

**REPRESENTATION AND DEPICTION  
OF 2D SHAPES USING PARTS**

**BY XIAOFENG MI**

**A dissertation submitted to the  
Graduate School—New Brunswick  
Rutgers, The State University of New Jersey  
in partial fulfillment of the requirements  
for the degree of  
Doctor of Philosophy  
Graduate Program in Computer Science**

**Written under the direction of**

**Doug DeCarlo**

**and approved by**

---

---

---

---

---

**New Brunswick, New Jersey**

**October, 2010**

## **ABSTRACT OF THE DISSERTATION**

### **Representation and Depiction of 2D Shapes using Parts**

**by Xiaofeng Mi**

**Dissertation Director: Doug DeCarlo**

We describe a 2D shape abstraction system that aims to clarify the structure without loss of the expressiveness of the original shape. To do this, traditional approaches in computer graphics typically use simplification techniques based on local adjustments of vertices, edges and faces. However, this thesis argues that an effective depiction can benefit from a computational representation compatible with a human's understanding of the shape. To support this argument, we first present a system that parses a 2D planar shape into a part-based structure that approximately respects the structural organization in human perception. Then we show that simplifications of this representation align with the common artistic practices in shape abstraction, in which only prominent parts are preserved and the visual structures are more clarified than using traditional simplification methods based on local adaptation of geometric details.

To compute the part structure of a given shape, we first propose that a part connects to the rest of the shape via its transition. Then we use a simple but general model to mathematically locate and describe this transition. We show that our model generalizes previously discovered theories on how the human visual system parses a shape into parts. It also provides a plausible way of explanatory shape analysis that requires clean pruning of parts without leaving attachment traces. Following insights from cognitive science, we have designed a set of heuristics to

resolve ambiguities in the representation of the shape. A set of stroke-based tools is designed so that the user can interact with the system to guide the shape analysis as well as to evaluate and optimize the performance of the parser.

Geometric thresholds and part selection tools are provided for the user to specify a subset of the part structure computed from the above step. The abstraction is done by simply reconstructing the shape from this subset. The reconstruction respects the geometric properties of the original part attachment and allows topological alternations of the structure resulting from elimination of less salient parts, which greatly improves the flexibility in the reconstruction. The perceptual study we have conducted confirms that human subjects indeed prefer our abstractions over the traditional 2D shape simplifications by Douglas-Peucker or Progressive Meshes, both of which try to approximate certain geometric properties during the simplifications.

## Acknowledgements

Working toward a PhD degree took me more than 6 years. Granted that research involves lots of fun, it is also accompanied by occasional frustrations and temptations to quit. I could by no means have gone through this process successfully without the support of many people.

I owe sincere gratitude to my advisor, Dr. Doug DeCarlo, for his continuous intellectual and financial support, his motivation, patience and understanding. Doug has actively been exploring how accounts of human perception can inform computer graphics systems for more natural and effective visual presentation. His insight in this area led directly to the high-level ideas behind the research presented in this thesis. I am grateful for his guidance in all aspects of this research, from topic selection to literature review, system implementation and thesis writing. I am particularly grateful for his step-by-step guidance on how to conduct research, how to organize and present ideas, how to write code in the right way and how to evaluate and debug programs to help understand our research models. (Unfortunately, I did not fully appreciate some of his advice until I began working at Adobe as an intern.) With his diligence, enthusiasm and integrity Doug set a perfect role model in front of me as a researcher. I feel so lucky to have had Doug as my advisor for the past 6 years.

I would also like to express my deep appreciation to Dr. Manish Singh. I got to know Manish in his seminar on shape understanding and spent another two semesters of independent study with him, from which I gained systematic knowledge of human vision and learned skills to design and run perceptual studies with human subjects. All of these have been essential to the completion of my dissertation research.

I am equally grateful to the rest of my thesis committee: Dr. Matthew Stone, Dr. Andrew Nealen and Dr. Adam Finkelstein. They have provided insightful comments and constructive criticisms at several stages of this research. Their comments were thought-provoking and

helped me flesh out the theories and approaches to the problem of part-based understanding and depiction. Drs. Stone, Nealen and Finkelstein were also amazingly considerate, squeezing space into their tight schedules to make it possible for my defense to happen before it was too late to keep my immigration status. I still remember that the first lecture I took in this country was given by Matthew. It was his artificial intelligence class that taught me how to come to the point in answering questions. His seminars were always interactive and helpful to graduate students in our academic life. His wisdom in language arts has been of great help in writing my publications and is very much appreciated.

I am indebted to many professors at Rutgers University: Dr. Ahmed Elgammal for educating me in foundational theories and methods in computer vision and machine learning; Dr. Peter Meer for showing me how to construct a robust computer vision system. Dr. Bill Steiger's amazing ability to convey concepts has been particularly impressive, and his computational geometry course, which I attended, has been most helpful to my research on smoothed local symmetries.

I am also indebted to my friends who volunteered to run the pilot sessions or served as human subjects in the perceptual study: Shu Chen, Yansong Chen, Lu Han, Junzhou Huang, Ray Huang, Ming Jin, Mingshun Jin, Luying Li, Baiyang Liu, Xueqing Liu, Fei Song, Alina Vamanu, Iulian Vamanu, Yi Xu. Wei Zheng, Guys, thank you so much!

Thanks also go to the department's graduate secretary, Ms. Carol DiFrancesco, and the payroll specialist, Ms. Aneta Unrat, for being always helpful and friendly to me. I am happy to have had Anthony Santella and Kevin Sanik as my lab mates: they made the VILLAGE lab a convivial place to work. I've also benefited considerably from our inspiring discussions about our research. Thanks, Anthony and Kevin!

This thesis was structured when I was working at Adobe as an intern in the second half of 2009. It was an outstanding personal and professional learning experience working with my mentors: Mr. Pete Falco from the Photoshop team and Dr. Sunil Hadap from the Advanced Technology Lab, as well as my manager, Mr. Dave Hackle. I am grateful to have had a chance to work with this prestigious team, and lunches with their computer scientists, Pushkar Joshi, Xinju Li, Yuyan Song and Nikolai Svakhin have given me joyful and unforgettable memories.

During the writing of this thesis, I have got a lot of editing help from Ms. Ellen Kanarek. Ellen did a great job, without which it was not possible for me to complete my dissertation in a timely manner. Thanks Ellen.

I regret that I was not able to spend enough time with my wife, Ying Zhang, during the thesis process. Here, let me express my full appreciation to Ying for her love, for her willingness to forgive, and for spending so much time in taking up the burden of doing housework and preparing delicious food, despite also being busy with her own PhD studies. I give my deepest gratitude to my parents, for their unflagging love and care throughout my life. They spared no effort to provide me with the best possible environment for growing up. As farmers from rural China, my parents had to work overtime and sometimes take multiple jobs to support my education. Watching their only child going abroad and leaving them alone, my parents never uttered a single word of complaint. Instead, they have always been encouraging of my research and have tried to present the best parts of their life to me in our phone calls so that I would not get distracted. I am very proud of my parents and no words could ever express my gratitude towards them.

# Table of Contents

<b>Abstract</b> . . . . .	ii
<b>Acknowledgements</b> . . . . .	iv
<b>1. Representation and Depiction</b> . . . . .	1
1.1. Shape, representation and depiction . . . . .	3
1.1.1. Structural representation in terms of parts . . . . .	4
1.1.2. Art as structural abstraction . . . . .	5
1.2. Related work . . . . .	9
1.2.1. Locally reflective symmetry and axial representation . . . . .	9
1.2.2. Medial axis transformation . . . . .	12
1.2.3. Structural pruning and robust medial representation . . . . .	14
1.2.4. Approximate symmetry . . . . .	15
1.2.5. Geometry simplification . . . . .	21
1.3. Non-photorealistic rendering . . . . .	23
<b>2. Identifying Parts using Contour Grouping Strength</b> . . . . .	26
2.1. Locating parts: existing approaches . . . . .	26
2.1.1. Medial axis revisited . . . . .	26
2.1.2. Boundary representations . . . . .	28
2.2. Symmetry in shape description and part separation . . . . .	30
2.2.1. Transitional region between parts . . . . .	31
2.2.2. Ambiguous shape explanation . . . . .	32
2.3. Invariant of that basic part with reflective symmetry . . . . .	34
2.4. The basic part with smoothed local symmetries . . . . .	37
2.4.1. Geometry of local symmetry . . . . .	37

2.4.2.	Tracing the smoothed local symmetries . . . . .	39
2.4.3.	Whole SLS vs SLS of basic parts . . . . .	43
2.5.	Locating the part body and transition . . . . .	44
2.5.1.	Chordal geometry . . . . .	44
2.5.2.	Geometry along smoothed local symmetries . . . . .	45
2.5.3.	Contour grouping strength and part transition . . . . .	49
2.5.4.	Locating part transitions using relatability . . . . .	51
2.6.	Extrapolation of part body using symmetry . . . . .	56
2.7.	Discussion . . . . .	58
2.A.	Derivations of chordal geometry . . . . .	59
2.A.1.	Derivatives of $L$ . . . . .	59
2.A.2.	Derivatives of $\theta$ and $\varphi$ . . . . .	59
2.B.	Curvature estimation and symmetry-curvature duality on polygons . . . . .	61
<b>3.</b>	<b>Shape Parsing in terms of Parts . . . . .</b>	<b>62</b>
3.1.	Shape analysis by sequential part pruning . . . . .	63
3.1.1.	Related work on structural pruning . . . . .	63
3.1.2.	Shape analysis with clean removal of parts . . . . .	66
3.1.3.	Termination condition for SLS tracing . . . . .	70
3.1.4.	Results . . . . .	72
3.2.	Valid and invalid deletions . . . . .	73
3.2.1.	Background parts . . . . .	74
3.2.2.	Part scale, area and contour grouping . . . . .	74
3.3.	Heuristics to prevent symmetry breaking . . . . .	77
3.3.1.	Prevent distortion of symmetry of neighboring parts . . . . .	77
3.3.2.	Prevent distortion of any salient axial symmetry . . . . .	79
3.3.3.	Measuring symmetry distortion . . . . .	80
3.3.4.	Deletions that improve symmetry . . . . .	81
3.3.5.	Preventing breaking of inter-part symmetries . . . . .	81
3.4.	Heuristics to avoid premature axis truncation . . . . .	82

3.4.1. Discussion . . . . .	84
3.5. Shape analysis by sequential constrained part pruning . . . . .	85
3.6. User guided shape analysis and performance assessment . . . . .	88
3.6.1. User guidance . . . . .	90
3.6.2. Assessing thresholds . . . . .	92
3.A. Locating global chordal symmetries . . . . .	95
3.B. Discovering salient axial structures with CDT . . . . .	97
<b>4. Abstracted Depiction in terms of Parts . . . . .</b>	<b>101</b>
4.1. Part attachment and shape reconstruction . . . . .	102
4.1.1. Parametrization of part transition . . . . .	103
4.1.2. Part attachment . . . . .	105
4.2. Selecting the right set of parts . . . . .	107
4.3. Summary of our approach . . . . .	109
4.4. Results and comparisons . . . . .	111
4.4.1. Comparing to abstractions with degraded features . . . . .	111
4.4.2. Comparing to traditional geometric simplifications . . . . .	113
4.5. Assessment . . . . .	119
4.5.1. The settings of the preference study . . . . .	121
4.5.2. Significance of shape abstraction . . . . .	124
4.6. Limitations of abstraction in terms of parts . . . . .	126
<b>5. Discussion and Future Work . . . . .</b>	<b>128</b>
5.1. Contribution . . . . .	128
5.2. Limitations of the representation . . . . .	130
5.2.1. The peanut dilemma . . . . .	130
5.2.2. Effects of figure/ground discrimination on abstraction . . . . .	130
5.3. Future work . . . . .	132
<b>Vita . . . . .</b>	<b>151</b>

## Chapter 1

### Representation and Depiction

An effective depiction system lets the viewer effortlessly grasp shape information from the image. Accurate photorealistic images (which are the ultimate goal of photorealistic rendering) do not necessarily achieve this best. Good examples of effective depictions are often found in hand-drawn technical illustrations, stylized caricatures, terrain and road maps and many other forms of visual arts.



**Figure 1.1:** Left: A map of the United Kingdom and Republic of Ireland; Right: a simplified version by a cartographer [Robinson et al., 1995]

Figure 1.1 shows two maps of the United Kingdom and Republic of Ireland. The left one is rendered using standard Azimuthal orthographic projection. The right one is the simplified version we obtained from Robinson and Morrison's cartography text [Robinson et al., 1995]. In this simplified version, many details that exist in the left one are eliminated, including removal of some of the islands (for example, those around the sea of the Hebrides). Also, notice that many bays (firths) that are kept in the abstracted version are opened up so that their presence is emphasized, such as those to the north of Edinburgh and the one to the south of Hull. Again, to the south, the Isle of Wight is made more separated from the coast of England.

According to Robinson and Morrison, cartographic abstraction is a procedure that involves two steps: first analyzing the shape of the outlines and then retaining those distinctive shapes on the map. When cartographers abstract shapes, they do not simply leave out details. Instead, they create bias in the sizes and relative configurations of the geometric components they want to convey. In addition, these components are purposely made more readable and pertinent. Geometric accuracy here is less important than using clean, smooth lines to emphasize the main component parts, which they call the “basic shapes” of the map.

Cartography is not alone in such practices. The same philosophy is often seen in other forms of artistic depiction. It is widely agreed that artistic drawing goes beyond the reproduction of contours. Instead, the subject is always broken down into the most basic form that the drawer perceives, and lines are the expression of the inner form of the awareness of this set of components and particularities [Taylor, 1987]. Art theorists exclusively advocate the *representational concept* [Arnheim, 2004a], which is claimed to be the essential building block of “image-making of any kind.” Representational concepts furnish the concepts of the visual forms being depicted, and the particular formation of this representation in the human visual system distinguishes artistic drawings from nonartistic imaging.

It is certainly desirable for a computer depiction system to follow the same theory so that more effective imaging is possible by systematic structural abstraction. To confirm the feasibility of adopting this philosophy, this thesis proposes a computational model of 2D shape whose representation is in terms of parts. We also describe an experimental shape abstraction system based on this representation. In respect of artistic practice, our depiction system generates the abstracted shapes from a subset of the component parts that is interactively specified by the user based on the geometric prominence and the importance of the part. The effectiveness of this approach is assessed through a perceptual preference study.

For such a system, it is essential that the shape to be depicted is properly represented in order to facilitate the abstraction. In building up our part-based representation, we borrow elements from various existing theories on human perceptual organization and structural shape models. In the rest of this chapter, we survey the state-of-the-art research on how shapes are represented by the human visual system, by computer graphics and by computer vision systems.

## 1.1 Shape, representation and depiction

In traditional computer graphics research, especially photorealistic rendering, shapes are represented in terms of geometric primitives, including pixels, points, curves and surfaces, among which polygons and triangular meshes are probably the most popular representations of shapes in 2D and 3D spaces, respectively. Since the goal is to simulate the appearance of the real world efficiently and realistically, shapes in these systems are represented in such a way that the geometric accuracy is maximized under a given *length* of representation, which typically uses geometric level of detail techniques.

Computer vision research, on the other hand, exclusively involves transforming raw input of shapes (i.e., digital images or geometries, which are also known as low-level representations) into higher-level representations that are suitable for comparison. Depending on the specific task or application, hundreds of different kinds of shape representation schemes have been proposed to facilitate shape indexing, classification, segmentation and recognition [Zhang and Lu, 2004]. For such systems, it is commonly agreed that stability is an important requirement for a good representation. Given the measures of similarities between shapes and between their representations, a small change in a shape should produce only a small change in its representation, and vice versa. This practice of representing shapes is often called the *integral invariants* of the shape, which are stable under various shape operations under certain distance measure in the shape space [Hoffman, 1983; Kuijper and Olsen, 2005; Manay et al., 2006; Ogniewicz, 1994; Shaked and Bruckstein, 1998]. However, the representation may or may not reflect the real behavior of human vision. For example, the widely used spectral transformation techniques, including Fourier and wavelet descriptors, reflect the hypothesis that at each stage of the visual perception, the output is a linear or nonlinear combination of the processing of the neurons at different scales. However, this hypothesis was found to be invalid, according to Albright and Gross [Albright and Gross, 1990]

A successful artistic depiction presents the visual information in a way that leads to a desired representation in the viewer's visual system. Likewise, the representation scheme underlying our graphics system must be compatible with the human visual representation and *initiates the*

*recognition* [Hoffman, 1983; Hoffman and Richards, 1984]. This means that the representation in our application should correlate well with the representation of the human visual system. This makes it possible to directly render, enhance or manipulate the features or structures that the recognition relies on to guarantee the recognizability and achieve the effectiveness of the depiction.

Quite a few perceptually driven shape rendering and model simplification methods have been proposed by applying knowledge of the limitations and capabilities of the human visual perception. Most of this work is grounded in physiological and psychophysical study on early vision [Ferwerda, 2001; Hubel, 1988] The perception of color, orientation, frequency, motion, etc., is studied and applied by enhancing the visual responses in building working visualization systems such as those by Gooch et al. [2004] and Rusinkiewicz et al. [2006], among many others.

Hubel and Wiesel's discovery of neuronal edge detectors [Hubel and Wiesel, 1968] suggests the very beginning of visual processing produces a shape representation in terms of detected edges and silhouettes, which in turn become the inputs of further perceptual activities. Recognition, however, does not rely directly on the boundary geometries of the shape since people see similarity between shapes with different geometries. Thus, in our work, we try to consider the possibility of how an effective computer depiction system can benefit from higher-level shape representations based on the understanding of higher-level perceptual behavior in psychophysics and cognitive science research.

### **1.1.1 Structural representation in terms of parts**

The mechanism of how the human visual system recognizes shapes is still not clear, but we do know that recognition of complex objects is done through the integration of shape components [Gross et al., 1985; Sundar et al., 2003; Tsunoda et al., 2001]. The idea of considering basic components and groups of components as the basic and higher-level structural units supports the theory of parts proposed by Hoffman and Richards [1984], who showed that parts can be considered as the basic shape unit in human perception. The visual system decomposes shapes into parts and the recognition proceeds in terms of matching the patterns or configurations of

the parts.

The idea of shape decomposition is consistent with (and goes beyond) the classic framework by Marr [Marr, 1977; Marr and Nishihara, 1978], which proposes representing the shape in terms of *generalized cylinders*—a hierarchical representation of elongated tube-like components that is formally a generalization of a cylinder, where a base (not necessarily circular as in a cylinder) is swept along some axis (not necessarily straight), allowing for some variations in the size of the base.

Biederman [1987] has made similar arguments about the process of shape perception, which suggests that recognition proceeds in terms of basic components, *geons*, of the shape. Dickinson et al. [1991] apply this idea to build a recognition system by matching shape components with a set of predefined 3D primitives in a hierarchical way. Shapes are decomposed into regions and then into groups of regions (or *aspects* [Dickinson et al., 1991]). The volume primitives and their connectivity are inferred from the aspect hierarchy.

Representation of shape in terms of parts is supported by research in cognitive science. Siddiqi et al. show that humans perceive parts consistently, and such consistency is not only intrapersonal, but also interpersonal. The viewer tends to decompose a shape into parts consistently at different times, and different humans decompose the shape in mostly the same way [Siddiqi et al., 1996], although the salience of each part and the subjects' understanding of its morphology may vary depending to its geometry [Hoffman and Singh, 1997; Siddiqi et al., 2001b].

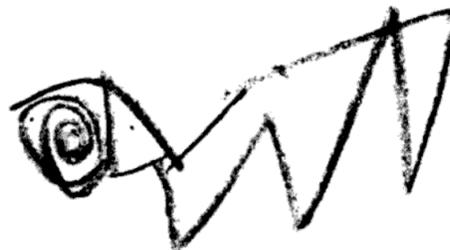
This implies that a shape representation compatible with human recognition should include the notion of a part structure of the shape. The enhancement, distortion or exaggeration of shape features should maintain the consistency of this structure to avoid loss of identifiability.

### **1.1.2 Art as structural abstraction**

Considering artistic depiction as an expression of the apprehended structure of a scene is supported by art theorists and practiced in art. Every drawing guidebook tells the student to learn to *see* the subject and practice seeing the composition of the shape. For example, according

to Delacroix [Arnheim, 2004a], in drawing the object, the artist must first be well aware of the “contrast of its principal lines” before he “sets pencil to paper”. Here, the *principal lines* refer to those frames and skeletons that help to sketch out the essential structure of the subject. Depiction is considered shaping the structure of the form that one has in mind.

The comprehension of the subject’s structure is particularly emphasized in drawing guide-books for children, as humans in their early stages often lack the ability to build this structure through analytical observation of subjects. That leads to drawings with “fused” parts and components [Eng, 1999; Speed, 1972]<sup>1</sup>. Eng argues that this perceptual fusion happens when the undeveloped visual system familiarizes itself with the complex form resulting from the combination of parts until it is capable of conceiving the combination as a whole. Guided by this (limited) perception, the drawing exhibits a degree of “over-abstractionism”, such as illustrated in Figure 1.2 [Arnheim, 2004a, Chapter IV]. A child’s drawing skill grows when his ability of observation becomes more developed and when he is able to break down more complex shapes into simple and meaningful units.



**Figure 1.2:** A horse drawn by a five-year-old boy. Legs are fused together. Courtesy of [Arnheim, 2004a], used with permission.

The example of children’s drawing shows that to achieve this goal, the awareness of the structural, part-based representation is the first essential step. This representation informs the system with “what is known” about the scene, as opposed to “what is seen”. According to the 20th century artist Picasso, the former, rather than the latter, is what an artist depicts [Durand, 2002]. In Chapter 4 we demonstrate that a depiction system shares a similar relationship between “what to depict” and the resulting depiction—the images may fail to communicate the

---

<sup>1</sup> This also partly results from the contour fusion in the motor act of drawing.

shape composition as having distinct parts, if they are not present in the underlying representation. As has been pointed out by Durand [2002], depiction can be considered as the *inverse of the vision* problem. The computer graphics system is responsible for the recognizability of the depicted results.

Certainly, depiction can go far beyond a visual presentation of “what is known” and it means a lot more than just conveying the knowledge about the shape completely and honestly. This is primarily because human vision is not merely a photographic sensory system; instead, it is an act of generalization, symbolization and associative thinking about the geometry concepts the human visual system perceives [Arnheim, 2004b; Imhof, 1982]. This is not only demonstrated in cognitive science, but is also acknowledged by artists. An artist always draws a component as a certain peculiar instance of a symbolic form that he has learned. For example, the shape of the human head is often conceptualized as “an elongated circle” [Ludwin, 1998; Willats, 1997]. One primary goal of depiction, as a form of expression, is to convey the shape geometry and structural composition that the artists choose to see and intend to show:

Art is not what you see but what you make others see.

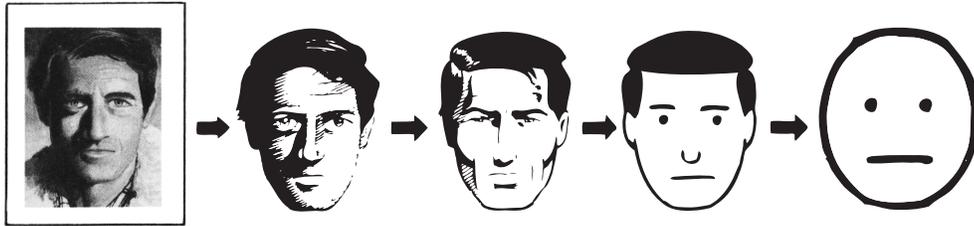
—Edgar Degas<sup>2</sup>

Artists do not attempt to, neither do they need to copy the geometry of the shape onto the paper as accurately as they can. Instead, shape generalization strategies like simplification, deformation, exaggeration, grouping and spatial rearrangement are widely used, as the reflection of the conceptualization at the “seeing” stage [Agrawala and Stolte, 2001; Imhof, 1982; Robinson et al., 1995].

However, improper placement of line strokes may result in a problematic image that leads to unintentional structural perception of the object. Recall that the undeveloped motor action could cause *fused* depiction that fails to articulate the composition of part structure of the visual form. Recall also that the human visual system relies on the part structure for recognition [Hoffman and Richards, 1984]. However, only a small number of parts is essential for this purpose and this set of parts determines how far the abstraction can go. In an abstracted depiction,

---

<sup>2</sup> [http://en.wikiquote.org/wiki/Edgar\\_Degas](http://en.wikiquote.org/wiki/Edgar_Degas)



**Figure 1.3:** A face shape with levels of abstractions. The left-most, framed, is a photograph of a human face. The next four are depictions of this face with four levels of abstraction. Picture courtesy of McCloud [1994], used with permission.

parts that are not essential are candidates for omission.

Figure 1.3 shows an example from the classic comic book “*Understanding Comics: The Invisible Art*” by Scott McCloud [McCloud, 1994] with four abstraction levels of a human face. As shown in the right-most depiction, a few lines and dots are readily perceived as a face by the human visual system.

From the examples in Figure 1.3, we also notice that an artist’s skill goes far beyond just leaving out irrelevant parts. Preserved parts are presented with sophisticated strokes and hatches, as well as exaggeration, idealization and symbolization. In the second depiction, for example, the concavity of the cheeks is exaggerated, and in the third and the fourth, the eyes are symbolized by two dots and the mouth is simply a horizontal line segment. Artists are also good at putting together the preserved parts in a coherent way. For example, although hairs and ears are left out in the right-most depiction, the gaps they leave are naturally filled and the parent part they attach to, the head, is presented with a single stroke.

All these are interesting observations that we consider could potentially help a computer graphics system to produce more effective rendering results. In particular, the task that maps “scene primitives” into corresponding “picture strokes” through omission, exaggeration, etc., is described using the *denotation system* [Willats, 1997, Part II], as opposed to the *drawing system* [Willats, 1997, Part I] which determines the location and orientation of each stroke in the picture space. However, as the first step towards a rendering system that explicitly targets for clarification and recognizability of the depiction, in our research we focus on how we can let our system analyze the structural representation in terms of parts and coherently compose the abstracted version with a selective set of parts. Also, we would like to provide a user interface

that allows the user to interactively guide the shape analysis and the selection of essential parts. We limit the scope to 2D shapes that contain no self-intersections. As an experimental system, we further limit the ability of our shape representation by considering only the hierarchical composition of parts. Modeling of other elements of perceptual organization, such as patterns, texture, etc., requires the development of a more powerful shape understanding system, and we leave this for future work.

