape. *Artificial Intelligence*, 34:213–247, 1988.
- [56] M. Leyton. Inferring causal history from shape. *Cognitive Science*, 13:357–387, 1989.
- [57] H. S. Lim and T. O. Binford. Structural correspondence in stereo vision. In *Proceedings of the DARPA Image Understanding Workshop*, pages 794–808, Los Angeles, California, 1987.
- [58] M. M. Lipschutz. *Differential Geometry*. Schaum’s Outline Series. McGraw-Hill, 1969.
- [59] D. G. Lowe. *Perceptual Organization and Visual Recognition*. Kluwer Academic Publishers, Hingham, MA, 1985.
- [60] A. K. Mackworth. Interpreting pictures of polyhedral scenes. *Artificial Intelligence*, 4:121–137, 1973.
- [61] D. Marr. *Vision*. W. H. Freeman and Co., San-Fransisco, CA, 1982.
- [62] D. Marr and K. Nishihara. Representation and recognition of the spatial organization of three-dimensional shapes. *Proceedings of the Royal Society of London*, B(200):269–294, 1977.
- [63] L. R. Nackman and S. M. Pizer. Three-dimensional shape description using the Symmetric Axis Transform. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 7(2):187–202, march 1985.
- [64] R. Nevatia. *Machine Perception*. Prentice Hall, 1982.
- [65] R. Nevatia and T. O. Binford. Description and recognition of complex-curved objects. *Artificial Intelligence*, 8:77–98, 1977.
- [66] B. Parvin and G. Medioni. A dynamic system for object description and correspondence. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pages 393–399, Maui, Hawaii, June 1991.
- [67] B. A. Parvin. *A Dynamic System for Object Description and Correspondence*. PhD thesis, University of Southern California, July 1991. IRIS Technical Report 286.
- [68] A. P. Pentland. Perceptual organization and the representation of natural form. *Artificial Intelligence*, 28:293–331, 1986.
- [69] A. P. Pentland. Recognition by parts. In *Proceedings of IEEE International Conference on Computer Vision*, pages 612–620, London, England, June 1987.

- [70] J. Ponce. On characterizing ribbons and finding skewed symmetries. *Computer Vision, Graphics and Image Processing*, 52:328–340, 1990.
- [71] J. Ponce and M. Brady. Towards a surface primal sketch. In T. Kanade, editor, *Three Dimensional Machine Vision*, pages 195–239. Kluwer Academic, New York, 1987.
- [72] J. Ponce, D. Chelberg, and W. B. Mann. Invariant Properties of Straight Homogeneous Generalized Cylinders and Their Contours. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(9):951–966, 1989.
- [73] K. Rao. *Shape Description from Sparse and Imperfect Data*. PhD thesis, University of Southern California, December 1988. IRIS Technical Report 250.
- [74] K. Rao and R. Nevatia. Computing volume descriptions from sparse 3-D data. *International Journal of Computer Vision*, 2:33–50, 1988.
- [75] K. Rao and R. Nevatia. Descriptions of complex objects from incomplete and imperfect data. In *Proceedings of the DARPA Image Understanding Workshop*, pages 399–414, Palo Alto, CA, May 1989.
- [76] A. A. G. Requicha. Representations for rigid solids: Theory, Methods, and Systems. *ACM Computing Surveys*, 12(4):437–464, 1980.
- [77] W. A. Richards and D. D. Hoffman. Codon constraints on closed 2D shapes. *Computer Vision, Graphics and Image Processing*, 31:265–281, 1985.
- [78] H. Rom and G. Medioni. Hierarchical decomposition and axial representation of 2D shape. In *Proceedings of the 8th Israeli Symposium on Artificial Intelligence and Computer Vision*, pages 261–271, Ramat Gan, Israel, December 1991.
- [79] H. Rom and G. Medioni. Hierarchical decomposition and axial representation of shape. In *Proceedings of the SPIE conference on Geometric Methods in Computer Vision*, San Diego, California, July 1991.
- [80] H. Rom and G. Medioni. Hierarchical decomposition and axial shape description. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pages 49–55, Champaign, Illinois, June 1992.
- [81] H. Rom and G. Medioni. Hierarchical decomposition and axial shape description. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(10):973–981, 1993.
- [82] A. Rosenfeld. Axial representations of shape. *Computer Vision, Graphics and Image Processing*, 33:156–173, 1986.