## 1.2 Related work

Our research follows the path of structural simplification. For the computation of our structural representation, we incorporate existing theories from cognitive science research on perceptual organization and shape description, such as the minimum principle and causality theory. On a high level, we also follow the understanding of the Gestalt psychologists to locate and utilize the geometric regularities, e.g., symmetry, proximity, etc., for cues of shape composition in terms of parts. This process begins with locating symmetries between edges of shapes [Blum, 1973; Kimia, 2003; Tek and Kimia, 1999].

### 1.2.1 Locally reflective symmetry and axial representation

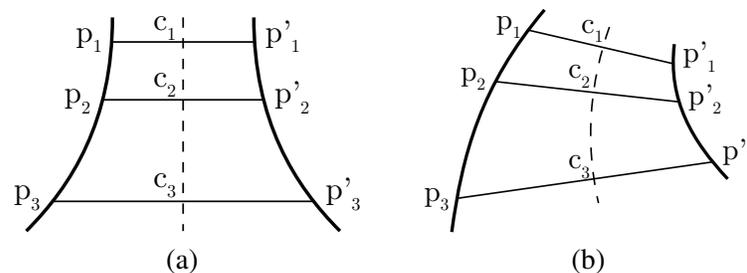
Before we go into more detail about computational approaches to symmetry related shape representations, we would like to point out that in computer graphics and human vision research, symmetry is a general geometric concept. Formally, in Euclidean geometry, symmetry refers to the *group* of transformations that are bijective and distance preserving (*globally isometric*) with the composition as the group operation<sup>3</sup> [Beckman, 1953]. Specifically, such transformations include reflection, translation, rotation and any of their compositions. In our research, however, we only consider the reflective symmetry, partly due to the limitation of the state-of-the-art techniques in discovering all potential symmetries. Fortunately, most single objects in reality

---

<sup>3</sup> This definition of symmetry can be extended to the group of any transformations (or functions) on any system that preserves all the features of interest of the system. See Rosen [2008] for more detailed discussions of symmetry and group theory in a more general setting.

are dominated by reflective symmetry. Symmetries corresponding to translational and rotational transformations are more often seen among different objects when they exhibit certain patterns. As explained in the previous section, an investigation into symmetries such as these is outside the scope of our research.

A 2D shape  $\Omega$  is said to be reflectively symmetric if there exists a line  $l$  such that for any point  $p \in \Omega$ , its reflective counterpart with respect to  $l$ ,  $p'$ , is also in  $\Omega$ , or  $d(\Omega, p') = 0$ . Figure 1.4(a) shows such an example, where each point on the left curve can be mapped one-to-one to its corresponding point on the right curve such that all these pairs share a same perpendicular bisector. However, for a more general 2D shape like shown in Figure 1.4(b), such a reflection does not typically exist. Nevertheless, it is still possible to find this one-to-one map between two curve segments along the shape boundary such that the perpendicular feet of all the pairs form a smooth curve. This was first studied by Brady and Asada who called these continuous, one-to-one mapped pairs of boundary points the *smoothed local symmetries* (SLS) [Brady and Asada, 1984]. The smoothed curve is called the SLS axis.



**Figure 1.4:** a: Reflection about a line; b: Reflection about a curve. Smoothed locally reflective symmetry (right) can be considered as a generalization of the globally reflective symmetry (left)

In the next chapter, we will come back to explore more about the differential geometry and its relationship to the proximity of pairs of contour segments along the smoothed local symmetries. Here we merely point out that Figure 1.4 also implies the generative viewpoint of the *axial* shape representation [Rosenfeld, 1986], which states that the shape could be *generated* by sweeping a template geometry of varying size along a curve called the *rib* or *spine*. This template geometry, called the *generator*, could take any shape (provided that the resulting shape is not degenerate or self-intersecting). In the case of Figure 1.4, it is a line segment. In most popular axial representations, the generator is reflectively symmetric along the axis with respect

to the the line passing through by the point on the curve in a certain direction. To the best of the author’s knowledge, most structural shape simplification approaches in this category exclusively choose axial shape representations. The structural simplification is thus reduced to axial pruning.

Axial representation is region-based. Formally, each axial segment  $ds$  corresponds to a subregion  $\mathfrak{R}_{ds}$  such that the whole shape  $\Omega$  is a union of these subregions, i.e.,  $\Omega_{\mathbb{S}} = \bigcup_{ds \in \mathbb{S}} \mathfrak{R}_{ds}$ . Here  $\mathbb{S}$  is the entire axis of the shape and  $\mathfrak{R}_{ds}$  is the generator at the axial segment  $ds$ . The geometric parameters of the generator at each point on the spine together with the axis itself fully define the geometry of the shape. The axial representation is said to *recover* the shape. This shape recovered by the axial representation is unique, but this is not true in the other direction. In general, one point of the shape could correspond to multiple axial segments since generators can overlap.

Most commonly, the subregion is locally symmetric to itself with respect to the axial segment it corresponds to. Symmetry axis representations are popular not only because they are mathematically well-defined, but also because symmetry is believed to be a crucial organizing principle of shape perception [Leyton, 1987]. This has been validated by psychophysical studies in the past decades. Kovács and Julesz [Kovács and Julesz, 1994; Kovács et al., 1998] have offered strong psychophysical evidence of our visual system’s sensitivity to medial properties. Burbeck and Pizer’s work suggests that the spatial symmetry is indeed computed and represented by the human visual system [Burbeck and Pizer, 1995]. They showed that pairs of responses from shape boundary signals (at a particular scale [Burbeck et al., 1996]) are gathered and transferred for subsequent visual processing. Specifically, at a particular scale, two boundary patches of which the perceived normals form the same angle to the line that connects them are correlated and encoded in the visual system. This directly supports the mathematical definition of a *local symmetry pair*, which will be described in detail later in Section 2.4.1 of next chapter. Lee et al. [1998] have presented evidence that early stages of visual processing are also involved in the perception of the symmetries of local features, though this involvement seems to decrease with increasing feature width. This work also implied that mass centers of local symmetry patches receive stronger visual responses.

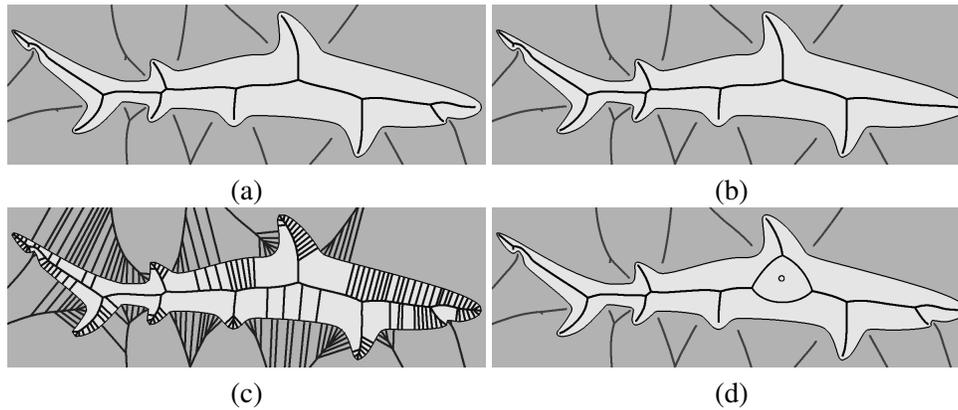
## 1.2.2 Medial axis transformation

The most popular axial representation is probably the medial axis transformation (MAT) proposed by Blum [1967; 1973]. It is defined as the medial axis, the loci of the centers of maximally inscribed disks, together with the radius function defined on the axis. Denote the shape as  $\Omega$ ; the corresponding medial axis as set  $\mathcal{M}$  and the radius function on  $\mathcal{M}$  as  $\mathcal{R} : \mathcal{M} \rightarrow \mathbb{R}^+$ . It turns out that  $\Omega = \bigcup_{\ell \in \mathcal{M}} C(\ell, \mathcal{R}(\ell))$ , where  $C(\ell, \mathcal{R}(\ell))$  refers to the circle centered at  $\ell$ . It has a radius of  $\mathcal{R}(\ell)$  and is tangent to the shape boundary on at least two points. To some extent the medial axis initiates the recognition by giving the intuitive indication of the shape organization with its graph structure. Each edge (called the *medial edge*) of its graph structure corresponds to some portion of the boundary. And for each boundary vertex, there must exist at least one corresponding point (called the *medial center*) or continuous curve segment (called the *ligature* [Blum, 1973]) on the medial edge.

In Figure 1.5(a), the foreground medial edges show nice correspondences to the head, body and fins of the shark. This way, the MAT naturally connects the region and the boundary of the shape and the inherent graph structure facilitates graph-based shape matching [Pelillo et al., 1998; Siddiqi et al., 1998]. With all these advantages, MAT plays an active role in computer vision and shape analysis [Blum, 1973; Blum and Nagel, 1978; Rosenfeld, 1986; Siddiqi et al., 1998].

However, this representation suffers from instability. The transformation from the shape boundary to the medial axis is not continuous under simple distance space like Hausdorff metric  $d_H$  [Munkres, 1999]. That is, for two shapes  $\Omega_1$  and  $\Omega_2$ ,  $d_H(\Omega_1, \Omega_2) \rightarrow 0$  does not guarantee that  $d_H(\mathcal{M}(\Omega_1), \mathcal{M}(\Omega_2)) \rightarrow 0$ . We call this the *differential instability* of the MAT. For example, unstable branches of the medial axis could appear or disappear due to small boundary perturbations, as shown in Figure 1.5(b). Instability of the MAT can also be revealed when a curve is discretized—the medial axis of a polygonal approximation (c) has many more hairs than that of the corresponding shape with a smooth boundary. Another case of instability is shown in (d) where a very small hole is added in the middle, but the medial axis changes substantially so that the topological structure of the axis agrees with the modified shape.

These perturbations are regarded as noise in image analysis [Choi and Seidel, 2001; Ogniewicz and Kübler, 1995; Shaked and Bruckstein, 1998]. All this research suggests that minor perturbations should be considered of less importance (or not considered at all) in the computed shape representation, and the corresponding portion of the medial axis should be pruned. Various approaches to medial axis pruning exist. However, they follow the same paradigm, that is, a significance function is defined on the medial axis, and the medial edges are pruned according to a given threshold.



**Figure 1.5:** The instability of the MAT. (a) The shape and its medial axis; (b) a small perturbation of the shark around the mouth results in a large change in the corresponding medial axis in both the foreground and background; (c) the medial axis of the polygonal approximation; (d) the medial axis of the shark with a tiny hole in the middle. In all cases we have  $d_H(\Omega_1, \Omega_2) \rightarrow 0$ , but the medial axes are very different

Different significance functions were defined. For example, Montanari’s approach of using propagation velocity of the symmetry axis in the grassfire model was typical in early attempts [Montanari, 1968; 1969]. These pruning methods removed unstable medial branches but also resulted in disconnected axes. However, Shaked [Shaked and Bruckstein, 1998] argued that an acceptable pruning method should preserve the axial structure and the result should also be continuous—small differences in pruning degree result in small differences in the axes. Ogniewicz and Kübler [1995] concluded that there were only a few significance measures which could be used in conjunction with a simple thresholding algorithm with no danger of disconnection. Such significance measures had a single local maximum on the axis. Ogniewicz and Kübler suggested a few residual functions on the axis as the definition of the significance, including the ratio of the potential residual value (the length of boundary unfolded by the axis)

and the chord residual value (the distance between the axis' anchor points on the medial edge) [Ogniewicz, 1994; Ogniewicz and Kübler, 1995], which were actually identical to the significance measure given by Blum and Nagel [Blum and Nagel, 1978].

Besides pruning, there exist other methods to avoid differential instabilities, such as defining a proper distance metric between axes [Choi and Seidel, 2001], or using certain curve matching approaches like the optimal curve alignment cost [Sebastian et al., 2003].

### 1.2.3 Structural pruning and robust medial representation

MAT pruning as described in the previous section tries to eliminate long axial edges that correspond to small perturbations to the shape region. However, even though the axis changes continuously with the object's boundary in the sense of Hausdorff distance, the instabilities of the medial axis are also found in the topological connectivity among the axial segments (or *internal structural alterations* of bifurcation points) caused by slight changes in the shape boundary [August et al., 1999]. We call such alterations of bifurcations the *structural instabilities*.

The structural instability not only makes it infeasible to use the medial axis to describe the hierarchy of the shape components, but also makes it hard to design shape matching algorithms that are based on the graph structure of the medial axis [Siddiqi et al., 1998; Zhu and Yuille, 1995]. Crafting different matching schemes between medial axes, so that the MAT representations of more similar shapes have closer matches, is one popular type of approach to graph-based shape matching.

The shock graph [Kimia et al., 1995] is probably the most popular and successful representation in this category. In practice, it is closely related to the medial axis transformation [Kimia, 2003; Serra, 1982]. Instead of assigning a radius to each point on the medial axis, each node in a shock graph is either a continuous segment of medial axis or a singular point. A *first-order* shock is an axial segment along which the associated radius changes monotonically. A *second-order* shock is a stationary axial point where the radius reaches a local minimum. A *third-order*

shock is an axial segment along which the associated radius stays constant. Finally a *fourth-order* shock is any medial point that does not belong to any of the other three categories<sup>4</sup>. The graph is constructed by connecting adjacent medial curves or points in the original medial axis.

Quite a few shock graph matching algorithms have proposed for object classification and recognition. General graph matching is known to be NP-hard [Garey and Johnson, 1979]. But for shock graphs, which can be converted to tree structures under the shock graph grammar defined by Siddiqi and Kimia [Siddiqi and Kimia, 1996], polynomial-time matching algorithms are possible. For example, Siddiqi et al. [1998] reduce the problem into matching of eigenvalue sums of sub-adjacency-matrices of the graph using semidefinite programming. Sebastian et al. [2001] propose using editing distance to measure the similarity between two shock graphs. Here the editing distance is defined to be the minimum number of atomic editing steps it takes to convert one shock graph to the other.

#### 1.2.4 Approximate symmetry

An alternative to complicated MAT matching algorithms is to modify the MAT representation to facilitate the shape matching. To the best of our knowledge, this family of approaches goes back to the shape *representation*; its proponents believe that the shape representation in this sense should help capture the “intuitive concepts” in the sense of how the human visual system analyzes and understands shapes [Richards and Hoffman, 1985; Zhu and Yuille, 1995]. Human perception is implicitly or explicitly modeled in such representations [Mumford, 1991]. In this approach there are two different categories of MAT-based representations which compared with

---

<sup>4</sup> The original definition of the shock graph by Kimia et. al. is given in a more general setting. In Kimia et al. [1995], a shock is defined as the touching point of distinctive boundary points during the process of shape deformation. According to the time of formation, a shock can be one of four types. A *first-order* shock is a boundary point with normal discontinuity. A *second-order* shock is a local minimum of time when it is formed. A *third-order* shock takes the same amount of time to form as its neighboring shocks. And when the process of deformation collapses to a single point a *fourth-order* shock is formed. The whole deformation procedure gives a complete shape segmentation by the sequence of such shocks. Since the structure of the shape remains invariant during the deformation, the resulting segmentation naturally respects the structure of the original shape as well. Obviously the resulting shock set depends on the specific deformation scheme. To preserve the shape structure, Kimia et al. proposed using a combination of constant boundary propagation (the grassfire) and heat diffusion (mean curvature flow). However, almost all the existing shock-graph based shape matching systems assume pure grassfire boundary propagation, which leads to the definition above in terms of the differential properties of radii along the medial axis.

the standard MAT or shock graphs, are more *robust* due to less sensitivity to small boundary perturbations.

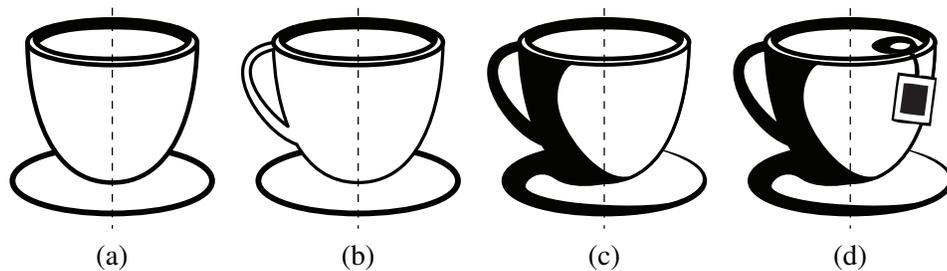
One family of such representations tries to retain only *good* portions of the medial axis. August et al. [1999] point out that the alternating structures of the medial axis are typically related to ligatures. They find that these are less intuitive as visual hints for a human to perceive the structure of the shape. Thus, they propose to use the set of non-ligature medial centers as a more stable, better perceptually-aligned shape representation. This line of research is further elaborated by Macrini et al. [2008] with a more robust ligature detection algorithm and a structured graphical representation called a *bone graph* that properly associates non-ligature medial edges. In such representations, the connectivity between *bones* is no longer defined by the connectivity of the medial edges. Instead, it pursues a more intuitive graph structure.

Approximate symmetry is a different approach. It trades accuracy in shape geometry for the simplicity of the representation. The theory behind this representation admits that the shape structure that the human visual system perceives is not necessarily the result of a precise geometric processing. A partial, approximate description could be more intuitive in capturing the structural representation that the human visual system uses to recognize and understand visual forms [Alt et al., 1988b; Mitra et al., 2006; Podolak et al., 2006].

One way to justify the approximate symmetry is the *minimum principle* [Feldman, 1999], which advocates, from a philosophy perspective, that simpler theories and shorter descriptions about the world tend to be closer to the truth. Approximate symmetry's another name, which tends to be more popular in physics research, is the *broken symmetry* [Rosen, 1998; Wigner, 1965], and it implies another justification: the causal explanation of the shape formation by Leyton [1992]. The causality theory considers the current state of the shape to be the result of a series of *symmetry breaking* events. These events generally break the previously established symmetry in geometry. Accordingly, the shape gets more and more complicated.

An example of such a procedure is illustrated in Figure 1.6, where a cup and a plate with perfect reflective symmetry, with respect to the plane that passes through the dotted line and parallel to the view direction, is shown in (a). This symmetry is broken when the cup is made into a mug with a handle attached on the left, as shown in Figure 1.6(b). The symmetry can

be further broken when the mug is placed into an environment where the lighting condition is asymmetric (c) or even further when it is used to serve a cup of tea (d).



**Figure 1.6:** A sequence of events that breaks the symmetry of the original shape.

Thus, the initial symmetry with respect to the dotted line is not the accurate symmetry of the final asymmetric scene. Instead, it represents the major, dominating symmetry. To recover the initial symmetry of (a) from the final result of (d), a shape analysis system must be able to “smartly” disregard the confusion caused by the asymmetric handle attachment, shading and teabag.

Now the question is, what the *best* approximate symmetry for a given shape is? An immediate answer would be, the one that seems most natural! More specifically, the approximate symmetry of a given shape should grasp the most essential and sufficient structural information of the shape. The region that the symmetry recovers represents our general knowledge about the shape of this object, thus *initiates* our recognition of it. This approximate symmetry is the key to the structurally based shape abstraction since its ultimate goal is to keep only these essential parts!

Unfortunately, this “naturalness” is not a criterion that a computer algorithm can accept. In fact, computing the approximate symmetry has been an open-ended question simply because the problem is not well defined.

One path of research in finding approximate symmetries is based on exact symmetry detection algorithms that try to identify the transformation between geometric primitives, such as Atallah’s algorithms in finding rigid transformation [Atallah, 1984] and reflective symmetries [Atallah, 1985] in an exact sense. The approximate versions of the above problems were then studied by Alt et al. [1988a], where the geometry was limited to point sets at first and then

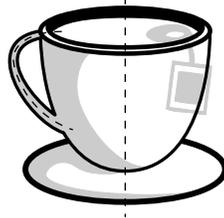
extended to polygons [Alt et al., 1991]. This line of research is more widely studied in computer vision for pattern recognition, registration, tracking, multiple view geometry [Goshtasby, 2005; Hartley and Zisserman, 2003], and robust regression techniques have been applied, such as the RANSAC [Fischler and Bolles, 1987; Stewart, 1999] to handle very noisy inputs, even with structural outliers [Chum and Matas, 2005; Georgescu and Meer, 2004; Jin, 2008].

As opposed to seeking correspondence between different objects, structure computation in our scope demands the computation of self-symmetry. Following the same transformation-based framework, for a given shape  $\Omega$ , this problem is typically formalized so as to find a transformation  $\mathfrak{T}$ , such that for all the points  $p \in \Omega$ , their maps under this transformation  $\mathfrak{T}(p)$  are not very far from  $\Omega$  itself, or that there is an exactly symmetric shape  $\Omega'$  under this transformation ( $\mathfrak{T}(\Omega') = \Omega'$ ) that is not very far from  $\Omega$  [Kazhdan et al., 2004]. The phrase “*not very far*” needs further formalization. The measurement proposed by Alt et al. [1988a; 1991] demands that the maximum of this distance should be below some threshold  $\epsilon$ . While the more popular PCA technique seeks minimization of the sum of all these distances, this is considered not robust [Meer, 2004] and less intuitive [Podolak et al., 2006].

Although the whole point of approximate symmetry is to simplify the representation of the transformation, sticking to one single reflection might be too simplified. Recall Figure 1.4(b) that shows a typical case when a single reflective transformation is not quite capable of describing the symmetry. Instead, the human visual system naturally perceives a one-to-one correspondence between points along the left and the right curves via continuous, smoothly varying transformations. These reflections represent a more complicated transformation that transforms the curves to each other with zero error. In this sense, the whole medial axis can be considered as one complicated (and unstable), but exact, reflective transformation.

Consider also the case of the tea mug shown in Figure 1.6. One would argue that the region that the single reflection could possibly recover is just the main body, which gives you Figure 1.6(a)—a cup. However, the handle is the feature that separates a mug from a regular cup. Its symmetry recovers the other important part of the mug shape that our recognition system relies on to correctly recognize the object as a mug. Thus, in this mug shape, there are two essential axial symmetries. Note that the axial symmetry that recovers the handle is a series of

smoothed local reflections, as shown in Figure 1.7.



**Figure 1.7:** Dotted lines indicate essential symmetries and corresponding parts that make up the mug shape.

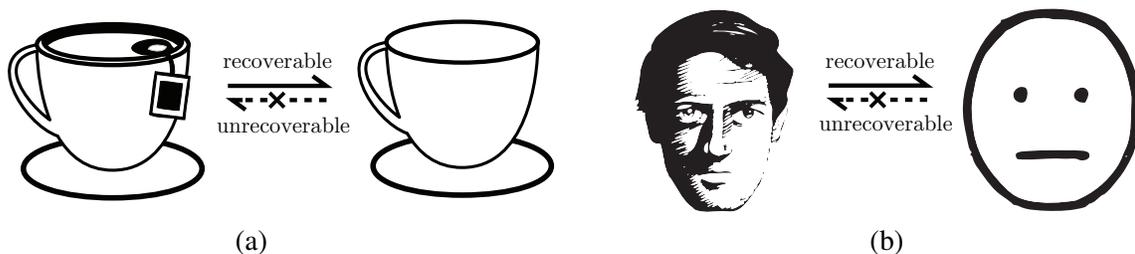
Less relevant parts, such as the shadow and the teabag, are shown in gray.

State-of-the-art research in symmetry description agrees that in order to compute intuitive symmetries of a given shape, the system has to provide a balance between the complexity of the symmetry descriptor (i.e. the minimum principle [Feldman, 1999]) and the accuracy in the recovery of the geometry from the shape descriptor. Keeping these two factors in mind, Feldman and Singh [2006] proposed a Bayesian probabilistic model in which the shape is explained to be generated through a stochastic growing procedure starting from the axial structure. Thus, computing the axis is reduced to the maximization of the joint probability of the axial structure and the shape generated from this axis. The objective probability function takes into account both the accuracy in recovering the shape and the complexity of the axial structure.

An alternative approach towards the approximate symmetry transformation is through sampling and voting. Mitra et al. [2006] proposes sampling symmetric pairs of boundary surface patches of 3D models; each sampled pair votes its local reflective transformation. Then local consensuses, or modes, are located to represent the dominating reflective symmetries. Podolak et al. [2006] finds the major global reflective transformation by a similar idea but instead of the surface boundary, the whole region (volume) of the 2D (3D) shape is sampled. These approximate symmetry detection systems do not produce any new shape representations based on a particular symmetry transformation. Instead, the symmetric structures they exploit are used for more intuitive shape matching [Podolak et al., 2006], symmetrization [Mitra et al., 2007] and more user-friendly interaction techniques [Gal et al., 2009].

Leyton [1992] proposes looking for traces of the shape manipulation from the asymmetry

of the shape. Differing from the axis transformation and generative shape descriptors, this approach takes the causality explanation directly and the recovered symmetry-breaking process is described in terms of a language governed by the *process grammar*. This is fundamentally different from the transformation based approaches described previously. According to Leyton [1992], each recovered shape element carries certain information about the shape under the process grammar that is innate to human perception. In symmetry analysis based purely on geometric transformations, the input is taken as a set of geometric elements, although in the resulting shape representation that those approaches produce, there could inherently be semantics attached, partly due to the symmetry-curvature duality [Leyton, 1987]. The *history* of the shape is the history that we *perceive*, rather than the actual formation of the object. Causality theory does not try to recover the true history of any shape. In fact, according to Leyton, procedures that make the shape more organized are not recoverable even by the human visual system. The shape is the type of “memory” that carries symmetry-breaking procedures only; symmetry-improving procedures, if any, can not be perceived by the human visual system.



**Figure 1.8:** Symmetry breaking and recovery as one-way procedures

Figure 1.8(a) shows two mugs with which we are already familiar. The left mug has a teabag in it, which breaks the symmetry of the mug body. From this figure, a human observer would immediately realize that the teabag is later attached to the mug. However, if you remove the teabag and restore the symmetry of the mug body, as shown on the right in (a), it is not possible for the human observer to infer that these “teabag attachment and removal” events have ever occurred, since removing the teabag improves the symmetry of the scene.

Human perception inevitably involves the recovery of the abstracted, conceptualized version of a given scene. The scene includes only essential visual elements, which the observer uses to index the object. If the observer were to describe the scene, from the viewpoint of Leyton

process grammar, this set of visual elements would sit on the top level of the syntax tree. To the observer, the differences between the original scene and the abstracted version merely serve a decorative purpose. The abstraction is deterministic and essential, but the decorations are arbitrary and optional. Figure 1.8(b) shows another example. In this case, the observer can choose to ignore the peculiarities of the original human face and only retain the conceptual version which a set of symbolized geometric elements suffices to express. However, you cannot ask an artist to draw a portrait by telling him only that the model has two eyes, a nose and a mouth without giving more information. Going from the abstracted shape to the original, concreate, detail-rich scene is simply not possible: you cannot tell whether the cartoon face in Figure 1.8(b) is young or old, male or female, Asian or white, etc.

As mentioned in the very beginning of this thesis, Figures 1.8(a) and (b) are two examples of shape abstraction. They are the results of a systematic, sophisticated removal of less relevant details while faithfully retaining important features. They contain more concise and organized forms of geometry, and so they are closely related to approximate symmetry.

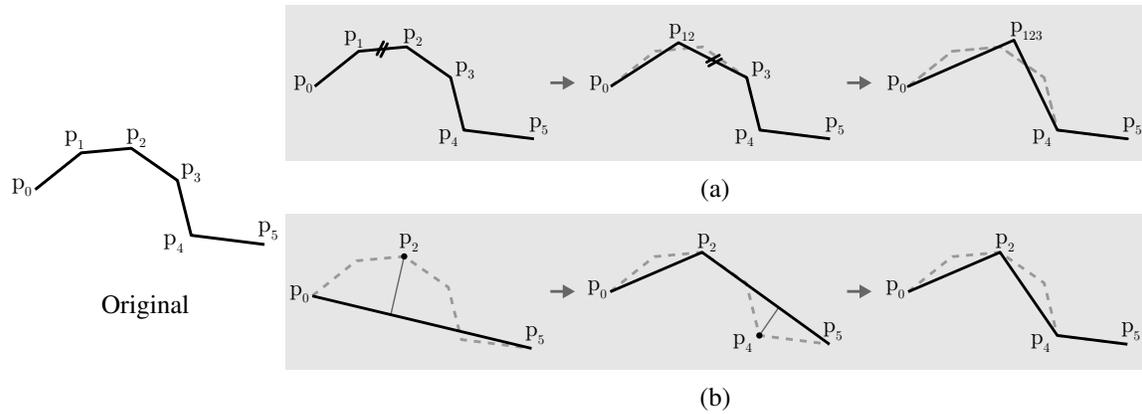
### 1.2.5 Geometry simplification

Figure 1.9 shows two typical shape simplification paradigms in traditional computer graphics. Simplification using Progressive Meshes [Hoppe, 1996] is a bottom-up edge collapse procedure. It starts from the original shape and each simplification step merges one edge into a new vertex. Given the original polygon with six vertices in Figure 1.9, edge  $p_1p_2$  gets first merged into vertex  $p_{12}$  and then this new edge  $p_{12}p_3$  into  $p_{123}$ . This procedure continues until the error in decimation is about to exceed a specified threshold or until the whole shape has collapsed.

Simplification using the Douglas-Peucker algorithm [Douglas and Peucker, 1973] goes in the other direction. It starts with a line segment <sup>5</sup>, and refines the shape by iteratively adding a point from the original shape that is furthest from the approximation. Figure 1.9(b) illustrates

---

<sup>5</sup>For a polyline, Douglas and Peucker [1973] used the line segment that connects the first and last vertices as the initial approximation. The problem of how to obtain the first line segment for a closed polygon was not addressed by [Douglas and Peucker, 1973]. In our implementation, we opt for the pair of vertices from the polygon that is furthest apart from each other



**Figure 1.9:** Decimating a 6-vertex polyline into 4 vertices using (a) Progressive Meshes; (b) Douglas-Peucker

two such steps. At the coarsest scale approximation, which is the line segment  $p_0p_5$ , Douglas-Peucker chooses to insert  $p_2$  since it is the furthest point from the original shape to  $p_0p_5$ . In most cases, choosing to add this furthest point effectively reduces the error of the approximation. Then  $p_4$  is chosen since it is the furthest from the refined polyline  $p_0p_2p_5$ . This procedure continues until the distance to the farthest point is below the specified error threshold or all the vertices have been added, which makes the approximation exactly the same as the original. In both methods, the judgement made in each step locally minimizes the error of the geometric approximation.

There are also shape simplification methods that work in the frequency space. High frequency details are filtered away in the simplified version. [Finkelstein and Salesin's](#) wavelet curve representation [1994] is one such example. Various shape smoothing and signal processing approaches fall into this category, e.g. [Desbrun et al. \[1999\]](#); [Ecker \[2004\]](#); [Jones et al. \[2003b\]](#); [Paris et al. \[2008\]](#), just to name a few. Similar to the Douglas-Peucker algorithm or Progressive Meshes, these approaches also simplify the shape based on local adaptation of boundary geometry.

Abstraction differs from traditional geometric simplification approaches in its processing units and in its reason for simplifying shapes. The goal of geometric simplification is to minimize the spatial distance from the original shape to the simplified version. Thus, the processing unit is typically the boundary curve of the shape. Abstraction, on the other hand, tries to maintain the observer's coarse-level understanding of the scene. It is essential that components

which are the building blocks of the scene in our concept are present in the abstracted version.

One way to better understand their differences is to look at the *requirements* when we simplify shapes. For geometric simplification in a given polygonal model, the requirement proposed typically involves the storage size or the objective distance from the decimated to the original, such as “decimating a 3D mesh from 1M vertices to 5K”, or “the error to the original should be bounded by the amount  $\epsilon$ ”. The requirements can be raised in a continuous manner except for the constraints of topology. For example, the number of geometric simplexes should be integers and they should maintain the Euler characteristics [Armstrong, 1983] if required. For shape abstraction, the requirements are typically to preserve concepts that the user wants to maintain. The abstraction system only removes parts and features that do not account for the object concepts, but the choice of the retained parts and features is implicitly associated with the concept. For example, when the purpose of abstracting the portrait in Figure 1.8(b) is just to conceptualize a human face, then the simplified version shown suffices. But if the requirement is to draw a male’s face, more concrete features will be needed. Likewise, the abstraction in Figure 1.8(a) is good enough to convey the scene “mug” without the teabag and the rim decorations in the original scene.

Instead of crafting a computation model of perceptual parts, Mehra et al. [2009] abstract shapes by extracting feature curves and reconstructs the shape from the set of most salient curves. Their approach accommodates 3D models with non-trivial topology and moderate noise, and still produces pleasing results. Since the structural representation is in terms of feature lines, the success of their system relies on the existence of salient feature lines on the model. Accordingly, the approach performs best on man-made shapes..

### 1.3 Non-photorealistic rendering

Our work on shape abstraction tries to directly address the problem of *what to draw* by transforming the scene into a shape description in terms of perceptual parts. As described in the very beginning of this chapter, this is obviously inspired by the artistic practice. Thus, this work is more appropriately considered as continuing the line of research in non-photorealistic

rendering(NPR) than geometric approximation.

Previous non-photorealistic research addresses related issues by feature removal and enhancement. As introduced in the previous section, smoothing the shape removes high frequency components. It is often used to bring more aesthetic visual presentation in nonphotorealism since human observers seem to favor shapes with such characteristics [Dickie, 1974; Takahashi, 1995]. However, it also blurs the boundaries between components and makes important features less distinctive. Various approaches have been proposed to deal with this issue.

Bilateral diffusion based on various anisotropic kernels is one of the biggest families of feature-preserving smoothing techniques for images [Paris et al., 2008; Tomasi and Manduchi, 1998; Weiss, 2006] and 3D meshes [Fleishman et al., 2003]. Non-iterative feature-preserving smoothing methods based on robust statistics have also been proposed [Fleishman et al., 2005; Jones et al., 2003a]. Smoothing that preserves features has proven to be very useful in producing nice looking, abstracted imaging and video rendering [Kang et al., 2007; Kyprianidis et al., 2009].

To avoid image pixels or surfaces that belong to different components or patches being diffused together, an alternative to feature-preserving smoothing is to compute the patches directly. Segmentation is another widely used technique for image and video abstraction [DeCarlo and Santella, 2002; Orzan et al., 2007; Wang et al., 2004; Winnemöller et al., 2006]. Since the imaging phase after segmentation involves averaging colors (a scalar property of each pixel) over each region, the scope of this technique is limited to image and video abstraction.

Image smoothing and segmentation have had some success in producing nice abstraction results. Due to the enhancement of edge features, on which the early stages of human vision rely, image smoothing and segmentation often help to clarify the structure of the scene. However, the structural clarification is based on replacement or diffusion of neighboring image patches or shape boundaries. The methods themselves do not necessarily provide a new representation of “what to draw”. Thus, these techniques do not seem to be flexible enough for the more creative shape abstraction and exaggeration that are often seen in artistic practice. They instead rely on very coherent user input to achieve satisfactory depictions.

The remaining chapters of this thesis will describe our work on how we can compute a structural, part-based representation about “what to draw” from a 2D shape boundary. Then we will present our shape abstraction system by simplifying this representation, along with the perceptual study we have conducted to evaluate this approach.

## Chapter 2

### Identifying Parts using Contour Grouping Strength

As we have introduced in the first chapter, the human visual system recognizes a complicated shape in terms of parts. In Hoffman [1983], where the concept is first introduced to the field of computational vision, a part is brought out as a “first index in the memory of form”, the “basic unit” for recognition of shape. This is not only about the description of the part, but also about its relation to the rest of the shape. Parts of a visual form are like words of a conversation. When we see parts, and if we are to describe a shape in English, we find that, naturally, our description is also in terms of parts.

However, it is not so natural for a computer system to take up this seemingly simple task to *see* parts. In this chapter, we present an approach for detecting the set of potential parts on a shape, which is the first step in solving the problem above. This work has appeared in [Mi and DeCarlo, 2007] and is presented here in greater detail.

#### 2.1 Locating parts: existing approaches

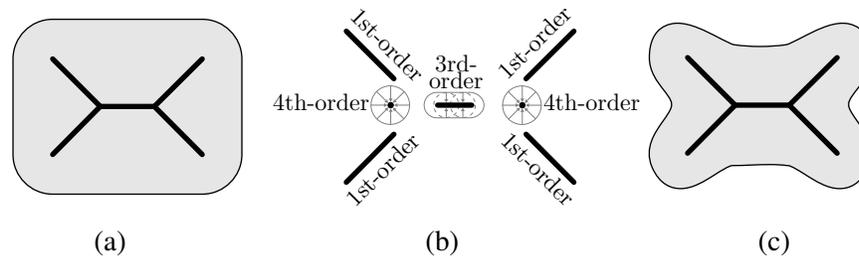
We first review existing computational models of parts, followed by our work on separating parts from 2D shapes, both mathematically and computationally.

##### 2.1.1 Medial axis revisited

The family of symmetry-based structural representations has been extensively surveyed in the previous chapter. In particular, since each branch of the medial axis often matches certain part-like visual components quite well, axial representations based on the medial axis transformation

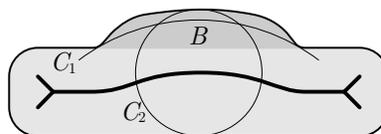
are now one of the most widely used families of structural representations in computer vision and graphics.

In spite of the fact that the symmetry revealed by MAT (called the *medial symmetry* is well supported from neuropsychological study [Burbeck and Pizer, 1995; Burbeck et al., 1996], and that it is widely used in shape representations, the medial axis does not directly map to what seems to be a reasonable representation of parts.



**Figure 2.1:** Different shapes may have identical medial axes (based on Blum and Nagel [1978]) and shock graphs

Let's look at an example to see how axes map to parts. Considering each branch as a visual component sometimes lacks specificity: some of these branches are not parts. Figure 2.1 shows two different shapes, (a) on the left is a single rounded rectangle, whereas (c) on the right is slightly more complicated in terms of part composition. The medial axis of the right figure intuitively predicts the structure of the shape. However, the left figure has several medial edges while the whole shape consists of a single part only. These two shapes share the same medial axis structure—the only difference is the radius function on the axis. Because these medial edges also have the same shock types [Kimia et al., 1995] as shown in Figure 2.1(b), simply attributing shock types to the medial edges does not seem to help in grouping the medial edges together to distinguish (c) from (a). In fact, the instability of the MAT topology also implies that some of the unstable medial edges do not predict any shape information.



**Figure 2.2:** The shallow bump  $B$  is not intuitively represented by MAT. The whole shape is represented as one single ribbon with varying-radius disks as generators

Besides the *precision* issue illustrated by Figure 2.1, and the instability we have discussed earlier in Chapter 1 due to perturbations, the medial axis representation also has problems in accuracy because it does not capture the complete set of symmetries. For example, Figure 2.2 shows a ribbon with a shallow bump  $B$  on the top. However, the symmetries associated with  $B$  are dominated by circles (like  $C_2$ ) along the major axis instead of those like  $C_1$ , and the MAT fails to represent this feature since the bitangent circles representing the local symmetries of  $B$ , such as  $C_1$ , go beyond the bottom of the shape. This inconsistency implies that a human's perception of symmetries and organizations does not necessarily prefer symmetry pairs of smaller distances.

### 2.1.2 Boundary representations

Since axial representations do not always identify visual components from a given shape, we examine geometries of the shape contours to uncover structural information. This idea was strongly supported by Pasupathy and Connor's [2002] findings about how visual neurons respond to boundary curvatures and angular positions. Koenderink and van Doorn [1982] made the argument from the artists' point of view that on the object surface there are elliptic portions with positive Gaussian curvatures and hyperbolic portions with negative Gaussian curvatures. The elliptic portions are typically considered as parts and the hyperbolic patches are perceived as transitional regions: the "glue" that keeps things together but itself is of little interest. This observation agrees quite well with the basic component model already discussed in Chapter 1: the generalized geons or volume primitives have positive normal curvatures everywhere in every direction on their surfaces, except for the bent primitives which have negative curvatures inside the bend. Note that the patches with major and minor negative curvatures do not have any corresponding portion on the outline [Koenderink and van Doorn, 1982; Richards et al., 1987]. Koenderink's theory gives the possibility of computational models based on the local geometry of the shape boundaries. Various boundary-based approaches have been proposed since his work.

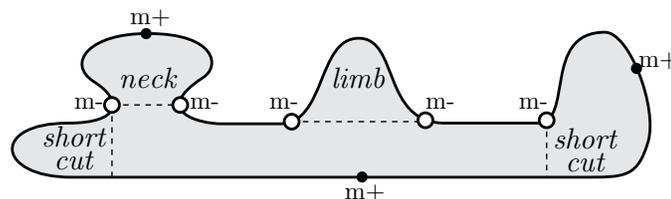
Richards and Hoffman [1985] decomposed the shape boundary at curvature minima into primitives called *codons*. According to the signs of the two curvature minima and the sign of

the curvature maxima in between, a codon can be one of 6 types. However, they also argued that a human does not necessarily partition the shape boundary into codons. Instead, only negative curvature minima are considered as the boundaries between shape components. The perceptual parts of a two-dimensional contour also tend to be the segments bounded by points of greatest negative curvatures. They called this the minima rule. In support of this claim, [Hoffman and Richards \[1984\]](#) present several compelling examples, including Rubin’s classical face-vase illusion.

An obvious follow up to the minima rule is to define pairs of consecutive negative curvature minima as the part boundaries. The portion with positive curvatures (if there is any) is defined as the part [[Latecki and Lakämper, 1999](#); [Rosin, 1999](#)]. However, as both [Latecki and Lakämper \[1999\]](#) and [Rosin \[1999\]](#) have pointed out, this might not always work. Symmetry is still needed to supplement boundary cues in locating and separating parts [[De Winter and Wagemans, 2006](#)].

This leads to two types of part partitioning schemes proposed by [Siddiqi et al. \[1995; 1996\]](#). In their model, a part can be either *neck-based* or a *limb*. In both cases, the parts are separated from the rest of the shape by local pairs of medial symmetries such that the two boundary points are local curvature minima. If the corresponding radius function reaches a local minimum, it forms a neck, otherwise, a limb. They correspond to the second-order and first-order shocks, respectively.

Figure 2.3 summarizes the above described theories about human perceptual organization based on geometric cues on the boundary. The dot lines separating parts are called *part lines* according to [Siddiqi et al. \[1995\]](#).



**Figure 2.3:** Boundary cues that suggest potential part cuts. Solid dots labeled with “m+”s are positive curvature minima while circles labeled with “m-”s are negative minima. Together they divide the shape boundary into the codons, while only negative curvature minima potentially divide the boundary into parts.

Singh et al. [1999] extended the minima rule and the part models of Siddiqi et al. [1995] with the *short-cut rule*, which states that the human visual system prefers to connect segmentation points that are in close proximity, in order to form a part. The two sides of the part boundary should form a symmetry pair and at least one side is a negative curvature minimum. As also shown in Figure 2.3, the symmetry pair adjacent to the leftmost “m-” separates a salient part, while only one side has negative curvature. This is also the case for the rightmost negative curvature minima.

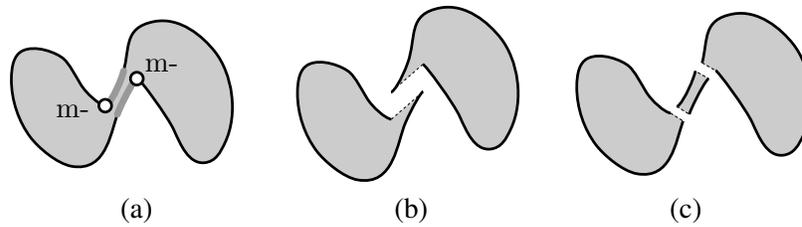
## 2.2 Symmetry in shape description and part separation

Chapter 1 described how the human visual system benefits from the reflective symmetry properties of the shape in building up the visual representation. This section, in particular, reinforces the key role of symmetry in the perception of parts.

Note the connection between the different roles symmetry plays in the shape description and the part separation by the visual system. When we use symmetry to *describe* a shape (either in the sense of causality explanation [Leyton, 1992] or as a mechanism of description minimization [Feldman, 1999; Mitra et al., 2006]), we effortlessly *relate* one side of the shape boundary to the other; and points of both sides are paired in a smooth and continuous manner.

The human visual system’s inclination in pairing symmetric boundary segments makes symmetric object contours a very strong cue in perceptual grouping. Arguments for this have been made by Geisler et al. [2001], Orabona et al. [2006] and Geisler and Perry [2008]. This also justifies computational models that *separate* the part from the rest of the shape along a symmetry pair. Otherwise, our visual system could not find a proper boundary segment that is symmetric with one side of the shape boundary, with the other side that is now in a separate piece due to asymmetric part cut—the originally strongly grouped pairs are cut apart!

We take an example from Singh et al. [1999], shown in Figure 2.4, to illustrate this further. Connecting the only two negative curvature minima (in Figure 2.4(a)) results in a strange cut (Figure 2.4(b)); the cuts shown in (c) following the short-cut rule looks more reasonable. The reason (b) is a bad cut is that the “bridge” that connects the two pieces forms a salient symmetry

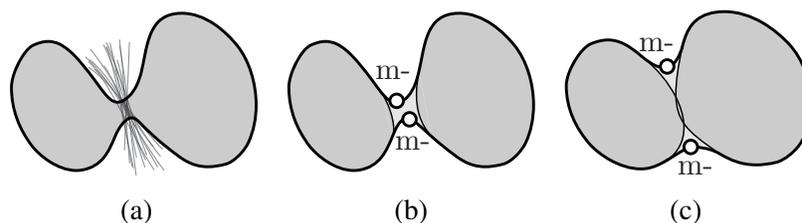


**Figure 2.4:** Do *not* separate strongly grouped contour segments apart!

itself. Due to this symmetry, its two sides, outlined in (a), are strongly grouped perceptually by the human visual system, which excludes any separation between them.

### 2.2.1 Transitional region between parts

A large-scale study on shape segmentation by human subjects shows that negative curvature minima and local symmetries together are effective at predicting the locations of visual parts [De Winter and Wagemans, 2006]. But it also demonstrates moderate inconsistency among subjects. Subjects make part cuts less consistently when the boundaries between parts are less salient and the curvature extrema along the shape boundary are more difficult to localize. Computationally, locating curvature extrema are known to be numerically unstable and unreliable since the estimation of curvatures involves second derivatives. More robust curvature estimation techniques such as proposed by Kalogerakis et al. [2007], however, typically involve some bandwidth parameters, the choice of which affects the location of the curvature extrema.



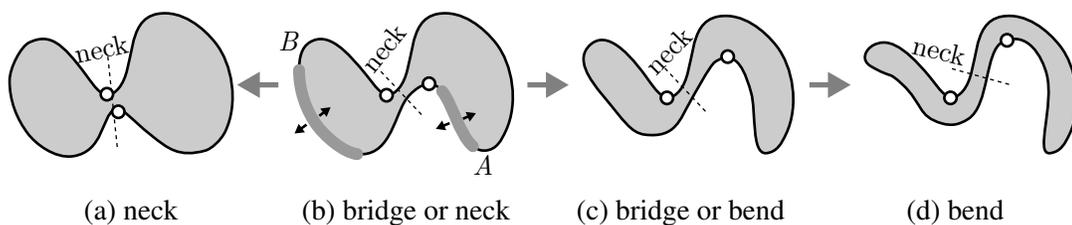
**Figure 2.5:** 15 subjects were asked to cut the shape in (a) into two by drawing a line. The lines they drew were traced and combined as shown in (a). The transition between parts, which is the region that connects these two parts, covers all these cuts (b). A transitional region also exists between overlapping parts (c). In both (b) and (c), dark and light gray colors indicate part body and transitional regions, respectively. Thick lines are shape boundaries.

To demonstrate this inconsistency, we prepared a two-part shape shown in Figure 2.5(a), and conducted a minimal, informal study that involved 15 subjects with a design similar to

De Winter and Wagemans [2006]. The connection between the two parts forms two negative curvature minima and a neck. The locations of the negative minima, however, do not perfectly agree with the location of the neck, which slightly increases the inconsistency among subjects in locating a cut between the two parts [De Winter and Wagemans, 2006]. As expected, subjects roughly agreed on the cut predicted by the neck, with moderate variations. The distribution of the cuts implies that instead of a single, determinant boundary line (which Siddiqi et al. [1995] call the *part line*), there is a region that sits between neighboring parts which accommodates all reasonable part segmentations. Geometrically, the *body* of each part takes its own form. Bodies of neighboring parts may or may not overlap with each other. For adjacent parts with disjoint bodies, as in Figure 2.5(b), this region serves as the *transition* from one part to the other. Figure 2.5 (c) is an illustration of transitions that exist between overlapping parts. They smooth out sharp corners formed by the intersections of part bodies, like fillets often found in mechanical systems [Brady and Asada, 1984].

We propose the concept of the transitional region between parts not only because it respects the perceptual organization of shapes by the human visual system, but also because it provides a mechanism to distinguish different parts of the shape, as will soon be discussed in Section 2.3.

## 2.2.2 Ambiguous shape explanation



**Figure 2.6:** Similar boundary curvature profiles give different perceptual shape explanations.

Due to the existence of a transitional region, a part segmentation is not binary—it’s not a partition of the shape. A point in the shape can belong to no parts, one part, or multiple parts. A morphological analysis on a gradually varying closed curve, illustrated by Figure 2.6, reveals another aspect of how we understand and describe shapes in terms of parts: separating a part from the rest of the shape is often an ambiguous task characterized by multiple explanations.

Singh et al. [1999] originally proposed an alternative yet still reasonable part decomposition for Figure 2.4(a): a cut could be made in the middle of the bridge to make both sides neck-based parts, as illustrated in Figure 2.6(b). The ambiguity can be resolved by “nudging” the two gray portions of the shape boundary as shown in Figure 2.6(b) towards the left or the right. Nudging them to the left results in the shape in Figure 2.6(a), which we have already studied in the previous section. The two negative curvature minima give a strong neck-based part decomposition, whereas the bridge no longer exists. Nudging them to the right makes the bridge feature less salient. However, the entire shape, as shown in Figure 2.6(c), starts to exhibit a bending structure—another ambiguous explanation! Of course, nudging the two boundary segments in gray even further right again, as in (d), disambiguates the description by excluding the compositional explanation based on the short-cut, which is salient in the original shape in (b). Such a shift in the categorization of a shape under deformation was first discussed by Kimia et al. [1990] and then further studied psychophysically by Siddiqi et al. [2001b].

The bending structure shown in (d) raises another interesting point: even with the negative curvature minima, the local symmetries do not form a strong part-cut. The human visual system would rather explain this whole shape as a bent form of a straight band. Here, the minima rule is ineffective.

Consider the justification for the minima rule based on *transversality* that Richards and Hoffman [1985] have originally proposed. In particular, when a part intersects with the rest of the shape transversally on only one side, this cut will only be triggered by a single negative curvature minimum. However, a single negative minimum may or may not trigger a part cut perceptually, as has been illustrated by Figures 2.6(b) and (d).

This leads to reconsideration of the cues that the human visual system uses to decompose a shape. In the rest of this chapter, we try to approach this question from two directions. First, instead of trying to find out what kind of geometric properties the transitions should have in order to separate parts, we look back at the original, and equivalent question: what kind of properties should the shape boundaries have to qualify as belonging to the same perceptual component, i.e., the *invariant* of the part body. The minima rule cannot be used when the cause of negative curvature minima is a structural bend (see Figure 2.6(d)). Thus, a computational model

that can effectively distinguish bending from transversality would be desirable.