- [83] P. Saint-Marc, J. S. Chen, and G. Medioni. Adaptive smoothing: A general tool for early vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(6):514–529, 1991.
- [84] P. Saint-Marc and G. Medioni. B-spline contour representation and symmetry detection. In *Proceedings of European Conference on Computer Vision*, pages 604–606, Antibes, France, 1990.
- [85] P. Saint-Marc, H. Rom, and G. Medioni. B-spline contour representation and symmetry detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(11):1191–1197, 1993.
- [86] H. Samet. The quadtree and related hierarchical data structures. *ACM Computing Surveys*, 16(2):187–260, 1984.
- [87] Y. Sato, J. Ohya, and K. Ishii. Smoothed local generalized cones: An axial representation of 3D shapes. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pages 56–62, Champaign, Illinois, June 1992.
- [88] S. A. Shafer. Shadow geometry and occluding contours of generalized cylinders. Technical Report CMU-CS-83-131, Carnegie-Mellon University, May 1983.
- [89] K. Siddiqi and B. B. Kimia. Parts of visual form: Computational aspects. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pages 75–81, New York, NY, June 1993.
- [90] F. Solina and R. Bajcsy. Recovery of parametric models from range images: The case for superquadrics with global deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(2):131–147, 1990.
- [91] M. Soucy and D. Laurendeau. Multi-resolution surface modeling from range views. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pages 348–353, Urbana-Champaign, IL, June 1992.
- [92] F. Stein and G. Medioni. Structural Indexing: Efficient Three Dimensional Object Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):125–146, February 1992.
- [93] K. A. Stevens. The visual interpretations of surface contours. *Artificial Intelligence*, 17:47–73, 1981.
- [94] J. B. Subirana-Vilanova. The skeleton sketch: Finding salient frames of reference. In *Proceedings of the DARPA Image Understanding Workshop*, pages 614–622, Pittsburgh, PA, September 1990.

- [95] D. Terzopoulos, A. Witkin, and M. Kass. Symmetry-seeking models for 3D object reconstruction. In *Proceedings of IEEE International Conference on Computer Vision*, pages 269–276, London, England, June 1987.
- [96] F. Ulupinar and R. Nevatia. Inferring shape from contour for curved surfaces. In *Proceedings of International Conference on Pattern Recognition*, pages 147–154, Atlantic City, NJ, 1990.
- [97] F. Ulupinar and R. Nevatia. Shape from contour: SHGCs. In *Proceedings of IEEE International Conference on Computer Vision*, pages 582–582, Osaka, Japan, 1990.
- [98] F. Ulupinar and R. Nevatia. Recovering shape from contour for constant cross section generalized cylinders. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pages 674–676, Maui, Hawaii, 1991.
- [99] H. B. Voelcker and A. A. G. Requicha. Geometrical modeling of mechanical parts and processes. *IEEE Computer*, 10(12):48–57, 1977.
- [100] G. Xu and S. Tsuji. Inferring surfaces from boundaries. In *Proceedings of IEEE International Conference on Computer Vision*, pages 716–720, London, England, 1987.
- [101] C. T. Zahn and R. Z. Roskies. Fourier descriptors for plane closed curves. *IEEE Transactions on Computers*, 21:269–281, 1972.
- [102] M. Zerroug and R. Nevatia. Volumetric descriptions of SHGCs from a single intensity image. Technical Report IRIS-301, University of Southern California, Los Angeles, California, May 1992.
- [103] M. Zerroug and R. Nevatia. Quasi-invariant properties and 3-D shape recovery of non-straight, non-constant generalized cylinders. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pages 96–103, New York, NY, June 1993.
- [104] M. Zerroug and R. Nevatia. Volumetric descriptions from a single intensity image. *International Journal of Computer Vision*, 1993. (To appear).