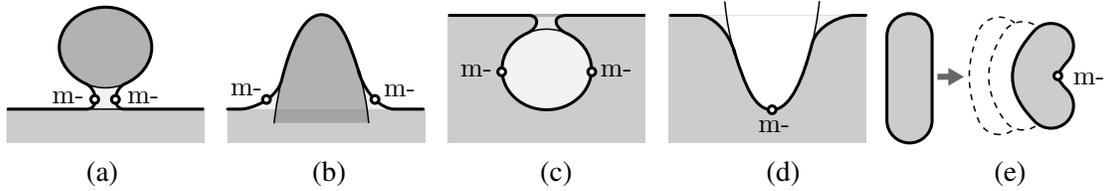
Second, we relate this model to the classic theory that gestalt psychologists have been advocating for almost a century: *contour grouping*. *Good continuity* is a strong hint that the human visual system should group together segments of boundary curves into identifiable shape components. We introduce results from cognitive science on the mathematical descriptions of contour grouping strength, and show its equivalence to the invariant of the part body that we use to separate parts.

### 2.3 Invariant of that basic part with reflective symmetry

A *basic part* is the primary unit in the composition of the shape: it is a part that could not be further decomposed into smaller parts. In the rest of this chapter, we will extensively discuss the geometry of the basic part and leave the more complicated compound parts for the next chapter.

From the viewpoint of causality theory, a negative curvature minimum in a basic part can have three types of explanations: additive composition, subtractive composition, or bending. Figures 2.7(a) and (b) show two types of additive compositions. In particular, the body of (a) is only connected to the rest of the shape through the transitional region, while in (b), the body intersects the main shape. However, it differs from the body-intersecting case in Figure 2.5(c) in that the body of this part is not closed. Typically this “open” body corresponds to a protrusion, or a limb, according to Siddiqi et al. [1996]; while a closed-body part connects to the rest of the shape through a neck. Despite the difference in both cases, the transitions are characterized by negative curvatures.

Figures 2.7(c) and (d) are the inverse of (a) and (b), where the parts are results of certain “cut-out” events. In the latter cases, the *transitions* between the parts and the main shapes have positive curvatures, while the *bodies* of the parts contain concave boundaries. These parts are not contained within the shape; instead, they are on the background and are thus called *background parts*. The regular ones contained in the shape are the *foreground parts*. Beside positive and negative part compositions, we regard all other negative curvatures as caused by *bending*,



**Figure 2.7:** Different types of “history” events that cause negative curvature minima on the shape boundaries. (a) disjoint additive composition; (b) intersecting additive composition; (c) disjoint subtractive composition; (d) intersecting subtractive composition; (e) part bending. (a)-(d) are the results of part attachment, while (e) is not. Dark and light gray colors indicate part body and transitional regions, respectively. Thick lines are shape boundaries.

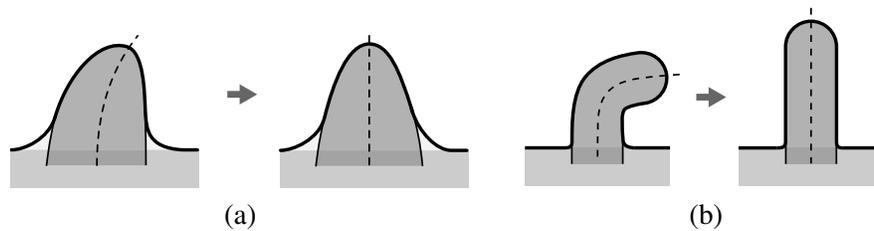
illustrated in Figure 2.7(e) (Obviously, like (c) and (d), there could also be a background version of this bending structure that leads to a positive curvature maximum).

The geometric properties of parts on the background are simply the reverses of the foreground parts and we can easily extend existing computational models, as well as our model of part to these cases with minimal modification. For example, the minima rule could be used to predict background parts by locating positive curvature maxima. However, the existence of negative parts does complicate the analysis of the shape structure. For now, we focus on the foreground parts only until the next chapter, when we will come back to deal with parts on both sides of the shape boundary. This simplifies the task: we only need to differentiate a bending structure from an additive part composition.

To avoid being “confused” by the negative curvature on the shape boundary that is caused by structural bending, we explore the axial structures that are present in the shape and refer to the geometry of the corresponding “unbent” version in order to determine the existence of the transition between parts. To recover the unbent version of an axial structure, we refer to a generative axial representation [Rosenfeld, 1986] and regenerate the shape with a straightened-out spine and the same set of generators. Note that although skeleton-driven deformation often faces difficulties in handling self-intersecting boundaries (typically when the radius of the curvature of the spine is smaller than the radius of the generator), the unbending process does not need to address this issue since the new spine is straight. Computing the axial description is an ill-defined problem and we will describe our approach to it in Section 2.4. Here, we point out that by ruling out the subtractive composition and canceling the bending of the axial structure, additive composition, or transversality, becomes the only cause of concave boundary contours.

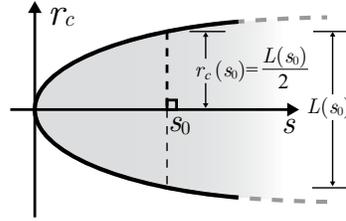
Figure 2.7(e) and Figure 2.8 illustrate several cases of the unbent process. In particular, Figure 2.8(a) shows a case where a bent axis does not accompany the concave boundary—curved axis does not always lead to negative curvatures on the boundary.

Associating a bent axial structure with its unbent, “symmetrized” version has been previously proposed by Blum and Nagel [1978] and is supported by Leyton’s theory on the causality explanation of the human visual system. This theory considers every shape as a result of a series of symmetry-breaking events, as we have discussed earlier in Section 1.2.4. A part with a bending structure is not in its primitive state. Instead, it serves as *storage* that records the event in the past that bent the part to its current state from its initial, more symmetric version with a straight axis [Leyton, 2001]. We can examine the boundary geometry of the unbent shape for cues of part compositions since this bending event does not introduce new parts.



**Figure 2.8:** The human visual system automatically “recovers” a previous, symmetrized state at the sight of a bending structure. Note that in (a), the body of the slightly bent part is convex everywhere even though it appears bent. Dashed lines are the centers of reflections, and the arrows denote the unbending preprocessing.

The unbent axial structure is perfectly reflective with respect to its straight axis. Under this condition, we can easily see that the non-concavity condition for the part body is equivalent to the non-concavity of the radius function. Let  $\mathbb{S}$  denote the line of reflection, and  $L(s)$  denote the distance between the symmetry pair with respect to a point  $s$  on  $\mathbb{S}$  (see Figure 2.9). Here, we call the line segment that connects the symmetry pair the *chord*;  $L(s)$  the *chordal length*. In the rest of this thesis, we will often refer to the *chordal radius*, denoted as  $r_c$ . It is defined as half of  $L(s)$ . Note that this is not the radius of the medial circle. The latter is defined as the distance from one side of the symmetry pair to the center of the medial circle, while the chordal radius is the distance to the center of the reflection. It is introduced mainly for simplicity of notation.



**Figure 2.9:**  $L$  and  $r_c$  are convex functions within a single part with perfect reflective symmetry.

The shape of the part body is completely defined by the chordal radius.  $r_c$  and  $L$  are concave functions with respect to  $s$ , i.e. the mathematical form of the part-body invariant:

$$\frac{d^2L}{ds^2} \leq 0 \quad (2.1)$$

## 2.4 The basic part with smoothed local symmetries

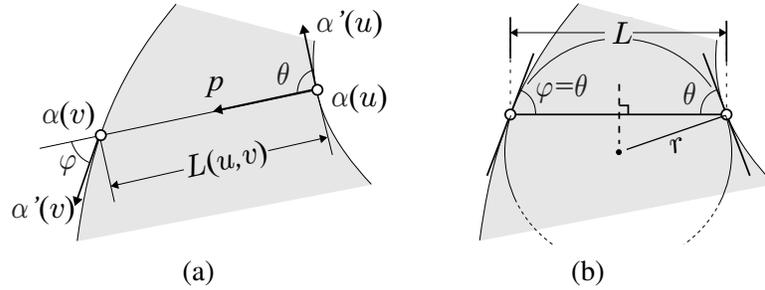
The invariant we proposed in Section 2.3 is general, but so far we assume we already have the unbent version of the part body with a straight axial structure. As explained in the previous section, this requires the computation of an axial representation of the original shape. Rosenfeld pointed out that this is an ill-posed problem since there typically exist multiple generative explanations [1986]. In our research, we use Brady and Asada's smoothed local symmetries (SLS) as it has been one of the most popular axial descriptions. Note that for some parts, e.g., a rounded parallelogram, a shearing deformation would be a better causal explanation. Such parts are better described by Brooks' *generalized ribbons* [Brooks, 1981]. We do not explore this possibility here.

### 2.4.1 Geometry of local symmetry

Although different definitions of local symmetries exist and they lead to different symmetry axis computations, Brady's SLS uses the canonical (and most popular) definition of local reflective symmetry, as is described in Section 1.2.1 [Blum, 1967; Brady and Asada, 1984; Giblin and Brassett, 1985]<sup>1</sup>. We represent the boundary of a simple 2D shape by a  $C^2$  continuous

<sup>1</sup> The MAT [Blum, 1967] and the SLS [Brady and Asada, 1984] both use a bitangent circle to define a pair of local reflective symmetries. However, the former restricts that the circle must be completely inside or outside the

curve  $\alpha : \mathbb{I} \rightarrow \mathbb{R}^2$ , where  $\alpha$  is arc-length parameterized with the domain  $\mathbb{I}$ . Consider a pair of points on the shape boundary with parameters  $u$  and  $v$ , where  $u, v \in \mathbb{I}$ . By definition, their locations in the Euclidean space are  $\alpha(u)$ ,  $\alpha(v)$  and tangent directions  $\alpha'(u)$ ,  $\alpha'(v)$ , respectively. We define the *chord* as the line segment between  $\alpha(u)$  and  $\alpha(v)$  and denote its length as  $L(u, v) = \|\alpha(v) - \alpha(u)\|$  and its direction as  $\mathbf{p}(u, v) = \frac{1}{L(u, v)}(\alpha(v) - \alpha(u))$ . See Figure 2.10(a) for an illustration.



**Figure 2.10:** Local geometry of a pair of boundary points with general position (a); and a local symmetry pair (b). The dotted line in (b) is the local perpendicular bisector.

Any two points  $u, v$  on the plane can be reflectively symmetric with respect to their perpendicular bisector. However, this does not qualify them to form a local symmetry. Instead, we further require that the angles  $\theta(u, v)$  from  $\alpha'(u)$  to  $\mathbf{p}(u, v)$  and  $\varphi(u, v)$  from  $\mathbf{p}(u, v)$  to  $\alpha'(v)$  be the same, where the range of  $\theta(u, v)$  and  $\varphi(u, v)$  are defined to be  $(-\pi, \pi]$ . This means that the tangent lines at these two boundary points should also form a reflection with respect to the same perpendicular bisector, as shown as the dotted line in Figure 2.10(b). This bisecting line passes through the center of the unique circle that is bitangent to both  $\alpha(u)$  and  $\alpha(v)$ .

To formalize the mathematical form of local symmetry, and to facilitate the notation when we study the geometric properties of the local symmetries further in the next section, we define the *asymmetry*  $\mathcal{A}$  of  $u$  and  $v$  as:

$$\mathcal{A}(u, v) = \varpi(\theta(u, v) - \varphi(u, v)); \quad (2.2)$$

Here  $\varpi(\cdot)$  is the *wrapping* operator that “pulls” any angle back to the interval  $(-\pi, \pi]$ . One implementation of this operation is  $\varpi(\phi) = \text{atan2}(\sin \phi, \cos \phi)$ . Naturally, when the value of

---

shape.

$\mathcal{A}(u, v) = 0$ ,  $u$  and  $v$  form a locally reflective symmetry.

Note that the values of  $\theta$ ,  $\varphi$  do depend on the order of the parameters  $u, v$ . Specifically,  $\theta(v, u) = \varpi(\pi - \theta(u, v))$ , and the same relation applies to  $\varphi(u, v)$ .  $\mathcal{A}$ , however, is not affected by this order. In the rest of this chapter we maintain the convention by placing  $u$  ahead of  $v$ , unless otherwise stated. When the context is clear, we simply omit the parameters  $(u, v)$  for  $L$ ,  $p$ ,  $\theta$ ,  $\varphi$  and  $\mathcal{A}$  etc.

The geometry of smoothed local reflective symmetries has been extensively studied by [Brady and Asada \[1984\]](#), as well as by [Giblin and Brassett \[1985\]](#). We want to extract these symmetries, so of relevance here are the differential correspondences between locally symmetric pairs. Starting from a local symmetry  $(u, v)$ , if one side  $u$  moves along the curve boundary for an infinitesimal amount  $du$ , how much should  $v$  move to remain symmetric with  $u$ ? This can be expressed by the following equation [[Giblin and Brassett, 1985](#)]:

$$\left(\frac{1}{r} - \kappa(u)\right) du + \left(\frac{1}{r} - \kappa(v)\right) dv = 0. \quad (2.3)$$

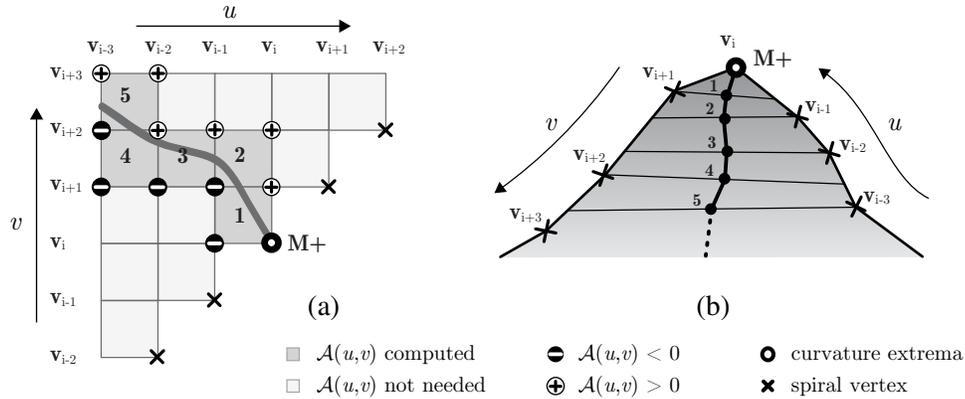
where  $r$  is the radius of the bitangent circle as shown in [Figure 2.10\(b\)](#), and  $\kappa(u)$  and  $\kappa(v)$  denote the curvatures at  $u$  and  $v$  respectively.

## 2.4.2 Tracing the smoothed local symmetries

Equation (2.3) is of more interest in revealing some mathematical facts about the smoothed local symmetries. When it comes to computing the set of smoothed local symmetries, however, a robust algorithm that directly applies this relationship to determine the set of smoothed local symmetries is not so obvious [[Giblin and Brassett, 1985](#)].

Numerical solutions for computing the full set of local reflective symmetries do exist, including the original pixel-based approach by [Brady and Asada \[1984\]](#), and later efforts based on boundary curves represented by piecewise cubic B-splines [[Saint-Marc et al., 1993](#)]. Both algorithms have  $O(n^2)$  complexities, with  $n$  being the number of line/curve segments, and require an additional step to connect computed local symmetries.

Our algorithm takes advantage of the *symmetry-curvature duality* theorem [Leyton, 1987] by *tracing* all the symmetries of basic parts of a given shape boundary represented by polygons. The symmetry-curvature duality theorem states that smoothed local symmetry axes only terminate at curvature extrema. For any curvature extremum of a planar curve, there is a unique smoothed local symmetry axis that terminates there. In other words, we do not need to trace the symmetry from spiral<sup>2</sup> boundary curve segments<sup>3</sup>.



**Figure 2.11:** Tracing smoothed local symmetries using marching squares. The first 5 steps of the tracing are visualized in both the parameter space (a) and the Euclidean space (b). Numbers 1, 2, ..., 5 show in each step which pairs of boundary points are examined.

We illustrate our algorithm in Figure 2.11, which shows the first five steps of the symmetry tracing starting from a positive curvature maxima  $v_i$  (also labeled with “M+”). Figure 2.11(a) shows a portion of the parameter space of the pair  $(u, v)$  arranged as a grid. Pairs with  $u = v$  are along the diagonal, each pair corresponding to a vertex of the boundary polygon. By the symmetry-curvature duality theorem, we know that there are smoothed local symmetries starting from each vertex of curvature extremum, and it is safe to skip other vertices since they are on segments of spiral curves.

<sup>2</sup> A segment of boundary curve is called spiral if it has monotonically varying curvatures.

<sup>3</sup> Unfortunately, not all smoothed local symmetries can be traced from shape boundaries, since SLS can form a loop that does not end at any boundary point, according to Giblin and Brassett [1985]. One notable case is an axis that encircles the hole in an O shape. We believe that any relevant axis that does not terminate at a curvature extremum will be present in the MAT. We do not pursue this further, and exclude shapes with holes from our work.

Note that extending the symmetry curvature duality to curves represented by polygons requires properly estimating curvatures at the vertices. Appendix 2.B briefly describes our curvature estimator. It also proves that  $\mathcal{A}(\mathbf{v}_{i-1}, \mathbf{v}_i)$  and  $\mathcal{A}(\mathbf{v}_i, \mathbf{v}_{i+1})$  must have different signs for any vertex  $\mathbf{v}_i$  such that  $(\kappa(\mathbf{v}_{i-1}) - \kappa(\mathbf{v}_i)) (\kappa(\mathbf{v}_i) - \kappa(\mathbf{v}_{i+1})) < 0$ , i.e., the estimated curvature reaches a local extremum at  $\mathbf{v}_i$ . It follows that there must be either a point on the edge  $\mathbf{v}_{i-1}\mathbf{v}_i$  that forms a local symmetry with  $\mathbf{v}_{i+1}$  or a point on the edge  $\mathbf{v}_i\mathbf{v}_{i+1}$  locally symmetric with  $\mathbf{v}_{i-1}$ . This pair initiates the tracing of the discrete smoothed local symmetries from  $\mathbf{v}_i$ .

To find the initial symmetry pair of a series of discrete SLS starting from a local extremum  $\mathbf{v}_i$ , we first examine the square adjacent to  $\mathbf{v}_i$  (square 1 in Figure 2.11(a)). We evaluate the asymmetry measure of the vertex pair corresponding to each corner of this square. Then, for each side of the square, we compare the signs of the two corners' asymmetries. If the signs are different, then there must be a zero crossing on that edge. In the case shown in Figure 2.11(a),  $\mathcal{A}(\mathbf{v}_{i-1}, \mathbf{v}_{i+1})$  and  $\mathcal{A}(\mathbf{v}_i, \mathbf{v}_{i+1})$  have different signs. Therefore, we know that vertex  $\mathbf{v}_{i+1}$  is locally symmetric to some point on the edge  $\mathbf{v}_{i-1}\mathbf{v}_i$ .

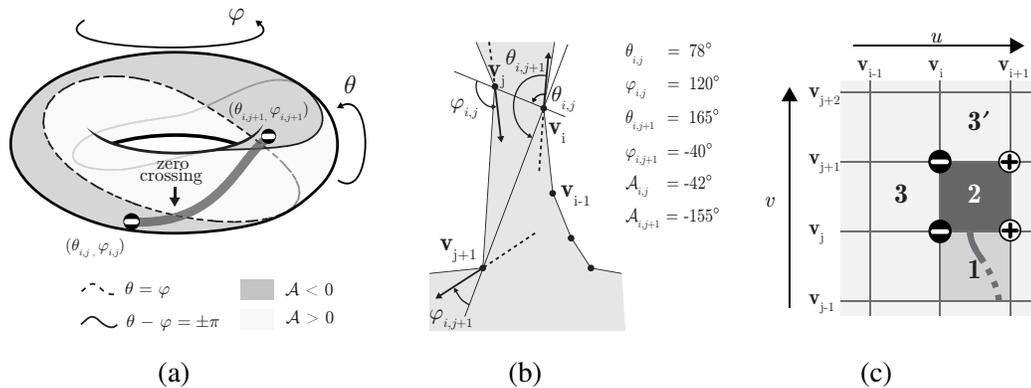
The benefit of this tracing approach lies in that after square 1 has been processed, the only square to process next is the one on the other side of the grid edge with different signs of asymmetries on either end, which is square 2 in Figure 2.11(a). Thus, a significant portion of asymmetries of vertex pairs is never computed. This process can keep going until we reach the other end of the smoothed local symmetries—another curvature extremum. The tracing will not end in the middle (away from the diagonal in the parameter space). There must always be another edge since the  $\mathcal{A}$ s on each side of the curve have different signs. This is also proven by [Giblin and Brassett \[1985\]](#).

When all four sides of the square across the zero-crossing edge have zero-crossings, an ambiguity needs to be resolved about which direction to continue the tracing. In our system, we simply choose the next side counterclockwise as the exit for tracings from positive curvature extrema, while for those from negative extrema, we choose the next side clockwise. This scheme effectively explores all the symmetries that end at curvature extrema.

It is obvious that the time complexity of this tracing algorithm is linear to the number of symmetry pairs that could possibly be discovered. One extreme case is a circle discretized into

an equilateral  $n$ -gon, in which case the complexity is  $O(n^2)$ .

The tracing algorithm is inspired by the traditional marching squares and cubes [Lorenson and Cline, 1987]. However, there is a potential issue in using the marching squares algorithm in the  $(u, v)$  space. Due to the introduction of the wrapping function  $\varpi$ , the asymmetry measure is not defined in Euclidean space, but rather on a torus parameterized by  $\theta$  and  $\varphi$ <sup>4</sup>. Figure 2.12(a) shows the topology of this toroidal  $(\theta, \varphi)$  space, where  $\theta$  is used as the radius of the cross section and  $\varphi$  as the radius of the ring. The solid curve on the torus shows where  $\theta = \varphi$ , which we call the *curve of symmetry*. Unfortunately, this closed curve does not split the space into two half spaces by itself. Instead, it does so together with the curve  $\theta - \varphi = \pm\pi$ , which we call the *curve of discontinuity*. Additionally, the two subspaces should not be characterized by the signs of  $\theta - \varphi$ , but by the signs of the asymmetries, since the range of  $\theta - \varphi$  is  $(-2\pi, 2\pi)$  while only half this interval—the range of the asymmetry measure  $(-\pi, \pi]$ —is meaningful.



**Figure 2.12:** Pitfalls of using the marching squares algorithm directly. The dashed closed curve (portions of which are hidden and thus drawn in translucent style) in (a) is the curve of symmetry while the solid closed curve is that of discontinuity. The value of each angle and corresponding asymmetry measures are shown in (b). The square where improper tracing happens is shown in (c).

For two pairs of  $(\theta, \varphi)$  with  $A$ s of different signs, there are two types of interpolative paths, one that intersects with the curve of symmetry and one that intersects the curve of discontinuity. The fact that these two pairs have the same sign of asymmetry measures does not mean that there are no symmetry pairs in between. Instead, for some badly sampled polygonal models, such as the example in Figure 2.12(b), there does exist a symmetry pair between  $\mathbf{v}_i$  and some

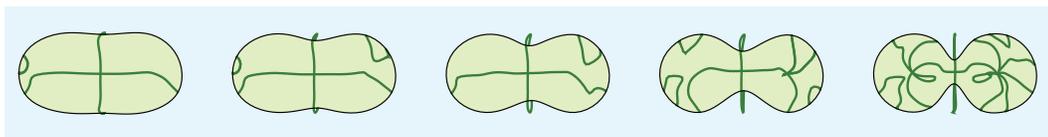
<sup>4</sup> Kuijper et al. [2006] called the  $(\theta, \varphi)$  space the *pre-symmetry set*

point on the edge  $\mathbf{v}_j\mathbf{v}_{j+1}$ . However, the true zero-crossing needs to go a long way on the torus surface by crossing in and again out the  $\mathcal{A} > 0$  subspace. On the other hand, the marching squares scheme we described above will trace down the wrong direction—at the square 2 in Figure 2.12(c); instead of taking 3 as the next square, it goes for 3'. Likewise, a marching square edge with different signs of asymmetry measures does not guarantee a true zero-crossing.

To address this problem, an adaptive refinement step is used to eliminate the possibility of jumps on the torus surface that are too large. Specifically, we watch the ratio between the edge length and the chordal length. If the edge is longer than half of the chord currently being discovered, we subdivide the edge. This effectively limits the length of the curve on the torus corresponding to the marching square edge. Hence, “zero-crossing overshooting” is guaranteed not to happen. A simpler solution (although less efficient) will be to resample the entire polygon before processing to enforce a specific minimum spacing between vertices.

### 2.4.3 Whole SLS vs SLS of basic parts

Figure 2.13 shows the SLS axes of five peanut shapes that our system computes. The set of smoothed local symmetries is quite redundant. (Also recall that the set of symmetries captured by the MAT fails to represent certain salient parts, e.g., shallow bumps). Brady and Asada [1984] pointed out that the boundary point that forms a symmetry pair with a given boundary point is not unique. Some of these symmetry pairs reflect global structures of the shape, some represent local features, and the rest of them do not convey any interesting structural information about the shape. Rom and Médioni [1993] proposed a method of pruning a symmetry set using the angle between the chord and the tangents, the ratio between the length of the smoothed local symmetries and the thickness of the ribbon, and other filtering schemes, which address some of the problems created by the non-intuitiveness of the symmetry set.



**Figure 2.13:** Redundant smoothed local symmetries of peanut shapes

In fact, tracing the smoothed local symmetries is only meaningful *within* a basic part, and the purpose of tracing this axis is to study the second derivatives of the chordal lengths: equation (2.1). We assumed that within a single part body, the human visual system can recover the symmetric version of the part. When the tracing goes into the transitional region, this assumption is no longer valid. Thus, when the tracing of the SLS reaches the end of a part body and enters its transition, that might be a good time to stop tracing. But this does not always work in all situations. In Section 3.1.3, we will further describe the condition to stop the SLS tracing.

Sections of SLS with negative values of equation (2.1) are the focus of our interest since they indicate potential transitions from parts to the rest of the shape. In the section that follows, we will first study in depth the differential geometries about the smoothed local symmetries and then return to locating those interesting part transitions.

## 2.5 Locating the part body and transition

The discussion in the previous section proposed an effective way of locating and describing the part and its transition to the rest of the shape. In practice, we found that using the second derivative of  $L$  with respect to the SLS axis is not ideally stable numerically. In this section we will explore further the geometry of the smoothed local symmetries. It not only facilitates the computation of the part structure, but also naturally connects our approach to the existing computational part models.

### 2.5.1 Chordal geometry

To prepare for a deeper study of the differential geometry of the smoothed local symmetries, we first provide some facts about the geometry of any pair of boundary points  $u$  and  $v$  on the  $C^2$  continuous shape  $\alpha$ , which we call the *chordal geometry*. More detailed derivations of these facts are provided in Appendix 2.A.

We define the *partial tangent circle*  $C(u, v)$  as the circle that is tangent to the curve at  $\alpha(u)$  and passes through  $\alpha(v)$ . (See Figure 2.14, which also shows  $C(v, u)$ ).  $C(u, v)$  and  $C(v, u)$  have radii  $\tilde{r}(u, v) = \frac{L}{2 \sin \theta}$  and  $\tilde{r}(v, u) = \frac{L}{2 \sin \varphi}$ , respectively. Here the radius is signed and is

negative if  $\theta$  or  $\phi$  is negative. If the circle is tangent to  $u$  or  $v$  from outside, then the radius is negative.

We can write the partial derivatives of  $\theta$  and  $\varphi$  in terms of these radii and boundary curvatures:

$$\begin{aligned}\nabla\theta &= \left(-\kappa(u) + \frac{1}{2\tilde{r}(u,v)}, \frac{1}{2\tilde{r}(v,u)}\right) \\ \nabla\varphi &= \left(-\frac{1}{2\tilde{r}(u,v)}, \kappa(v) - \frac{1}{2\tilde{r}(v,u)}\right),\end{aligned}\tag{2.4}$$

Here  $\kappa$  is the curvature of  $\alpha$  and we follow the convention that convex shape boundary segments have positive curvatures. It naturally follows from (2.4) that:

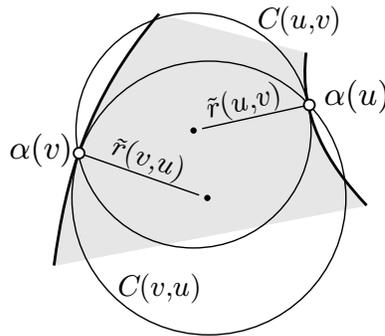
$$\begin{aligned}\nabla(\theta + \varphi) &= (-\kappa(u), \kappa(v)) \\ \nabla(\theta - \varphi) &= \left(-\kappa(u) + \frac{1}{\tilde{r}(u,v)}, -\kappa(v) + \frac{1}{\tilde{r}(v,u)}\right).\end{aligned}\tag{2.5}$$

## 2.5.2 Geometry along smoothed local symmetries

Now we are ready to derive the second derivative of the choral length  $L$  with respect to a small amount of movement along the symmetry, or  $ds$ . Here we denote the curve parameters of the positions of the boundary point  $u$  as  $u + du$ , and  $v$  as  $v + dv$ .

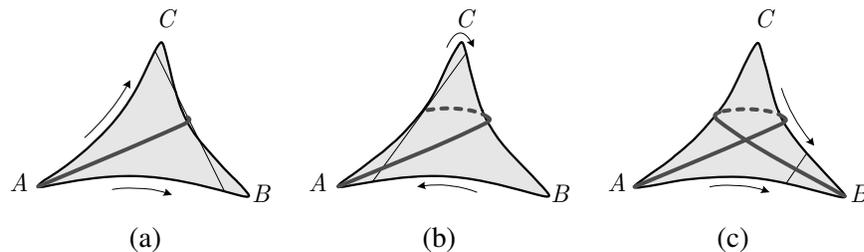
Since  $u$  and  $v$  form a local symmetry, we have  $\theta = \varphi$ . When this is the case, the two partial tangent circles overlap, producing  $r = \tilde{r}(u, v) = \tilde{r}(v, u)$ , which is bitangent with the shape boundary at both  $u$  and  $v$ . Accordingly we call  $r$  the bitangent radius.

Along any symmetry,  $\theta - \varphi = 0$ , thus the change of  $\theta - \varphi$  is also 0, which means  $\nabla(\theta - \varphi) \cdot (du, dv) = 0$ . Using the second equation of (2.5) and replacing  $\tilde{r}(u, v)$  and  $\tilde{r}(v, u)$  with  $r$ ,



**Figure 2.14:** Geometry of partial tangent circles  $C(u, v)$  and  $C(v, u)$

we get the relation of  $du$  and  $dv$  as given by equation (2.3)—just as what Gibbons [1985] had given previously. However we would like to bring to the reader’s attention how the direction of each side of the smoothed local symmetries continues along the boundary. Normally,  $du$  and  $dv$  are of different signs and when one side of the symmetry pair moves counterclockwise along the shape boundary, the other side moves clockwise. The condition for this occurrence is when the bitangent radius is either larger or smaller than the radii of curvatures at both  $u$  and  $v$ . For example, the symmetries captured by the MAT are of this type since all the corresponding bitangent circles are inscribed. However, if the bitangent radius is between the radii of the osculating circles at the two tangent points, then we have  $du$  and  $dv$  have different signs. As one side of the symmetry pair moves on, the other side turns back. These “skewed” local motions of the symmetry pairs on the shape boundary are illustrated by Figure 2.15.



**Figure 2.15:** Motion of the symmetry pair along the boundary. The axis sweeps along intuitive symmetries when the signs of  $du$  and  $dv$  differ, as in (a) and (c). When  $du$  is of the same sign as  $dv$ , local symmetries continue in a “skewed” fashion, as shown in (b).

Brady and Asada [1984] assumed that only the first case could happen, by referring to the work by Bookstein [1979] on algorithms for constructing the medial axis. They did realize, however, that under this assumption, the set of symmetry pairs is not continuous in the  $(u, v)$  space. Thus their definition of the smoothed local symmetries virtually assumes that the local symmetries are examples of the first case. In fact, skewed continuation of local symmetries does not reveal any intuitive symmetry of the shape, so in our implementation, we also choose to ignore them. Still, they need to be traced if we wish to uncover symmetries beyond them (as in Figure 2.15).

Now we take a closer look at the local coordinate system at the SLS axial point,  $s$ , of a symmetry pair  $u$  and  $v$ . Consider an infinitesimal movement  $ds$  along the axis, bringing the

two points to  $u + du$  and  $v + dv$ , respectively, We have:

$$\begin{aligned} \mathbf{s} &= \frac{\boldsymbol{\alpha}(u) + \boldsymbol{\alpha}(v)}{2} \\ \mathbf{p} &= \frac{\boldsymbol{\alpha}'(u) + \boldsymbol{\alpha}'(v)}{2 \cos \theta} \\ d\mathbf{s} &= \frac{1}{2} (\boldsymbol{\alpha}'(u)du + \boldsymbol{\alpha}'(v)dv) \end{aligned}$$

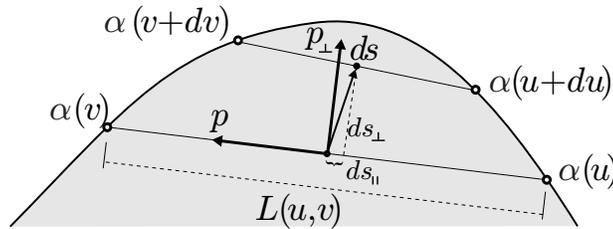
We know that  $u$  and  $v$  form a locally *reflective* symmetry with each other. The center of the symmetry is given by  $\mathbf{s}$  as shown above. The direction goes along the bisector, which is perpendicular with the unit-length chordal direction  $\mathbf{p}$ . This direction, denoted as  $\mathbf{p}_\perp$ , is:

$$\mathbf{p}_\perp = \frac{\boldsymbol{\alpha}'(u) - \boldsymbol{\alpha}'(v)}{2 \sin \theta} \quad (2.6)$$

However,  $\mathbf{p}_\perp$  is not parallel to  $d\mathbf{s}$ . When we project  $d\mathbf{s}$  onto  $\mathbf{p}$  and  $\mathbf{p}_\perp$ , we have:

$$\begin{aligned} ds_\perp &= \frac{1}{2} (du - dv) \sin \theta, \\ ds_\parallel &= \frac{1}{2} (du + dv) \cos \theta \end{aligned} \quad (2.7)$$

where  $s_\perp$  and  $s_\parallel$  are the perpendicular and parallel components of  $d\mathbf{s}$ , respectively.



**Figure 2.16:** Local differential structure of the SLS axis

Figure 2.16 illustrates all the components related to the structure that describes a local movement of the SLS axis. Ideally, we hope that  $ds_\parallel \approx 0$ , since we want the SLS axis to be a good visual representation of the locally reflective symmetry. Whenever  $du$  and  $-dv$  are approximately equal, the parallel component is close to zero. In particular, when the curvatures of the two boundary points are small relative to the reciprocal of the bitangent radius  $r$ , then according to (2.3), the SLS axis does a good job of reflecting the symmetry structure. However, in cases

when  $ds_{||}$  is relatively larger than  $ds_{\perp}$ , we find that the SLS axis is no longer a good representation of the symmetry structure. A typical situation where this occurs is when the “tip” of the part is circular. The exact placement of the SLS is unstable. An example of this instability can be seen on the sides of the peanut shapes in Figure 2.13.

Here, we commit to the reflective symmetry and choose to omit the parallel component of the axial structure. Thus, instead of considering the second derivative of the chordal length with respect to  $ds$  as the part-body invariant expressed by (2.1), we consider it with respect to  $s_{\perp}$ , which revises the invariant to

$$\frac{d^2L}{ds_{\perp}^2} \leq 0 \quad (2.8)$$

In fact, both (2.1) and (2.8) have the same sign, so qualitatively they are equivalent. However, as will be discussed below, by ruling out the non-intuitive portion, the new version (2.8) better aligns our perception of the salience of visual parts.

Appendix 2.A gives the gradient of the lengths of a chord  $(u, v)$  with respect to  $u$  and  $v$  as:

$$\nabla L = (-\cos \theta, \cos \varphi) \quad (2.9)$$

Along the SLS axis, we have  $dL = \nabla L \cdot (du, dv) = (-du + dv) \cos \theta$ ; dividing this by  $ds_{\perp}$  (see equation (2.7)), we get:

$$\frac{dL}{ds_{\perp}} = -2 \cot \theta. \quad (2.10)$$

Also from equation (2.4), we have

$$\frac{d\theta}{ds_{\perp}} = \csc \theta \frac{\left(-2\kappa(u) + \frac{1}{\tilde{r}(u,v)}\right) du + \left(\frac{1}{\tilde{r}(v,u)}\right) dv}{du - dv}$$

Considering that when  $u$  and  $v$  form a local symmetry,  $\tilde{r}(u, v) = \tilde{r}(v, u) = r$ , and also that the relative motion of the two boundary points across the symmetry is given by equation (2.3), we get:

$$\begin{aligned} \frac{d\theta}{ds_{\perp}} &= -\csc \theta \frac{\kappa(u)du - \kappa(v)dv}{du - dv} \\ &= -\csc \theta \frac{\kappa(u) \left(\frac{1}{r} - \kappa(u)\right) + \kappa(v) \left(\frac{1}{r} - \kappa(v)\right)}{\left(\frac{1}{r} - \kappa(u)\right) + \left(\frac{1}{r} - \kappa(v)\right)} \end{aligned} \quad (2.11)$$

This, together with (2.9), and denoting  $w_u = 1/r - \kappa(v)$  and  $w_v = 1/r - \kappa(u)$ , results in:

$$\frac{d^2L}{ds_{\perp}^2} = -2 \csc^3 \theta \frac{\kappa(u)w_u + \kappa(v)w_v}{w_u + w_v} \quad (2.12)$$

This completes the derivations of  $\frac{d^2L}{ds_{\perp}^2}$  in terms of the boundary curvatures and bitangent radius.

### 2.5.3 Contour grouping strength and part transition

The cocircularity of the partial tangent circles across the chord, which implies local reflective symmetry, is a strong cue for perceptual grouping. This motivates us to observe the concavity of the chordal length function by tracking along the SLS axis. Research in cognitive science suggests that although the cocircularity cue dominates, under good cocircularities good *collinearity* further improves the grouping strength. Interested readers can refer to the quantitative study conducted by [Elder and Goldberg \[2002\]](#) on how cocircularity and collinearity (which they call “parallelism”) cues interact in determining the contour continuation.

Collinearity is a measure of how two segments of 2D contour line up with each other. A common practice (such as the measure used by [Elder and Goldberg \[2002\]](#)) is to use the total turning angle from one to the other. Formally, given a pair of boundary points, this is exactly the sum of  $\theta$  and  $\varphi$  shown in Figure 2.10(a). Here we follow existing conventions in research on occluding contour grouping [[Kellman and Shipley, 1991](#); [Singh and Hoffman, 1999](#)], and use the term *relatability* to measure the grouping strength. Originally, the relatability between points  $\alpha(u)$  and  $\alpha(v)$  is defined by  $\mathcal{R}(u, v) = \min_c \int_c |\kappa(\gamma)| d\gamma$ , where  $c$  is any interpolating curve connecting  $\alpha(u)$  and  $\alpha(v)$  such that it is  $C^1$  continuous and does not include any inflections. For a reflectively symmetric pair of contour segments that have the same sign of  $\theta$  and  $\varphi$ , we have  $\mathcal{R}(u, v) = \theta + \varphi$ . This assumption is valid along smoothed local symmetries because of the cocircularity that is already present; the relatability thus depends only on the collinearity.

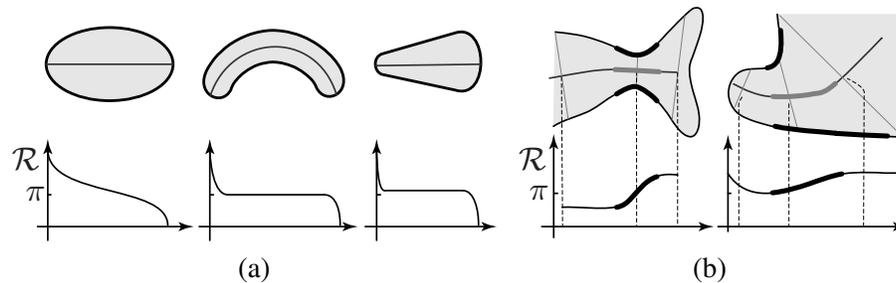
To derive the changes of  $\mathcal{R}(u, v)$ , we use methods similar to those used when we derived equation (2.12) from equation (2.11), resulting in:

$$\frac{d\mathcal{R}}{ds_{\perp}} = -2 \csc \theta \frac{\kappa(u)w_u + \kappa(v)w_v}{w_u + w_v} \quad (2.13)$$

We will further analyze the factors that affect (2.13). For now we point out the relationship between (2.13) and (2.12):

$$\frac{d^2 L}{ds_{\perp}^2} = \csc^2 \theta \frac{d\mathcal{R}}{ds_{\perp}} \quad (2.14)$$

The similarity in forms is not surprising since (2.13) is essentially twice  $\frac{d\theta}{ds_{\perp}}$  evaluated using (2.11). This has a very interesting implication. The classical minima rule, after the “normalization” by the curviness of the symmetry axis, can be explained by perceptual grouping. Namely, if the grouping strength of the local symmetry gets stronger, or at least stays the same along the SLS axis, the axis stays within the body of a part. Otherwise, the axis enters into the section of transition, which leads to the rest of the shape.



**Figure 2.17:** Shape and reliability profiles along symmetry axes

Figure 2.17 further illustrates this theory. The top row shows different fragments of 2D shapes along the SLS axes. In particular, shapes in (a) are geometric primitives, each of which consists of a single part. Their SLS axes are also shown, and below each shape is the corresponding reliability profile along the SLS axis. As we can see, at the starting point in each shape the reliability value is  $2\pi$  and then it monotonically decreases (it may or may not include an interval with a constant value of  $\mathcal{R}$ ), and finally drops to 0, at which point the local symmetry pair becomes most reliable and closes up the shape. Figure 2.17(b), in contrast, shows two fragments of shapes each of which perceptually involves multiple parts. Similarly, the SLS axis is plotted and the corresponding profile of the  $\mathcal{R}$  value is shown under each shape. But unlike examples in (a), both shapes have increasing intervals of  $\mathcal{R}$ , and as we can perceive, these sections with increasing  $\mathcal{R}$ s are exactly where parts come together—transitions.

### 2.5.4 Locating part transitions using relatability

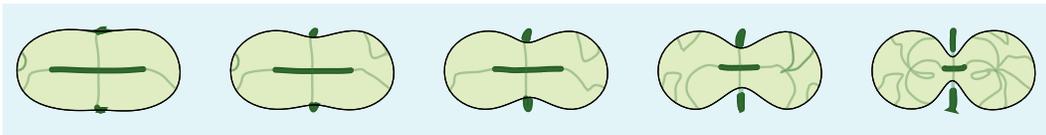
The signs of equations (2.12) and (2.13) help to identify the section of the SLS axis that corresponds to part transition; we also hope to have a measure of the transition strength. Due to the curvature factor in the right hand side of both equations, they are scale-dependent. We simply multiply them with the bitangent radius  $r$  to remove the scale dependency. In spite of the qualitative equivalence between (2.12) and (2.13), we believe that quantitatively,  $\frac{d\mathcal{R}}{ds_{\perp}}$  aligns better with the sensitivity of the visual system towards the shape transition. Putting these together, we give our model of the strength of a part transition at a point on a local symmetry:

**Definition 2.5.1.** The *local strength of the part transition*,  $\mathcal{B}$ , of a symmetry pair is defined as the first derivative of relatability with respect to  $ds_{\perp}$ , normalized by the radius  $r$ :

$$\mathcal{B} = r \frac{d\mathcal{R}}{ds_{\perp}}.$$

In practice, within a part body we would like to accept slight increases of relatabilities (or concavity of chordal lengths), since we expect to encounter noise or fluctuations at different scales. Thus, we use a small threshold  $b_0 > 0$ . Only those with  $\mathcal{B} > b_0$  are considered symmetries of part transition. Useful values of  $b_0$  range from 0.2 to 1. Later in this section we will describe in detail how we set the value  $b_0$ .

Figure 2.18 shows the SLS axes of the same set of the shapes from Figure 2.13, except that only those with  $b_0 > 0$  are highlighted. Note that a local symmetry with  $\mathcal{B} > b_0$ , that qualifies to be transitional, does not necessarily result in a plausible part cut. Although a local maximum of  $\mathcal{B}$  could arguably be a good candidate for separating a part from the shape, as we have explained earlier in Section 2.2, it is often more desirable to locate a continuous section of part transition.  $\mathcal{B}$  could also be spuriously large if one point of the symmetry is a corner vertex (even



**Figure 2.18:** The smoothed local symmetries of the peanut shapes that have positive values of  $\frac{d\mathcal{R}}{ds_{\perp}}$  are highlighted while the rest, most of which redundant, are in a lighter color

a shallow corner), which typically has a very large curvature value under concurrent curvature estimation methods. We want our method to be more robust to such fluctuations.

Instead of looking at individual symmetry pair, we integrate  $\mathcal{B}$  along adjacent local symmetries with  $\mathcal{B} > b_0$ . This leads to the following definition:

**Definition 2.5.2.** The *total strength of a part transition*,  $\mathcal{T}$ , between two parts is defined as the change of relatability across the transitional section:

$$\mathcal{T} = \int_c \frac{d\mathcal{R}}{ds_{\perp}} ds_{\perp},$$

where  $c$  is the axis segment with  $\mathcal{B} > b_0$ .

Note that by integrating equation (2.13) with  $s_{\perp}$ , the scale dependency is removed. Thus, normalization with  $r$  is not needed. The integral is a measurement of the change in total turning angle,  $\theta + \varphi$ , along the transitional section of local symmetries, which is considered one of the the part saliency measurements proposed by Hoffman and Singh [1997].

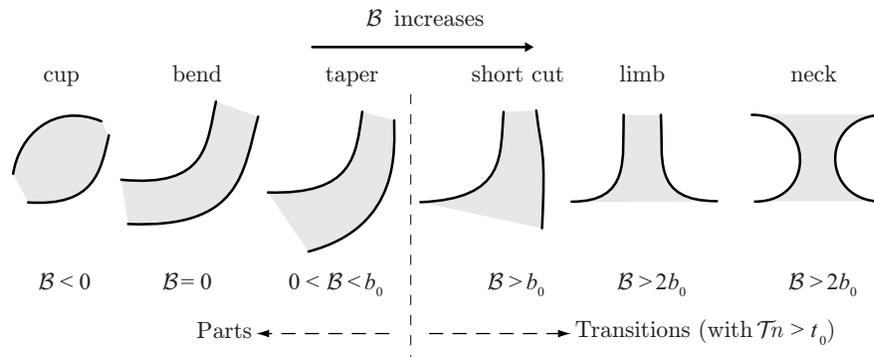
In our system, only a section of symmetries with significant total strength is considered as a potential part transition. However, using the difference in turning angles as the total boundary strength is not conducive to detecting shallow parts. In practice, we use the normalized total strength  $\mathcal{T}_n = \frac{\mathcal{T}}{\overline{\mathcal{R}}}$ , which is normalized by:

$$\overline{\mathcal{R}} = \frac{\int_c \mathcal{R} ds'_{\perp}}{\int_c ds'_{\perp}}$$

This is the average value of relatability along the symmetry. Similarly, we employ a threshold  $t_0$ . A valid part transition has  $\mathcal{T}_n > t_0$ .

The form equation (2.13) indicates that the local strength of the part transition is not determined by just one side of the shape boundary. The other side, across a continuous sequence of local symmetries, also matters. In particular, the factors include a weighted sum of curvatures  $\kappa(u)$  and  $\kappa(v)$  where the weights on both sides are proportional to their speeds in the symmetry tracking, which becomes even more obvious if (2.3) is rewritten as  $\frac{\kappa_u}{\kappa_v} = \frac{w_u}{w_v}$ .

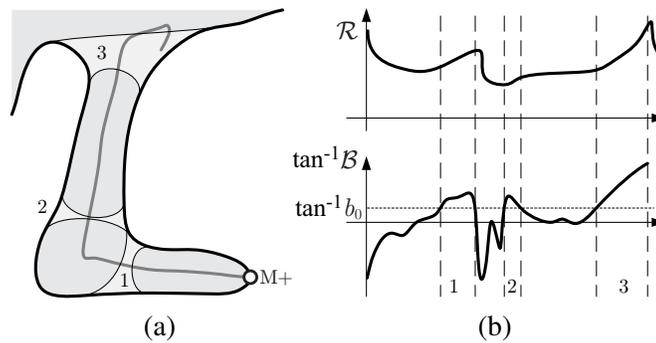
It naturally follows that the human visual system does not perceive a potential part transition



**Figure 2.19:** How both sides of the symmetry turn determines whether they form a part transition, and if so, how strong the transition is. Note that the boundaries of bend, taper, short-cut and limb share the same shape geometry and turning angle on their left side, so the boundary strength is determined by the right side.

if both sides of a local symmetry are convex. This corresponds to the “cup” shape, according to [Brady and Asada \[1984\]](#). See the first shape in [Figure 2.19](#) for one such example. The rest of the shapes in [Figure 2.19](#) all have concave boundaries. If the total weighted sum of the curvatures cancels each other out, the integral of the change in total turning angle is 0. A bending structure, where one side of the shape boundary is concave, is one such case—no section of transition is perceived here either. The convex side does not completely cancel out the concavity of the boundary it is symmetric with. But if it *almost* does so, and  $\mathcal{B} < b_0$ , then as we mentioned earlier, the transitional strength is so weak that the visual system does not perceive a convincing part transition. The “taper” shape, again from [Brady and Asada \[1984\]](#), is one typical example ([Figure 2.19](#)). Another such case is the shallow neck shown in [Figure 2.6\(c\)](#). Note that a neck is where the chordal length reaches a local minimum, hence (2.12) is positive, which indicates positive local transitional strength. If it is not significant enough, a shallow neck like this will not be perceived as a transitional region either. However, if the convex side is almost flat, such as the fourth shape in [Figure 2.19](#),  $\mathcal{B}$  will be significantly positive. If the integral of a continuous section of symmetries with  $\mathcal{B} > b_0$  is large enough, then this section is perceived as transition. This agrees perfectly with the short-cut rule proposed by [\[Singh et al., 1999\]](#). Of course, if both sides of the symmetry are concave, together they contribute to the transitional strength. Depending on their relative position, the part on one side of the transitional section might be perceived as a limb or a neck-based, which are the two cases proposed by [Siddiqi and Kimia \[1995\]](#).

Figure 2.18 suggests the implementation of locating sections of smoothed local symmetries that correspond to transitional regions between parts. We further clarify this procedure with

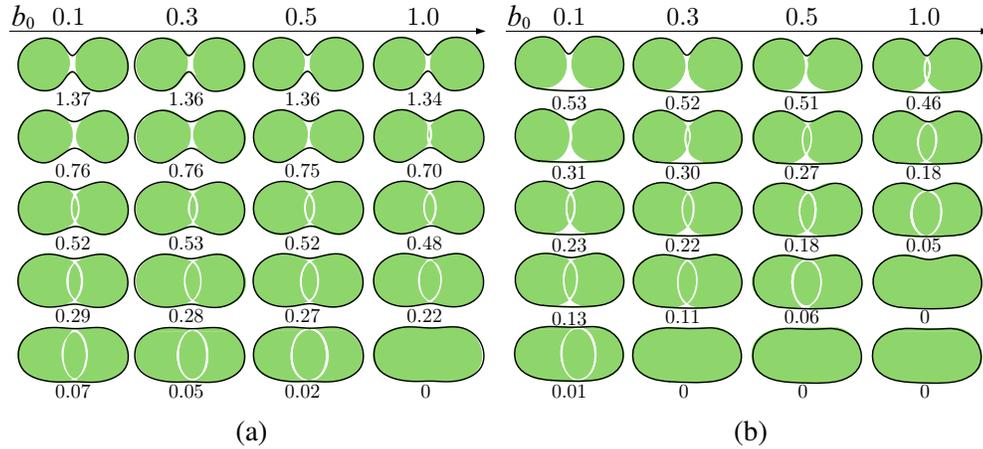


**Figure 2.20:** The reliability and local strength along the SLS (b) help to locate the transitional regions between parts (a).

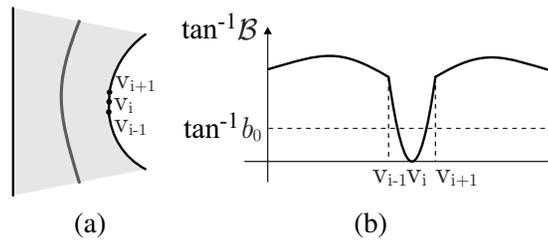
Figure 2.20. The algorithm begins with the tracing of the smoothed local symmetries from each curvature extremum, such as the SLS traced from the local positive curvature maximum from the toe of the foot shown in Figure 2.20(a). Then we compute the local transitional strengths of symmetries along the SLS, followed by locating continuous sections with  $\mathcal{B} > b_0$ . For each section, we compute the normalized total strength  $\mathcal{T}_n$ , and only consider sections with  $\mathcal{T}_n > t_0$ . As shown in Figure 2.20(b), the traced SLS contains three potential transitional sections. However, the total strength of section 2 is lower than  $t_0$ , thus is not considered as a part transition.

The two strength thresholds jointly determine the existence and extent of the transition between parts. Figure 2.21(a) shows the same set of shapes as Figure 2.18, but arranges them in rows. Each row in (a) has the same shape but uses different local transitional thresholds:  $b_0$  ranges from 0.1 through 1.0. As  $b_0$  increases, fewer local symmetries are included in the transition section, and thus the total transitional strength  $\mathcal{T}_n$  decreases. For a threshold of  $t_0 = 0.15$ , the last shape is rejected from being considered as a composition of two parts in any choice of  $b$ , since it lacks sufficient total transitional strength.

Figure 2.21(b) shows another group of shapes under the same threshold settings. In fact, these shapes are the results of “filling” the bottom indentations (from the background) of the shapes on the left. As expected, since only one side of the symmetry contributes to the local transitional strength of each shape, the total strengths are roughly below half of those in (a).



**Figure 2.21:** Effects of varying transition strength and  $b_0$ . Below each shape is the corresponding value of normalized total strength  $\mathcal{T}_n$  (or zero for no transition).



**Figure 2.22:** Split transition due to local curvature estimates.

Consequently, four of them are considered to be single-part shapes.

It is worth noting an implementation issue related to the sensitivity of the system due to curvature estimation. Figure 2.22(a) shows a perceptually continuous section of transitional region. However, the right boundary of the shape happens to have three consecutive vertices lining up, in which case, our curvature estimator evaluates as  $\kappa(v_i) = 0$ . Accordingly, its local symmetry has transitional strength 0, which breaks the transitional section into two pieces. Instead of trying to incorporate a wider window on the shape boundary in curvature estimation (This makes it difficult to trace the SLS properly, as is explained in Appendix 2.B), we estimate the local transitional region in a more robust way. For any local symmetry  $ds$  with chordal length  $L$ , we look at a neighborhood of size  $\frac{L}{5}$  along the SLS axis and use the median as the transitional strength.

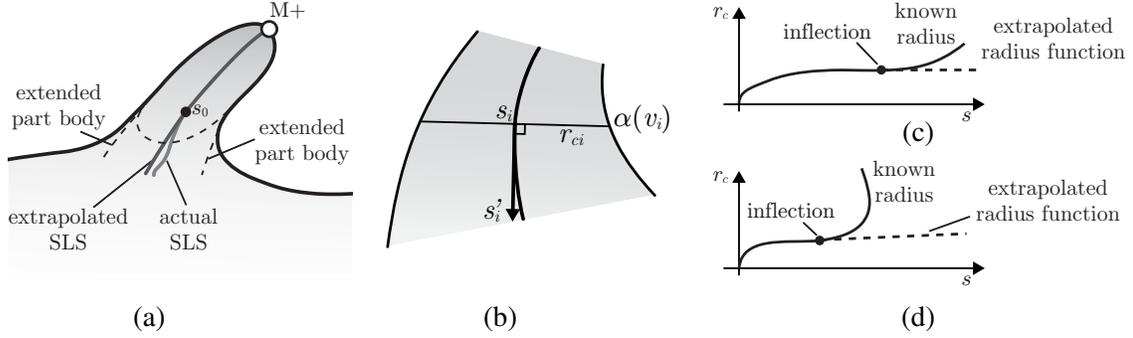
## 2.6 Extrapolation of part body using symmetry

We have described our part transition model which is used to locate the sections of smoothed local symmetries shared between parts. But it does not give the geometric description of the part body itself. Figure 2.7 shows how different forms a part body could possibly take and how different parts could be connected with each other differently. The parts can be closed as in (a,c) or open (b,d); they can be parts of the foreground (a,b) or the background (c,d).

Since the portion of the part body that “merges” into the transition is unknown, we can construct a reasonable connection by extrapolation. Following the model in Singh and Fulvio [2005] of the extrapolation of contours of 2D shapes by human vision, we use a logarithmic spiral to compute the unknown boundary of the part body. There are two cases; the first occurs when a “salient symmetry” is present. The SLS axis of a part forms a “salient symmetry” when the parallel component of  $s_{||}$  is small relative to  $s_{\perp}$ . If the local symmetry  $s_0$  where the SLS enters into the transition has  $\left. \frac{ds_{||}}{ds_{\perp}} \right|_{s_0} < 0.1$ , then we choose to extrapolate this coherent part of the SLS axis using a logarithmic spiral; but the chordal lengths along the axis are extrapolated linearly. The boundary of the part body is regenerated from the extrapolated axis and chordal lengths. Otherwise, in the second case, if we cannot locate any salient symmetries, we simply extrapolate the boundary using a logarithmic spiral.

Gray et al. [2006] describes logarithmic spirals in detail. The derivative of the curvature of the logarithmic spiral with respect to the arc length is given by  $\frac{d\kappa}{ds} = -b\kappa^2$ , where  $b$  is the constant term that determines how fast the curvatures decay along the curve. Note that logarithmic spirals are characterized by their self-similarity, and  $b$  is scale independent. We have verified that any  $b$  larger than 1 produces reasonable looking results (we simply set  $b = 2$ ). We have also tried to determine  $b$  based on the existing segment of the SLS axis. However, the start point of extrapolation could have the highest curvature along the whole axis, which makes  $b$  negative. This conflicts with the conclusions of Singh and Fulvio [2005]: human vision extrapolates a contour segment by a curve with decaying curvatures, which means  $b$  must be positive.

Figure 2.23 further explains the subsequent procedures. Each point  $v_i$  of the part boundary



**Figure 2.23:** Extending the body of a part is a four-step procedure. First, the SLS axis is extended (a); then part boundaries on both left (c) and right (d) sides are transformed into radius functions along the axis (b); At each side, the inflection point of the chordal radius function is located separately; the concave portion is abandoned, replaced with a linear extrapolation; Finally, the boundary is regenerated using the extended axis and chordal radii.

is then represented as an axial point  $s_i$  (with unit tangent  $s'_i$ ) and a chordal radius  $r_{ci}$ , such that

$$\alpha(v_i) = s_i \pm r_{ci} \cdot R_{90}s'_i \quad (2.15)$$

where  $R_{90} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$  is the rotation matrix that transforms the axial tangents to the perpendicular vectors towards the boundary points on the  $u$  side of the corresponding local symmetries. This map gives us the radius function along the axis. Recall the part body invariant described in Section 2.3. We want to redefine this function by getting rid of the transition—the convex portion. To do this, we simply locate the inflection point (after which the radius curve becomes convex) and extrapolate linearly from there, as shown in Figure 2.23(c) and (d). Then we regenerate the shape boundary using (2.15). Since now the radius is concave along the entire symmetry axis, there would not be any extended boundary with positive transitional strength.

In general, the inflection points of the radius curves of both sides are not mapped to the same axial point (see Figures 2.23(c) and (d)). Thus, we evaluate the chordal radii and extrapolate the radius functions using (2.15) separately. In fact, as indicated by Figure 2.23(a), the extrapolated axis does not align with the actual SLS axis from  $s_0$ . Beyond this axial point, the two extrapolated chordal radius at any axial point may not be the same. It is also possible that the whole radius along the axis is concave. In this case, the extrapolation starts from the last pair

of the transitional local symmetry. This typically happens to the non-concave side of the transitional boundary—the concave side across the symmetries makes the SLS transitional, like the “short-cut” case in Figure 2.19. If either of the radius extrapolation leads to negative values, we simply abandon extending the body, and close it with a cubic Hermite curve. A neck-based part such as in Figure 2.7(a) and (c) typically falls into this case. While some parts close up naturally, we can always close the body of a part even if it is extendable, as also shown in Figure 2.23(a). A part with a closed-up body has the description of its own form, while the part extension always implies the attachment to other parts. However, the part extension gives us flexibility in editing the way the part is attached, and so we always extend a part when possible.

## 2.7 Discussion

Our shape analysis algorithm tries to incorporate theories and findings from concurrent cognitive science research. With different previous theories and computational models proposed to interpret the composition of parts by geometric properties, our model fits them all into one simple framework. At the highest level, the design of the system is guided by theories and models in geometry and perception (the minima rule, the symmetry curvature duality, relatability).

Good contour continuity is one of the strongest sources of perceptual grouping. Since perceptually grouped contour segments are perceived as elements of the same visual component, any deviation in good continuity along the contour will thus be perceived as a result of separated components. This insight forms the base of our approach to separating parts from the rest of the shape.

A transitional model in a shape analysis system is as important as a lexical analyzer for language processing. It finds geometric units—parts. A shape-understanding system does not stop here, however. As the human perceives any shape as a well-organized multiple-part structure, we hope our system also organizes the parts into a structure that respects the perceptual organization by the human visual system. This is a challenging task even for simple polygonal shapes, since understanding many shapes involves resolving ambiguities in the representation. The next chapter reveals the details on how our system works towards this goal.

## 2.A Derivations of chordal geometry

This appendix gives detailed derivations of the local geometry about a pair of boundary points (not necessary a local symmetry) we used in Section 2.4.1 and Section 2.5.2. Again, throughout this appendix, the boundary of a simple 2D shape is represented by a  $C^2$  continuous curve  $\alpha$  with unit-length parametrization.

### 2.A.1 Derivatives of $L$

Recall  $p(u, v) = \frac{1}{L(u, v)}(\alpha(v) - \alpha(u)) = \frac{\alpha(v) - \alpha(u)}{\|\alpha(v) - \alpha(u)\|}$  (see Section 2.4.1). We have:

$$\begin{aligned} \frac{\partial L}{\partial u} &= \frac{\partial \|\alpha(v) - \alpha(u)\|}{\partial u} = \frac{\partial \sqrt{(\alpha(v) - \alpha(u)) \cdot (\alpha(v) - \alpha(u))}}{\partial u} \\ &= \frac{1}{2} \frac{-2 \alpha'(u) \cdot (\alpha(v) - \alpha(u))}{\|\alpha(v) - \alpha(u)\|} \\ &= -\alpha'(u) \cdot p = -\cos \theta \\ \frac{\partial L}{\partial v} &= \frac{\partial \|\alpha(v) - \alpha(u)\|}{\partial v} = \alpha'(v) \cdot p = \cos \varphi \end{aligned}$$

Thus, we get  $\nabla L = (-\cos \theta, \cos \varphi)$ . Note that  $\frac{\partial L}{\partial v}$  can be derived in a manner similar to deriving  $\frac{\partial L}{\partial u}$ ; or by reversing the order of local symmetry tracing and substituting  $u$  with  $v$ :

$$\frac{\partial L(u, v)}{\partial v} = -\frac{\partial L(v, u)}{\partial v} = -\cos(\pi - \varphi) = \cos \varphi$$

This “anti-symmetry” trick is used throughout this appendix.

### 2.A.2 Derivatives of $\theta$ and $\varphi$

To get higher order derivatives of  $L$ , we need to determine derivatives of  $\theta$  and  $\varphi$  for an arbitrary pair of points on the boundary. Since  $\theta$  cannot be expressed by a simple function due to the introduction of the wrapping function  $\varpi$ , we first derive the derivatives of  $\cos \theta$ . Considering that  $\nabla \cos \theta = -\sin \theta \nabla \theta$ , it follows that  $\nabla \theta = \frac{1}{-\sin \theta} \nabla \cos \theta$ .

We know  $\|\alpha'(u)\| = 1$  by the unit-length parametrization. It follows that  $\cos \theta = \alpha'(u) \cdot p$ .

We have:

$$\begin{aligned}\frac{\partial \cos \theta}{\partial u} &= \frac{\partial \boldsymbol{\alpha}'(u) \cdot \mathbf{p}}{\partial u} = \frac{\partial \boldsymbol{\alpha}'(u)}{\partial u} \cdot \mathbf{p} + \boldsymbol{\alpha}'(u) \cdot \frac{\partial \mathbf{p}}{\partial u} \\ \frac{\partial \cos \theta}{\partial v} &= \frac{\partial \boldsymbol{\alpha}'(u) \cdot \mathbf{p}}{\partial v} = \boldsymbol{\alpha}'(u) \cdot \frac{\partial \mathbf{p}}{\partial v}\end{aligned}\quad (2.16)$$

Now we need  $\frac{\partial \mathbf{p}}{\partial u}$  and  $\frac{\partial \mathbf{p}}{\partial v}$ . Since

$$\nabla \frac{1}{L} = -\frac{1}{L^2} \nabla L$$

We have:

$$\begin{aligned}\frac{\partial \mathbf{p}}{\partial u} &= \frac{(\boldsymbol{\alpha}(v) - \boldsymbol{\alpha}(u)) \cos \theta}{L(u, v)^2} - \frac{\boldsymbol{\alpha}'(u)}{L(u, v)} = \frac{\mathbf{p} \cos \theta - \boldsymbol{\alpha}'(u)}{L(u, v)} \\ \frac{\partial \mathbf{p}}{\partial v} &= \frac{-(\boldsymbol{\alpha}(v) - \boldsymbol{\alpha}(u)) \cos \varphi}{L(u, v)^2} + \frac{\boldsymbol{\alpha}'(v)}{L(u, v)} = \frac{\boldsymbol{\alpha}'(v) - \mathbf{p} \cos \varphi}{L(u, v)}\end{aligned}\quad (2.17)$$

To expand (2.16), we further write  $\boldsymbol{\alpha}'(u)$  in terms of the curvature  $\kappa(u)$ :

$$\frac{\partial \boldsymbol{\alpha}'(u)}{\partial u} \cdot \mathbf{p} = \boldsymbol{\alpha}''(u) \cdot \mathbf{p} = -\kappa(u) \mathbf{n}(u) \cdot \mathbf{p} = \kappa(u) \sin \theta \quad (2.18)$$

where  $\mathbf{n}(u)$  is the outward pointing normal vector at  $u$ .

Now all the elements are ready for  $\nabla \theta$ . Plugging (2.17) and (2.18) into (2.16), we get:

$$\begin{aligned}\frac{\partial \theta}{\partial u} &= \frac{1}{-\sin \theta} \left( \kappa(u) \sin \theta + \frac{\cos^2 \theta - 1}{L(u, v)} \right) = -\kappa(u) + \frac{\sin \theta}{L(u, v)} \\ \frac{\partial \theta}{\partial v} &= \frac{1}{-\sin \theta} \left( \frac{\cos(\theta + \varphi) - \cos \theta \cos \varphi}{L(u, v)} \right) = \frac{\sin \varphi}{L(u, v)}\end{aligned}$$

Considering that  $\tilde{r}(u, v) = \frac{L(u, v)}{2 \sin \theta}$  and  $\tilde{r}(v, u) = \frac{L(u, v)}{2 \sin \varphi}$ , we can simplify the expression of  $\nabla \theta$ :

$$\frac{\partial \theta}{\partial u} = -\kappa(u) + \frac{1}{2\tilde{r}(u, v)}, \quad \frac{\partial \theta}{\partial v} = \frac{1}{2\tilde{r}(v, u)} \quad (2.19)$$

Equation (2.19) gives both components of  $\nabla \theta$ . To get  $\nabla \varphi$ , we again use the ‘‘anti-symmetry’’ trick by considering the other side of the chord  $(u, v)$ . To use the expression of  $\nabla \theta$ , we exchange  $u$  and  $v$ , and replace  $\theta$  with  $\pi - \varphi$ . This gives  $\frac{\partial(\pi - \varphi)}{\partial v} = -\kappa(v) + \frac{1}{2\tilde{r}(v, u)}$ . Thus,

$$\frac{\partial \varphi}{\partial u} = -\frac{1}{2\tilde{r}(u, v)}, \quad \frac{\partial \varphi}{\partial v} = \kappa(v) - \frac{1}{2\tilde{r}(v, u)}. \quad (2.20)$$

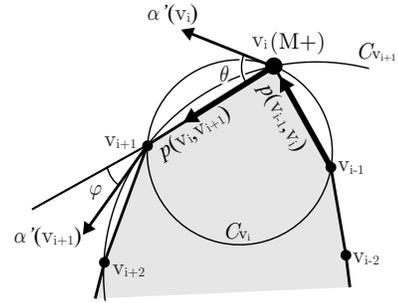
## 2.B Curvature estimation and symmetry-curvature duality on polygons

To estimate the discrete curvature at a vertex  $\mathbf{v}_i$ , we use the circumcircle formed by  $\mathbf{v}_i$  and its two adjacent neighbors  $\mathbf{v}_{i-1}$  and  $\mathbf{v}_{i+1}$  to approximate the osculating circle at  $\mathbf{v}_i$ . The curvature is simply the reciprocal of the radius of this circumcircle. Accordingly, the estimated curvature at  $\mathbf{v}_i$  is given by

$$\kappa(\mathbf{v}_i) = \frac{2 \sin(\frac{\phi}{2})}{\sqrt{\|\alpha(\mathbf{v}_i) - \alpha(\mathbf{v}_{i-1})\| \|\alpha(\mathbf{v}_{i+1}) - \alpha(\mathbf{v}_i)\|}} \quad (2.21)$$

where  $\phi = \arccos(\mathbf{p}(\mathbf{v}_{i-1}, \mathbf{v}_i) \cdot \mathbf{p}(\mathbf{v}_i, \mathbf{v}_{i+1}))$ ; and  $\alpha'(\mathbf{v}_i)$  is estimated as the osculating circle's tangent direction at point  $\alpha(\mathbf{v}_i)$ . This local geometry estimation scheme, as has been widely used in previous research [McCrae and Singh, 2009; Mullineux and Robinson, 2007; Tookey and Ball, 1997], makes the symmetry-curvature duality theorem applicable to curves represented by polygons. Considering the SLS tracing algorithm described in Section 2.4.2, we only need to show that  $\mathcal{A}(\mathbf{v}_{i-1}, \mathbf{v}_i)$  and  $\mathcal{A}(\mathbf{v}_i, \mathbf{v}_{i+1})$  must have different signs if  $\mathbf{v}_i$  is a local curvature extremum given that the curvature is estimated by equation (2.21). Here we only examine the case where  $\mathbf{v}_i$  is a positive curvature maxima by showing that  $\mathcal{A}(\mathbf{v}_{i-1}, \mathbf{v}_i) < 0$  and  $\mathcal{A}(\mathbf{v}_i, \mathbf{v}_{i+1}) > 0$ . Other cases can be derived similarly.

Figure 2.24 shows the osculating circles,  $C_{\mathbf{v}_i}$  and  $C_{\mathbf{v}_{i+1}}$ , of vertex  $\mathbf{v}_i$  and  $\mathbf{v}_{i+1}$ , respectively, as well as the chordal geometries between these two vertices. By definition,  $\mathcal{A}(\mathbf{v}_i, \mathbf{v}_{i+1}) = \theta(\mathbf{v}_i, \mathbf{v}_{i+1}) - \varphi(\mathbf{v}_i, \mathbf{v}_{i+1})$ . Note that according to the inscribed angle theorem,  $\theta$  is the same as the inscribed angle of  $C_{\mathbf{v}_i}$  subtended by the chord  $\mathbf{v}_i\mathbf{v}_{i+1}$ , and  $\varphi$  is the same as the inscribed angle of  $C_{\mathbf{v}_{i+1}}$  subtended by the same chord. Since  $\kappa(\mathbf{v}_i) > \kappa(\mathbf{v}_{i+1})$ , the radius of  $C_{\mathbf{v}_i}$  is smaller than that of  $C_{\mathbf{v}_{i+1}}$ . It follows that  $\theta > \varphi$ , which means  $\mathcal{A}(\mathbf{v}_i, \mathbf{v}_{i+1}) > 0$ .  $\mathcal{A}(\mathbf{v}_{i-1}, \mathbf{v}_i) < 0$  can be shown in a similar manner.



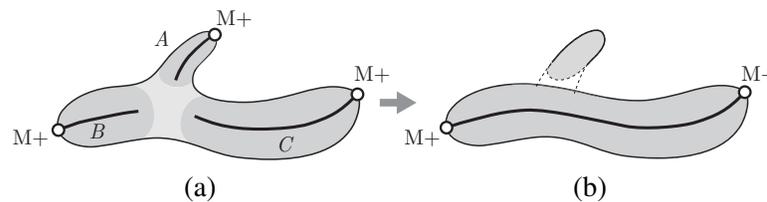
**Figure 2.24:** Osculating circles and chordal geometries around  $\mathbf{v}_i$

## Chapter 3

### Shape Parsing in terms of Parts

For any given shape we can find all the basic parts using the model proposed in Chapter 2, but the set of geometric components includes more than just parts. Each of them might be a real part used by the human visual system to describe the shape, such as part *A* in Figure 3.1(a); it might also be only a portion of a part, such as *B* and *C* in Figure 3.1(a). Together they are perceived as a single part, since the most natural description of the whole shape would probably be this part with a protrusion, part *A*, on the top, as depicted in (b). We treat part-like elements such as *A*, *B* and *C* as determined using the part body invariant from Chapter 2 as *potential part candidates*.

This chapter describes our method of identifying real parts from among these part candidates and constructing a comprehensive, structural description of the shape in terms of these parts. Our shape parser follows morphological frameworks [Rom and Médioni, 1993; van der Poorten and Jones, 2002] that identify and delete parts from the whole shape one by one in the order of their sizes. The set of deleted parts are used collectively to describe the shape. Section 3.1 will introduce this approach in detail, with an improved part separation scheme thanks to the transitional model described in Chapter 2. However, Section 3.2 shows a few problems with this



**Figure 3.1:** Not every populated element is perceived as a real part. Of the three candidates in (a), only part *A* is recognized as a part by the human visual system. *B* and *C* are explained as portions of a bigger part, shown in (b)

approach. In particular, it proposes taking both parts of the shape and parts of its background into account. It also points out that using only size to order the deletions can produce parts that do not respect the perceptual organization. Furthermore, when we use this description as a representation for abstracted depiction in Chapter 4, the resulting abstraction appears distorted. Section 3.3 and Section 3.4 proposes a set of heuristics as our models of perceptual organization. These models are implemented using constraints that restrict deletions that would disrupt the perceptual organization. A graph-based algorithm is described in Section 3.5 to implement these constraints for the set of deletable parts. In Section 3.6, we further demonstrate that this approach can be adapted to allow a user to guide the decomposition using strokes drawn on the shape. This approach has been presented in [Mi et al., 2009] and here we describe it in more detail.

### 3.1 Shape analysis by sequential part pruning

In Figure 3.1(a),  $B$  and  $C$  become part candidates of the whole shape because of the existence of part  $A$ .  $A$ 's transition into the base forms contour segments of negative curvatures. Because of these concave curve boundaries, the smoothed local symmetries traced from the lower two positive maxima also become transitional, and these transitional SLS axial segments give rise to the candidates  $B$  and  $C$ . If only  $A$  were not there!

The description of the shape in the beginning of this chapter hints that in order to arrive at a correct explanation about candidates  $B$  and  $C$ ,  $A$  would need to disappear. This naturally leads to the paradigm of “shape analysis by pruning”: if we can somehow cleanly delete part  $A$ , then the negative curvatures between  $A$  and the base formed by candidates  $B$  and  $C$  are gone. The SLS axes of  $B$  and  $C$  no longer contain any transitional sections. Instead, they are merged as a complete part (Figure 3.1(b)), which correctly describes the base of the shape.

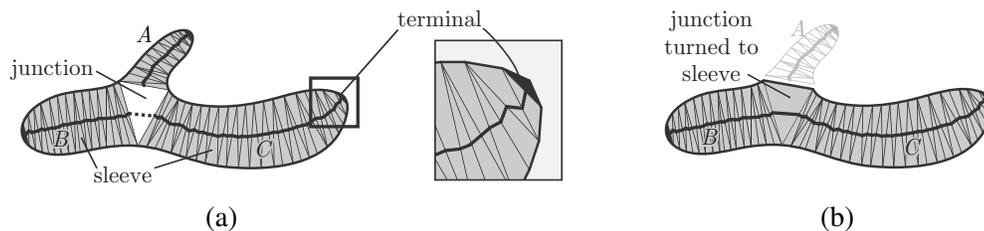
#### 3.1.1 Related work on structural pruning

The medial axis pruning described in Chapter 1 is meant for shape simplification. The pruned axis of the simplified shape does not change the relationship between retained medial edges.

Compared to this, the *chordal axis transformation* and morphological pruning of the chordal axis are more related to the shape analysis.

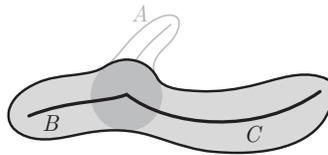
The chordal axis transformation (CAT) is very similar to the MAT. They both recover the boundary of the shape using the same set of local symmetries—those on which the bitangent circles are maximally inscribed [Leyton, 1992]. The difference is only in what generates each pair. For the MAT, the generator is simply the bitangent circle; for the CAT, the generator is the bitangent circle only when this circle is shared with other symmetry pairs; otherwise, the CAT takes the local symmetry chord as the generator. In the latter case, the center of the generator lies exactly on the SLS axis, which means that this portion of the chordal axis is a subset of the SLS axis.

For shapes represented by polygons, in Prasad's [1997] implementation, the CAT pruning is based on the discrete CAT, which is simply the CAT with the discrete set of local symmetries. Specifically, a pair of vertices of a polygon is considered locally symmetric in terms of the discrete CAT if and only if they form a constrained Delaunay edge in the constrained Delaunay triangulation (CDT) [Chew, 1989] of this polygon with edge connectivity constraints. Thus, topologically, computing the CDT is the essential step in the computation of the discrete CAT. The CDT divides the shape into so called *constrained Delaunay triangles*, each of which can be one of four types, according to the number of polygon edges they share. An *isolated triangle* is a Delaunay triangle whose three edges are all polygon edges. Since this only happens with triangular polygons, we choose to ignore this type of polygon since the analysis is trivial. A *terminal triangle* shares two polygon edges; a *sleeve triangle* shares one, while a *junction* shares none—all its edges are internal Delaunay edges.



**Figure 3.2:** (a) shows the CAT of the shape in Figure 3.1. Each of the three sleeves is led by a terminal triangle. They can be related to parts *A*, *B* and *C* naturally. (b) shows that pruning the sleeve on the top from the CAT simplifies the topology of the sleeve structure and the chordal axis, but this pruning also suffers from the stump artifacts.

Figure 3.2(a) is an example of the CAT. The only junction is shown in white; sleeve and terminal triangles are shown in medium and dark gray, respectively. The virtue of the discrete CAT is that major local discrete symmetries are naturally represented by chains of adjacent sleeve triangles (in the rest of this chapter, we simply call them *sleeves*). When a sleeve is adjacent to a terminal triangle, we say that the terminal triangle *leads* the sleeve. Considering the connection between the CAT and the SLS, a sleeve of the discrete CAT led by a terminal triangle can often be related to the smoothed local symmetries of a basic part candidate traced from some curvature extremum<sup>1</sup>. In Figure 3.2(a), *A*, *B* and *C* can each be related to a sleeve led by a terminal triangle, and these sleeves join at a junction triangle. Figure 3.2(b) further shows that once the sleeve related to *A* has been deleted, the sleeves related to candidates *B* and *C* are now combined into a single part. In fact, this junction becomes a sleeve triangle if any one of its adjacent sleeves (each led by a terminal triangle) is deleted. After the deletion of any one sleeve, the other two sleeves are merged, and the sleeve structure of the whole shape becomes simplified, as pointed out by Prasad [1997] and applied by van der Poorten and Jones [2002] in their map simplification system. Prasad and Rao [2000] also point out that this structural removal implies the hierarchy of the shape. In the case of *A*, *B* and *C* in Figure 3.2(b), removing the top sleeve and leaving a whole more salient sleeve structure suggests interpreting the shape as: a single sleeve containing both *B* and *C*, with a branch that contains *A*.



**Figure 3.3:** The simplification of the topology of the axis structure, as well as the stump artifact of the MAT pruning

However, removing *A* does not lead to a clean, smooth sleeve. There are often stumps left after the pruning of discrete CAT, as pointed out by van der Poorten and Jones [2002]. In the domain of continuous chordal axis pruning, such stump artifacts were noticed even earlier

<sup>1</sup> In most well sampled polygons, one of the vertices of a terminal triangle is also a curvature extremum, and a smoothed local symmetry can be traced from it. But unfortunately, this connection between the smoothed local symmetries and sleeves in the discrete CAT is not guaranteed.

by Rom and Médioni [1993] in their work on morphological 2D shape analysis. As shown in Figure 3.3, the regenerated shape with the top medial edge pruned also exhibits a similar stump artifact, although in this case, the stump has a round form. As shown in Figure 3.3, the junction of the three medial edges is the result of the existence of all three part candidates  $A$ ,  $B$  and  $C$ . Portions of the medial edges of  $B$  and  $C$  that are close to part  $A$  are also affected by  $A$ . In the shape with  $A$  pruned, however, the regeneration is based on the medial edges of  $B$  and  $C$ , which are “contaminated” by  $A$ . In other words, the medial edges are computed based on the whole shape, but these medial edges never get reevaluated during the pruning. This results in the “stump” artifacts of the MAT pruning. Given the equivalence between the MAT and CAT, shape simplification by CAT pruning shares similar problems with the medial axis pruning, including the stump artifacts and the unrepresented shallow bumps (recall Figure 2.2 in Section 2.1.1), as well as structural instability<sup>2</sup>.

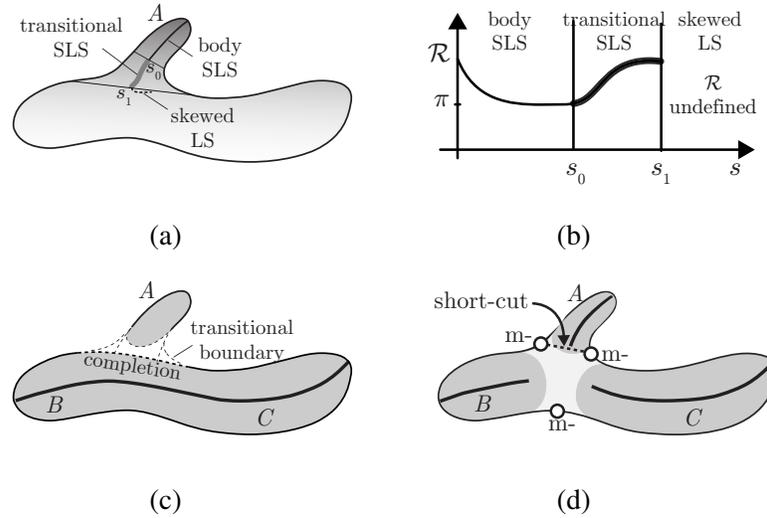
### 3.1.2 Shape analysis with clean removal of parts

The existence of stump artifacts results from the failure to remove  $A$  cleanly. Rom and Médioni [1993] attempt to eliminate this problem by pruning more of the boundary by a fixed amount, which achieves limited success. Our transitional part connection model provides a more elegant solution to this issue.

Figures 3.4(a) and (b) show the SLS axis starting from the very tip of the shape, and the reliability  $\mathcal{R}$  of the local symmetries along this axis. The transitional region, indicated in thick gray, signals the existence of a potential part candidate, which is  $A$  in the figure. As described in Section 2.5.4, this transitional section is detected by the interval with significant positive values of  $\mathcal{B}$ . If we were to follow the traditional binary segmentation of the shape in terms of parts, any chord of local symmetry along this section would be reasonable. In this case, however, a stump would remain, revealing traces of the prior existence of  $A$ , except for the last pair (as indicated by  $s_1$ ) where the transition is completely removed.

---

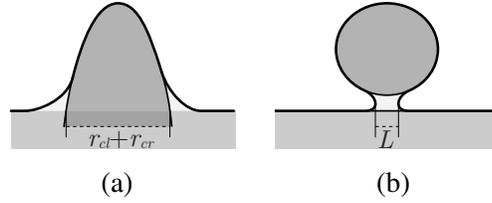
<sup>2</sup> Prasad [2007] tries to resolve the structural instability of the discrete CAT by introducing the approximate co-circularity of edge-adjacent triangles, which shows an idea very similar to the bone graph proposed by Macrini et al. [2008].



**Figure 3.4:** (a) The transitional SLS of candidate  $A$ . (b) The corresponding values of reliability along the axis. (c) A clean part removal. (d) The decision to delete  $A$  is based on the short-cut rule, or rule of proximity.

Following the removal of part  $A$ , instead of completing the shape with a line segment, we use a cubic Hermite spline with  $C^1$  boundary continuity. In our implementation, we have also tried other interpolative curves, including the Euler spirals [Kimia et al., 2003] and curves of elastica [Mumford, 1994]. We do not see significant differences between the cubic spline and the elastic interpolations. Due to the local symmetry of the chord, the Euler spiral is always a circular arc, which we find to be visually distracting in some occasions.

The short-cut rule naturally justifies our choice of  $A$ , rather than  $B$  or  $C$ , as the part to be deleted. The short-cut rule [Singh et al., 1999] points out that closer local symmetry triggers stronger contour grouping. This suggests that parts should be deleted in the order of their *scale*. We use the chordal length of the local symmetry as our scale measure. Note that since the chordal length of  $s_1$  does not represent the size of the part body, but rather that of the transition, this length is affected by how the part joins the main body of the shape. It would be more reasonable to use the scale of the part body to order the deletion. Thus, instead of using this radius, we evaluate the extrapolated radii at  $s_1$  on both sides and order the part deletion by their sum. For a part that is not extendable, which is typically neck-based, we use the smallest chordal length of the transitional region. See Figure 3.5 for illustrations of both cases. Notice that for extrapolated part candidates, our definition of the scale measure is independent of the



**Figure 3.5:** For extendable parts, we use  $\tilde{L} = r_{cl} + r_{cr}$  as the scale measure, where  $r_{cl}$  and  $r_{cr}$  are the extrapolated left and right radii at the last symmetry pair (a). If such extension does not exist, we simply use the minimal chordal length of the transitional section. For a neck-based part, this is just the distance across the neck.

forms that the transitions take.

As pointed out by [Rom and Médioni \[1993\]](#), the deleted components naturally form a hierarchy according to the relationship between the completion curve after a previous deletion and the deleted boundary of a subsequent deletion. We can also build this relationship between parts, but in a little more complicated way due to the introduction of the transitional region. For two subsequently deleted parts  $P_1$  and  $P_2$  where  $P_1$  is deleted before  $P_2$ ,  $P_2$  is considered the parent of part  $P_1$  only if the *body* of  $P_2$  shares any segment of curve that has been used to complete the shape after the deletion of  $P_1$ . The transitional boundary of  $P_2$  sharing  $P_1$ 's completion curve does not qualify  $P_1$  to be a parent part of  $P_2$ . If this is the case, we assign the parent of  $P_1$  as  $P_2$ 's parent.

For neighboring parts with similar scale measures, our system opts to delete them together, in part because it is often a better arrangement to consider neighboring parts at the same level in the shape hierarchy rather than to treat one as the parent of the other, which is implied by the sequential deletion. At this moment, we simply consider that two scale measures  $\tilde{L}_1 > \tilde{L}_2$  are similar if  $\tilde{L}_1 < 2\tilde{L}_2$ . In [Section 3.3.5](#), we will revisit this issue and come up with a better justification for multiple part deletion. For multiple neighboring parts, we simply use the smallest scale as the scale of the whole group, and we use a single Hermite curve to complete the gap left after the deletion of these parts.

We summarize the procedure for computing the shape description, given a simple 2D shape, as follows:

### Part detection

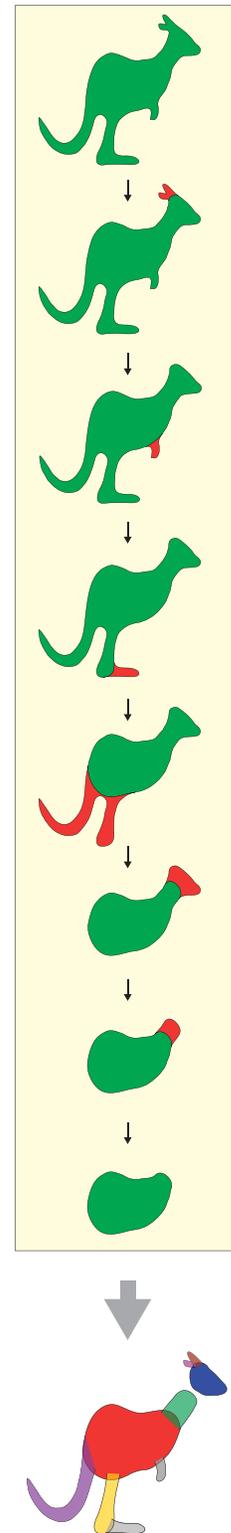
1. Trace local symmetries from each curvature extremum including those with negative curvatures. At this moment we only consider parts from the foreground, and most foreground parts are described by SLS axes traced from positive curvature extrema. But in some situations, a single SLS axis that starts at a negative curvature minimum may continue on to describe symmetries of a part in the foreground.
2. Locate transitional sections of the SLS with significant strength  $\mathcal{T}_n > t_0$ . If any end of the transition correspond to a basic part candidate, build the data structures related to this candidate based on the transition and the SLS on that end, including the necessary part extension, where possible. Note that a basic part candidate does not contain any other potential parts.

### Part decomposition

3. Find the group of neighboring parts with the smallest scale as measured by the smallest part in the group. Replace their boundary with the completion curve.
4. If the shape is not empty, go to step 1.

The whole set of deleted parts, as well as the hierarchy the deletion implies, collectively constitutes the shape description in terms of parts.

We apply a sequence of part deletions to a kangaroo shape, which gives us the step-by-step results as well as

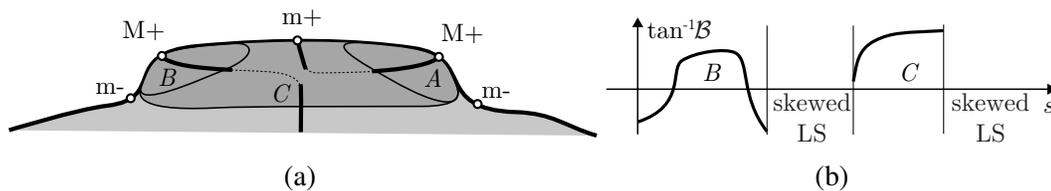


**Figure 3.6:** Sequential deletion of parts reveals the composition of the kangaroo shape

the final decomposition in Figure 3.6. Note that in the bottom figure, extendable parts overlap with their parent parts. Examples of deletions that involve multiple parts can also be found in the first and fourth steps. Because of this arrangement, the two ears of the kangaroo are both subparts of the head. The leg and the tail are deleted together as a group due to their similarity in scale. Since their transitional boundaries overlap, the hierarchical relationship built according to Rom and Médioni [1993] will consider the leg as a subpart of the tail if the former is deleted after the latter.

### 3.1.3 Termination condition for SLS tracing

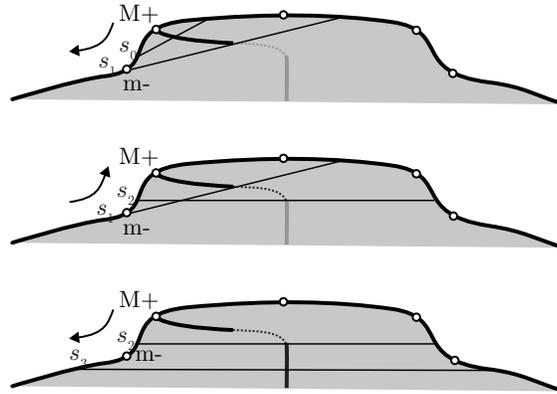
Section 2.4.3 has portrayed the high complexity of the SLS axis, together with the good news that all we need to focus on are the local symmetries related to the basic parts. Thus, instead of locating transitional sections and basic parts after the whole set of SLS is traced out, we do so on the fly. We stop tracing when we do not expect there could be any more basic parts along the rest of the SLS. The problem now is to determine when we do not have to trace any further. An immediate answer would be: stop tracing once we locate a section of transition with significant total strength  $\mathcal{T}$ .



**Figure 3.7:** Tracing SLS stops when no other basic part is expected. Note again that at sections of skewed local symmetries,  $\mathcal{R}$  and  $\mathcal{B}$  are not defined. The bodies of the parts are closed, instead of extended, for better visualization.

However this simple solution might fail. Figure 3.7 illustrates this with a shallow, flat hump. Along the boundary of the hump there are five curvature extrema. Two of them are negative curvature minima, which turn out to be irrelevant in discovering the part of the hump when tracing the SLS from them.

The SLS from the right curvature maximum contains a transitional region which predicts a potential part candidate at corner A. Then the local symmetries proceed in a skewed way,



**Figure 3.8:** Keep tracing until at least one side of the SLS stably passes a segment of boundary curve with negative curvatures

shown by the dotted curve, which is skipped according to the convention, previously explained in Section 2.5.2. The SLS terminates at the positive curvature minimum in the middle, which means that tracing from this extremum is no longer necessary. The local symmetries from these two extremum are thus not relevant to the hump either.

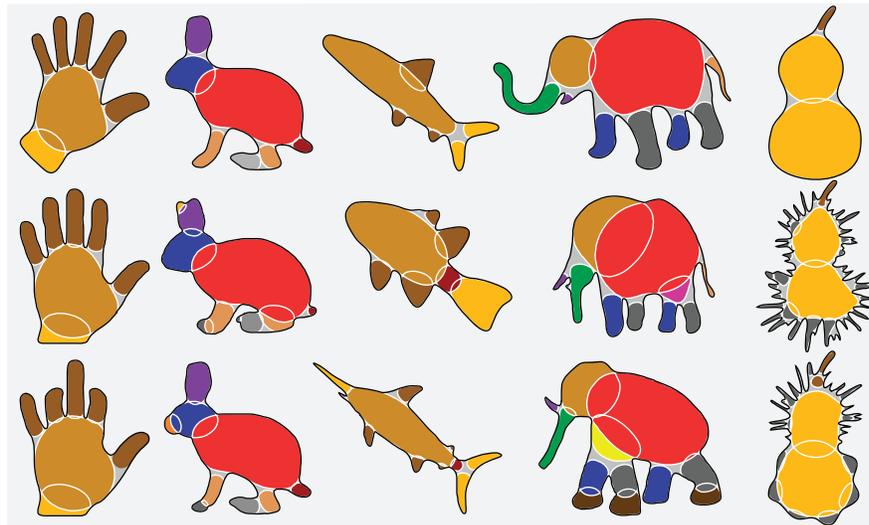
The SLS tracing from the left curvature maximum behaves similarly in the beginning—an immediate transitional section of SLS indicates a potential corner candidate  $B$  followed by a sequence of skewed local symmetries. However, after that the local symmetries resume normal continuation and enter into another transitional section. As shown in Figure 3.7(b), this section corresponds to the hump  $C$ . We would not have discovered this hump if we had followed the simple solution and stopped at  $B$ .

Observe the movement of the left end of the SLS chord when being traced from the left curvature maximum (let's call it the *tip* here). It first goes counter clockwise and passes the left negative curvature minimum ( $s_0$  to  $s_1$  in Figure 3.8); as the tracing proceeds in a skewed manner, it “shrinks” back towards the tip ( $s_1$  to  $s_2$ ). After that, the local symmetries enter the transitional section of  $C$ ; the left end passes the left minimum again. But this time, it does not shrink back ( $s_2$  to  $s_3$ ).

In our system, after we have discovered a transitional SLS, we look ahead to see if the symmetries proceed in a skewed fashion. If this is the case, and the skewed local symmetries shrink back, we keep tracing to discover more transitional sections.

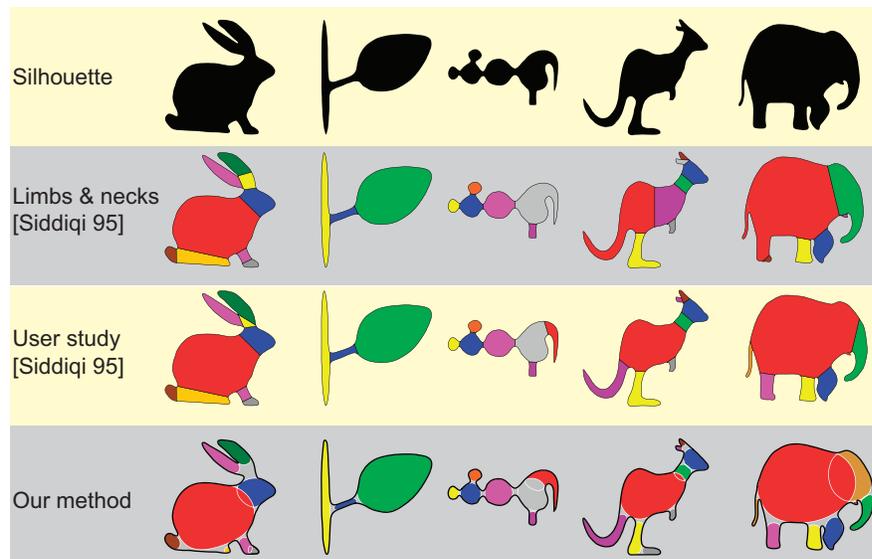
### 3.1.4 Results

With all the elements in place for the first version of the shape analysis system, we are ready to test it on some shapes. We have assembled five sets of shapes, each consisting of three shapes of the same category, on which we think our shape parser described thus far would do a reasonable job. Figure 3.9 shows the resulting decompositions, arranged column-wise. In particular, shapes on the first four columns are from Kimia’s 99 dataset [Sebastian et al., 2001; 2003]. Observe the heads of the elephants, the legs and feet of the rabbits, and the fingers and wrists of the hands. The body of each decomposed part takes its own form. It may overlap or be disjoint with the bodies of the neighboring parts. Finally, the transitional regions glue all these parts together into a coherent shape. In the right-most column are three gourd shapes. Two of them are contaminated with noise of varying magnitudes. Since our system does not specifically handle noise, these bumps and protrusions are treated as parts. Note that the major recovered part of the lower compartment of the middle gourd loses its original roundness. This happens in part because at the moment our system only considers foreground parts, but the noise contaminates the boundary in both inward and outward directions. Indeed, our system is not meant to handle noise. Curvature flow does a better job here.



**Figure 3.9:** Results of shape composition in terms of parts. The part bodies are shown in colors; transitional regions are shown in light gray.

Figure 3.10 further compares the parts of our model with the shape description in terms of



**Figure 3.10:** Comparison between binary part segmentation by Siddiqi et al. [1995] and shape description in our part model, which explicitly describes their bodies and the transitions in between. Also notice that Siddiqi et al. [1995] only considers two particular types: limbs and neck-based parts. Some parts are missing in their results.

part segmentation [Siddiqi et al., 1995]. The comparison between our results and those of Siddiqi et al.'s reveals that our model discovers not only limbs and neck-based parts, but also some others that are not described in the previous approach. For example, we successfully decompose the foot and the tail of the kangaroo, and the trunk of the elephant. Our system makes specific distinctions between part bodies and transitions. While in the segmentation of Siddiqi et al., each region of the binary segmentation has to contain some of the transition. It seems to be the representation scheme itself that inevitably limits the power of a shape segmentation system in describing the forms of overlapping parts, such as the heads of the rabbit and the elephant.

### 3.2 Valid and invalid deletions

The first version our shape parser has demonstrated a capability to describe parts in a more meaningful way. To further assess its performance in analyzing shape structure, we test our procedure on some other shapes, and consider the role of background parts in shape analysis. These tests further demonstrate the need for a method that orders parts for removal that does

not simply look locally on the shape (using measures such as size).

### 3.2.1 Background parts

The analysis of the shark shape in Figure 1.5(a) begins with the deletion of a portion of the top tail fin as shown in Figure 3.11(b), because it narrows down and forms a neck in the middle due to the indentation. Notice that the word “indentation” reveals the most natural explanation of the cause of the negative curvature minimum labeled in Figure 3.11(a). Recall that in Section 2.3 we demonstrated three possible causes of negative curvatures along the shape boundary, and we have chosen to ignore the subtractive composition here (refer to Figures 2.7(c) and (d)). Although the shape analysis could still proceed beginning with the deletion of candidate *B* shown in Figure 3.11(b), the resulting description of the top tail fin would end up looking very strange. Here, considering candidate *A* as a part of the background seems to be the only plausible way to match the most salient perceptual organization of this region.

Considering parts from both the foreground and the background implies that perceptual figure/ground assignment is not globally consistent<sup>3</sup>. This is supported by recent studies on local geometric factors that affect the assignment and saliency of perceptual parts and, in turn, figures [Hoffman and Singh, 1997; Kim and Feldman, 2009].

### 3.2.2 Part scale, area and contour grouping

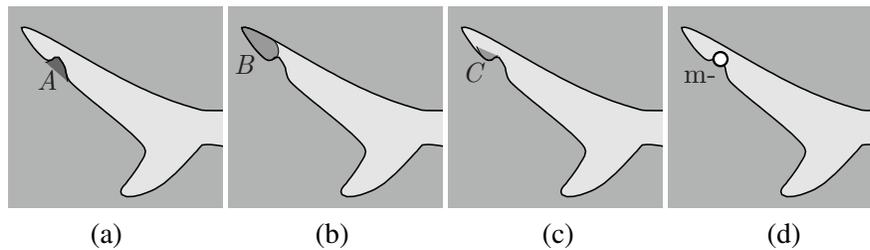
However, taking parts of both foreground and background into account does not necessarily make the indentation the first potential candidate to be removed. The base of the indentation is wider than the neck on the foreground it causes. This again would cause the top half of the upper fin, candidate *B*, to be deleted next.

It is perfectly reasonable to doubt the validity of the deletion order based on the scale, or to accept the rationale of the scale order but to doubt the scale measure based on the chordal

---

<sup>3</sup> Most traditional research assumes that figure/ground assignment is globally consistent along the entire shape contour, meaning that, for the entire shape, human vision interprets one side of the boundary curve as figure and the other side as ground. However, local geometric cues are proven to be influential to the figural assignment. Accordingly, the border ownership may locally reverse, creating an inconsistency in figure/ground assignments along the contour [Kim and Feldman, 2009; Peterson and Gibson, 1994].

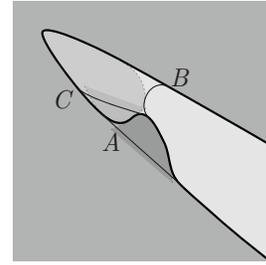
length of the base. An immediate response would be to try the area measure, which seems to be a commonly considered criterion to order deletions, according to previous research in medial or chordal axis pruning [Prasad, 1997; Rom and Médioni, 1993; Shaked and Bruckstein, 1998]. Unfortunately, although indentation  $A$  has a smaller area than that of  $B$ , the area of the little corner candidate  $C$  is even smaller. If ordering the deletion by area,  $C$  would be deleted first, leaving a wider indentation. Continuing the deletion in this order would fill the wider indentation. This new step is acceptable since in the final description of the shape, it explains this part of the shape as a whole fin minus an indentation. However, the description also includes corner  $C$ , which implies that after the deletion of the wide indentation, a corner grows from the foreground, making the indentation narrower. This is not intuitive. Consequently,  $C$  does not seem to be an intuitive part in its shape context either. Deleting parts in the order of area might be a valid approach to simplify geometry because the area difference is locally minimized (in a greedy manner), but for a shape description system that works by removing perceptual parts one by one, it is also flawed.



**Figure 3.11:** The first potential candidate to be removed should be  $A$ , from the background (a). Ordering by extended chordal length would select  $B$  (b), and part  $C$  would be chosen by area. Neither is ideal. The system populates them as valid basic part candidates because of the negative curvature minima (d)

As we have pointed out in Section 3.1.2, although the tracings of SLS populate significantly more part candidates, only deleted elements are parts used to describe the shape. This makes the deletion a non-compromising task—every deleted candidate will be associated with a perceptual unit, explaining the formation of the shape in the final description [Leyton, 2001]. If a deletion does not turn out to be associated with any shape unit naturally, it is an *invalid* deletion. Of course, our goal is to make as many valid deletions as possible.

We can give different reasons why the deletion of the corner  $C$  in Figure 3.11(c) does not make sense perceptually. One argument is to defend the scale-based deletion order. Recall that we chose this order by the short-cut rule, which is well grounded in gestalt psychology about proximity.



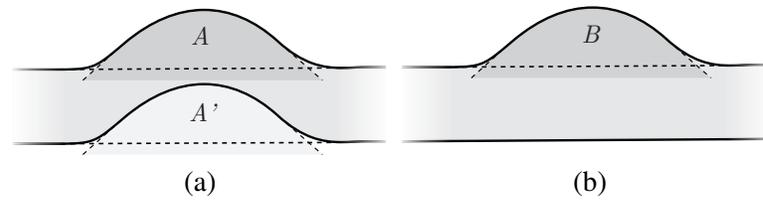
**Figure 3.12:** Part  $A$ 's completion curve connects more strongly grouped contour segments.

Figure 3.12 offers a closer observation of the geometry around the top tail fin. It also shows the completion curves of the candidates  $A$ ,  $B$  and  $C$ . Now consider the question: of all three pairs of contour segments connected by these completion curves, which pair is most strongly grouped? It seems easy to agree that the answer is  $A$ . The short-cut rule is still valid, but the distance measured in Euclidean space, such as the chordal length can align badly with the perceptual distance.

However, instead of attempting to improve the perceptual grouping measure, we decide to stick to the chordal length across the symmetry to order the deletion for two reasons. First, the perceptual grouping of contour segments is affected by many factors. We have already incorporated the most prevailing geometric cues, including the cocircularity (by separating parts only across local symmetries), collinearity (with our model of part transition) and distance (implied by the deletion order). There are other related cues, however, including curvature, orientation, etc. We have yet to see a systematic, quantitative study of a robust measure of contour grouping strength that takes all related factors into account.

Second, even if we have a perfect grouping measure, there are still other factors that affect the salience of a potential part. As further demonstrated by Figure 3.13, although the candidates  $A$ ,  $A'$  and  $B$  have exactly the same geometries of their bodies and transitions (except that  $A'$  is a candidate of the background), only  $B$  is a good candidate for deletion. The shape of (a) becomes that of (b) after  $A'$  is deleted. Including  $A'$  in the description means that (a) is the result of an indentation made to the shape (b) from the bottom, which does not sound very persuasive.

The lesson learned here is that deletion of a part candidate is not a decision that can be made solely on its local geometry. Its neighboring candidates, the global structure of the shape



**Figure 3.13:** Neither  $A$  nor  $A'$  is a good candidate for deletion since (a) is interpreted as a bending structure.  $B$  is a salient part, since it is interpreted as a hump on (b).

surrounding it, as well as the shape before and after the deletion, are the factors that should be taken into consideration.

The reason relates to recoverability theory [Leyton, 1992] described in Section 1.2.4. The human visual system does not perceive Figure 3.13(b) as a symmetrized version of (a), thus, (a) does not come from (b), which invalidates the deletion of  $A'$  as a part. Similar arguments can be made against the deletion of candidate  $A$ . This is another reason why the deletion of candidate  $C$  in Figure 3.11 is not appropriate—the top tail fin does not become more symmetric because of this deletion.

### 3.3 Heuristics to prevent symmetry breaking

Choosing the next part to delete in our system involves a difficult decision. The concerns described in the previous section shed light on resolving the ambiguity and lead to our shape parsing strategy by introducing a set of heuristics as constraints on the deletion. These constraints are imposed so that the theory of recoverability and the explanation towards the shape structure by the human visual system are strongly encouraged. This section and Section 3.4 describe these constraints.

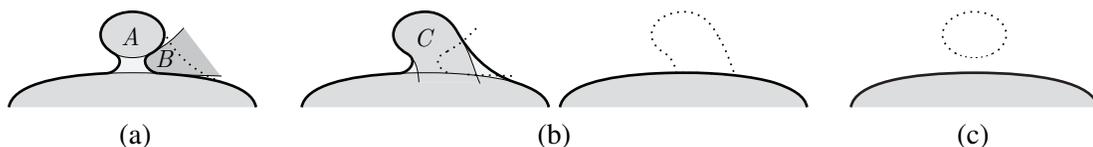
#### 3.3.1 Prevent distortion of symmetry of neighboring parts

The main point of the theories of causality and recoverability, is that the human visual system only perceives *symmetry-breaking* events of the shape. Attaching a part to a shape does not typically make it more symmetric. Accordingly, when parsing a shape deleting parts, the shape should only become more symmetric as we remove parts. Thus, when selecting parts for

deletion, a valid part deletion does not distort existing symmetries.

Thus in order to assess a part removal, we need to measure how removing a part in fact distorts the symmetry of other parts. Unfortunately, we lack a principled measurement of how symmetric a general shape is. Sebastian et al. [2003] presents a shape categorization system by aligning contour curves, but their algorithm is more suitable for matching curve segments. It does not seem to be straightforward to use their system here. Instead, we introduce a heuristic to constrain part deletions.

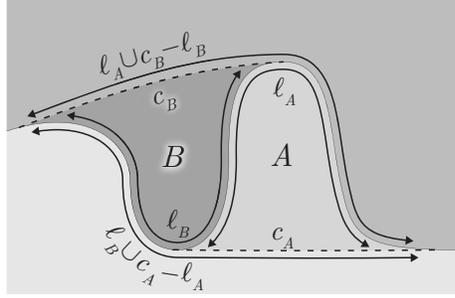
We start with an example shown in Figure 3.14, where (a) shows a neck-based part. Since we are considering parts from the background,  $B$  is also recognized as a potential part candidate. However, the explanation suggested by deleting  $B$ —describing  $A$  as an indentation made to  $C$ —is not persuasive to the human visual system since this step does not make  $A$  more symmetric but only distorts its axial structure. As shown in Figures 3.14 (b) and (c), a successive deletion of this distorted protrusion  $C$  gives the same result as if the system had chosen to delete  $A$  in the first place. Note that when  $A$  is deleted,  $B$ 's symmetry is not distorted, since  $B$  no longer exists. Since the deletion of  $B$  would not effectively simplify the structure, this step is further discouraged by the minimum principle: deleting part candidate  $A$  is a better move since the description implied by this deletion is shorter. We conclude with our first heuristic: *if the deletion of a part candidate distorts the symmetry axis of any of its neighbors that do not disappear, this deletion is considered invalid*<sup>4</sup>.



**Figure 3.14:** Removing  $B$  from the shape in (a) distorts the symmetry of  $A$  and results in a new part  $C$ , as shown in (b), which is subsequently removed; Removing part  $A$  causes  $B$  to disappear (c).

To measure whether the deletion of a part candidate makes a neighboring candidate disappear, we introduce the *dependency*, or  $\delta$ . Roughly speaking,  $\delta$  measures how much a potential part can be explained as the transition of a neighboring potential part from the other side of

<sup>4</sup> This corresponds to the second heuristic in [Mi et al., 2009] which is called: Don't distort parts.



**Figure 3.15:** Elements of the definitions of  $\delta_{B|A}$  and  $\delta_{A|B}$

the shape boundary. For example, part  $B$  in Figure 3.14 only exists because of the transitional region of  $A$ . Given a foreground part  $A$  and a background part  $B$ , the dependency of  $B$  on  $A$ , denoted as  $\delta_{B|A}$ , is defined by:

$$\delta_{B|A} = 1 - \frac{\int_{\ell_B \cup c_A - \ell_A} \min(0, \kappa(\gamma)) d\gamma}{\int_{\ell_B} \min(0, \kappa(\gamma)) d\gamma}$$

where  $\ell_A$ ,  $\ell_B$ ,  $c_A$  and  $c_B$  are the boundary and completion curves, respectively, of candidates  $A$  and  $B$ . Figure 3.15 further illustrates these curve elements.  $\int_{\ell} \min(0, \kappa(\gamma)) d\gamma$  measures the total negative curvature along a curve  $\ell$ . Here we care about concave curve segments only since  $B$  would not be a background candidate without body curves of negative curvatures. Likewise, if  $B$  is in the foreground, the max operator will be used instead. A larger value of  $\delta_{B|A}$  means that  $B$  depends more on  $A$ . For instance, in Figure 3.14,  $\delta_{B|A} = 1$ ; after  $A$  is removed, there will be no trace of  $B$  left. In practice, we are not so strict and say that  $B$  “almost” disappears after  $A$  is deleted, if  $\delta_{B|A} > \delta_0$ . Useful values of  $\delta_0$  vary from 0.6 to 0.75. In our system, we use  $2/3$  as the threshold  $\delta_0$ .

### 3.3.2 Prevent distortion of any salient axial symmetry

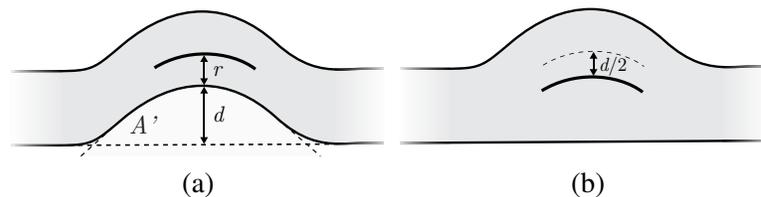
Avoiding distortion of the symmetry should be further extended to symmetries that have not yet been discovered as a basic part candidate. We have already seen such cases illustrated by Figure 3.13(a), where the axis of the bending structure is moved significantly if either the foreground part  $A$  or the background part  $A'$  is deleted. In particular, we say that deleting  $A$  *intrudes* on the bending structure since they are both on the foreground and the deletion narrows

the structure down. If the part is on the other side of the shape boundary, such as  $A'$ , we say that the deletion *disrupts* the structure.

Note that although only the SLS of basic part candidates are computed when we trace them from curvature extrema, and deleting a candidate may distort other axial symmetries, preventing distortion of this extended set of symmetries does not mean that we should fall back to compute the whole set of SLS. Appendix 3.A describes our approximate solution to find symmetries that may be distorted. Here, we assume that we have this SLS already. We need to check if it is distorted when a part is deleted.

### 3.3.3 Measuring symmetry distortion

To measure how much a local symmetry is distorted, we introduce another term, the *symmetry distortion*, or  $\lambda$ . This is defined as the ratio between the amount of the boundary movement normalized by the sum of this movement and the chordal radius. Specifically,  $\lambda = \frac{d}{r+d}$ , where  $d$  is the maximum movement of the shape boundary due to the deletion of a part and  $r$  is the chordal radius, illustrated in Figure 3.16(a). A deletion with  $\lambda > \lambda_0$  is considered significant symmetry breaking and is thus restricted. Useful values of  $\lambda_0$  range from 0.1 to 0.3. In our system,  $\lambda_0 = \frac{1}{5}$ <sup>5</sup>. Section 3.6.2 will further justify our choice of the threshold values. Note that  $\lambda$  measures the distortion of both the symmetry of a part and any other salient axial structure computed according to Appendix 3.A.

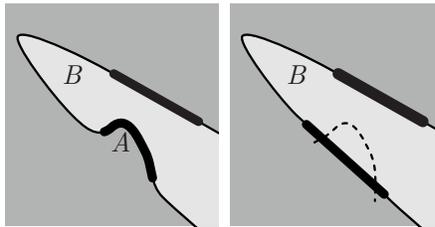


**Figure 3.16:** (a) Measuring the symmetry distortion of the bending structure due to the removal of part  $A'$ ; (b) the movement of the SLS axis

<sup>5</sup> The amount of movement of the *axis* is  $\frac{d}{2}$ , and the corresponding normalized distortion of the axis should be  $\frac{d}{2r+d}$ . The threshold of  $\frac{1}{9}$  under this measurement will predict symmetry distortion equivalently.

### 3.3.4 Deletions that improve symmetry

The argument against the deletion of  $B$  in Figure 3.14(b) includes the complaint that this deletion does not lead to a more organized structure, and this is justified by the idea of recoverability. This implies that if a deletion does lead to more symmetric geometry, it should be allowed, even if it distorts the less symmetric axial structure. We have already seen one such example: deletion of part  $B$  in Figure 3.13(b) is perfectly valid since it leads to a more salient axial structure, even though the symmetry distortion measurement  $\lambda$  is well above the threshold  $\lambda_0$ . This also explains in part why deleting part  $A$  from the top tail fin in Figure 3.11(a) (and filling the indentation) seems to be a compelling next step, aside from the justification given by Section 3.2.2 in terms of the contour grouping strength.

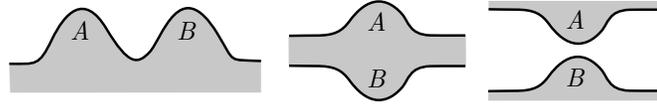


**Figure 3.17:** The curves of the symmetries before and after the deletion of  $A$

To measure whether a deletion leads to a more symmetric structure, we locate the approximate medial symmetry that this deletion affects (again, see Appendix 3.A for details), and this relates two boundary curves as shown in bold Figure 3.17. Here, the curve-matching algorithm by Sebastian et al. [2003] is appropriate for our purpose. We compute the energy of alignment between these two curves using Sebastian’s algorithm. If this energy decreases after the deletion of  $A$ , then this step is considered as improving the symmetry it affects, which makes the deletion valid.

### 3.3.5 Preventing breaking of inter-part symmetries

Deletion of a part could also break the symmetries along the sleeve, when a pair of part candidates across the sleeve form a reflective symmetry which aligns with the symmetries of the sleeve itself. Such inter-part symmetries also exist between neighboring parts, as described earlier in Section 3.1 (although we only considered foreground parts then). Our system actively



**Figure 3.18:** Symmetries between parts advocate group deletion. The left figure shows neighboring symmetric candidates; the middle and the right figures are instances of part candidates symmetric across a symmetry axis.

searches for these symmetries and avoids deleting its symmetric counterparts.

Figure 3.18 illustrates all possible configurations of part candidates that we consider removing. In particular, we consider two neighboring part candidates locally symmetric and delete them in the same group if they are on the same side of shape boundary, and both their scales  $\tilde{L}$  (see Figure 3.5 in Section 3.1.2) and protrusions differ by a factor of less than 2. The protrusion of a part candidate is measured as  $\int_c ds'_\perp$ , the perpendicular component of the SLS axial length<sup>6</sup>. Here we use a grouping scheme similar to that of the original in Section 3.2.2, except that now we also consider the amount of protrusion. In addition, we have tried measuring the symmetry between two candidates by how well the boundary curves of the two candidates align with each other [Sebastian et al., 2003], but this does not produce a significant improvement over our existing simple measurement.

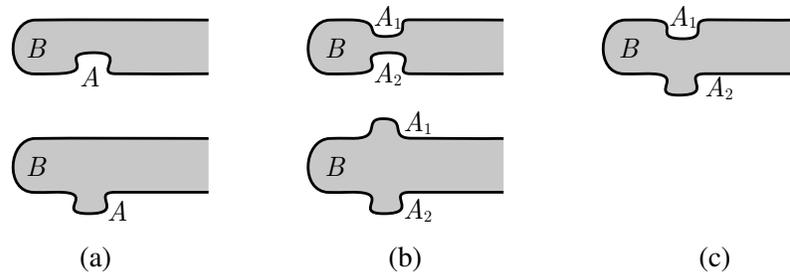
For remote candidates across a sleeve, we further require that the width of the sleeve is smaller than the sum of the candidates' scales. More remote pairs of candidates are not considered symmetric across the sleeve. This is justified by the psychophysical study by Siddiqi et al. [2001a] on the shape triangle. A pair of bumps on the boundary only form a part when their distance is small compared to their scale.

### 3.4 Heuristics to avoid premature axis truncation

Let us continue studying the example of the top tail fin in Figures 3.11 and 3.17. Part *B* is not a good candidate for deletion not because it distorts the symmetry of part *A* (since the deletion causes *A* to disappear as well), but because it truncates the whole top tail fin. However, the

---

<sup>6</sup> Note that our definition of protrusion differs from that in Hoffman and Singh [1997] where it is defined as the ratio of the perimeter to the base.



**Figure 3.19:** Cases of part removal that extends, from left to the right, the symmetry of a salient axial structure. In each case, the candidate that is constrained to be deleted to avoid premature axis truncation is labeled as  $B$ . Its neighboring candidates, deletion of which extends the symmetry of  $B$ , are labeled as  $A$ ,  $A_1$  or  $A_2$ .

most natural description of the tail would ideally include a complete tail fin, which exhibits a salient axial structure. We would like to preserve this structure by invalidating the deletion of  $B$ . Figure 3.1 on page 62 is another example. Unless  $A$  is removed in advance, neither part  $B$  nor  $C$  should be deleted, since each of them represents a portion of the base with a salient axial structure.

Figure 3.19 summarizes the general configurations of salient axial symmetries of the foreground, with both foreground and background part candidates sitting on them. In the case of (a), where there is only one part on either side, the part candidate  $B$  is constrained from being deleted only when the removal of a neighboring candidate  $A$  improves the global symmetry. Recall that in this case,  $A$  is not considered to distort any symmetry, and is allowed to be removed (see Section 3.3.4). In Figure 3.19(b), there are part candidates, labeled as  $A_1$  and  $A_2$ , on both sides of the candidate  $B$ . And they are on the same side of the shape boundary—both on the foreground or on the background. The system simply holds off deleting candidate  $B$  and waits for candidates  $A_1$  and  $A_2$  to be removed first.

In both cases,  $A$ , or  $A_1$  and  $A_2$ , form negative curvatures on the boundary segments. It is these parts that cause the transitional region of candidate  $B$ , which belongs to a longer and more salient axial structure and should not be truncated prematurely. According to causality theory, it makes more sense to remove the cause of the event in the shape history, which encourages deleting  $A$  or  $A_1$  and  $A_2$  first. Analyzing the shape this way also preserves the explanation that considers  $B$  as belonging to the main structure. Thus, instead of accepting  $B$  as a real part, candidate  $A$ ,  $A_1$  and  $A_2$  should be removed and put in the shape description first. We conclude

with the second heuristic: *if a part candidate belongs to a holistic, salient axial structure, and deleting the candidate truncates this structure in the middle, this deletion is invalid*<sup>7</sup>.

The problem here is that this “salient axial structure” is typically not available in those traced SLS axes, partly due to early termination of the SLS tracings<sup>8</sup>. We resort to using the CAT computed using the CDT, the discrete chordal axis transformation, to find these salient structures. Appendix 3.B describes this approach in detail. How to identify the part candidates (such as  $A$ ,  $A_1$  and  $A_2$  in Figure 3.19) that potentially extend the symmetry of their neighboring part (such as  $B$ ) are also explained.

Note the difference between the case demonstrated in Figure 3.19(b) and the symmetric part candidates across a symmetric axis as illustrated in Figure 3.18. They are not more related than their seemingly similar appearance. In fact,  $A_1$  and  $A_2$  in Figure 3.19(b) do not have to be symmetric. Likewise, symmetric candidates  $A$  and  $B$  in Figure 3.18 may not have a neighboring candidate that should be preserved from premature truncation.

### 3.4.1 Discussion

Figure 3.19(c) reveals one aspect of the limitations to our approach: salient axial symmetry detection using the CAT (Appendix 3.B) needs improvement. This figure shows a case when our system cannot make intuitive judgements. This happens when there are part candidates on both sides, but one is a protrusion and the other an indentation. Two factors complicate the situation. First, the two part candidates may form a bending structure. Second, the generalization of the junction by its sleeveiness and the turning angle criterion could fail here, as it typically does on poorly sampled polygonal shapes. Recall that in Section 1.2.4 of Chapter 1, we argued that removing parts can be considered a procedure for recovering the set of parts that are essential for the recognition of the shape. Accordingly their symmetries are only approximate in describing the entire shape, since irrelevant parts are not described. During the process of the part removal, we want to avoid breaking the symmetries of these essential parts.

---

<sup>7</sup> This corresponds to the first heuristic in [Mi et al., 2009] which is called: Dont truncate axes prematurely.

<sup>8</sup> In fact, it turns out that SLS axes are not ideal for computing salient axial structures, as is further explained in Appendix 3.A.

If we call recovering the structure of the shape by removing parts a “bottom-up” procedure, we need an auxiliary “top-down” process to tell the “bottom-up” approach in advance what to avoid disrupting, which is a dilemma since what will be recovered later is not available. The computation of the generalized salient sleeve of the CDT is a more or less temporary solution, in the good hope that salient symmetries are captured by these sleeves. We believe that the body of research in approximate symmetry inference and computation directly from the original shape [Feldman and Singh, 2006; Lee and Liu, 2009; Mitra et al., 2006] is the right track towards a better solution.

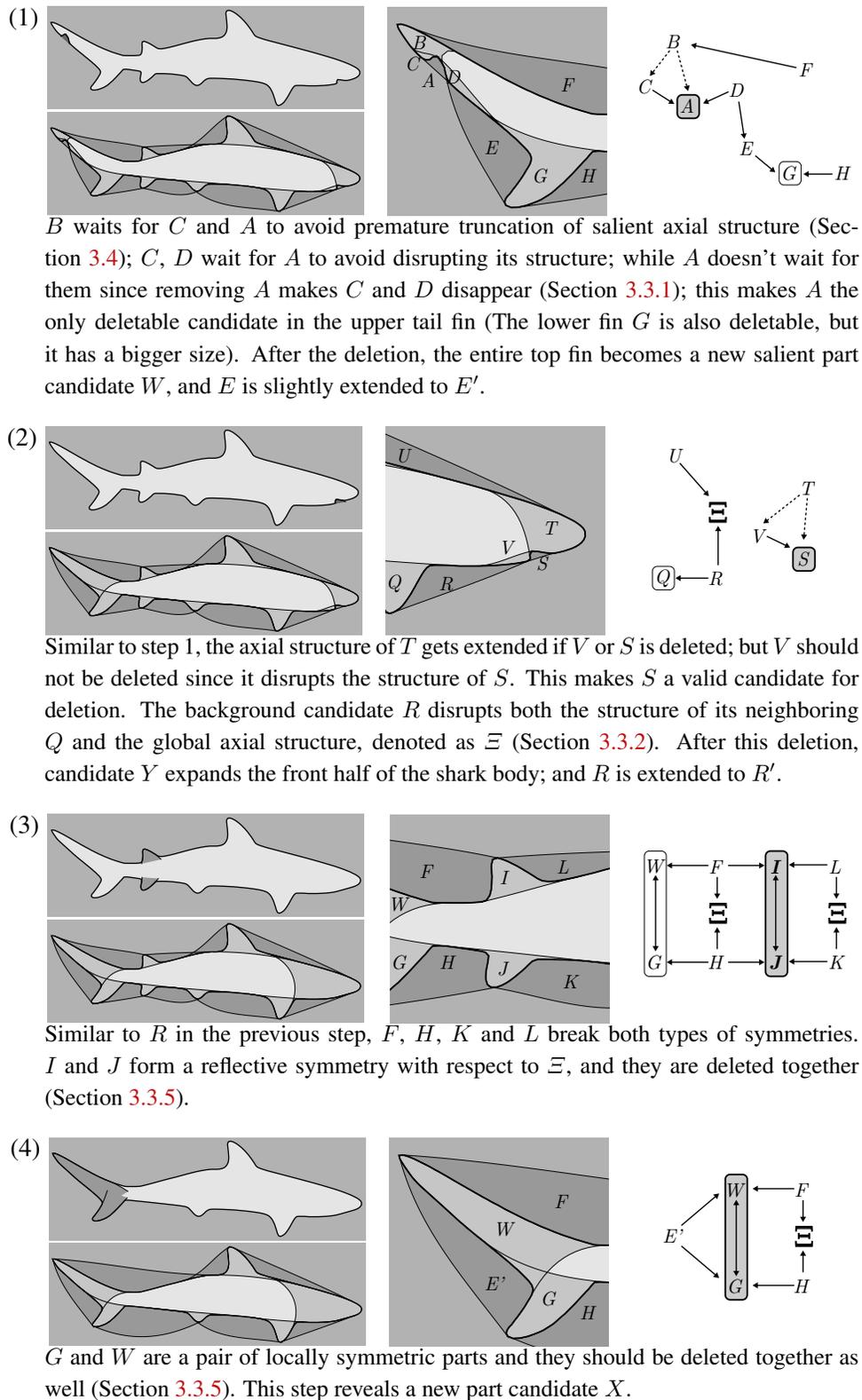
### 3.5 Shape analysis by sequential constrained part pruning

The heuristics described in the previous two sections impose constraints on how the basic part candidates should be deleted. We construct a direct graph [Cormen et al., 2001, page 1080 – 1084] to help the system sort out the set of parts from the pool of candidates, such that all the constraints are observed. The graph is constructed as follows:

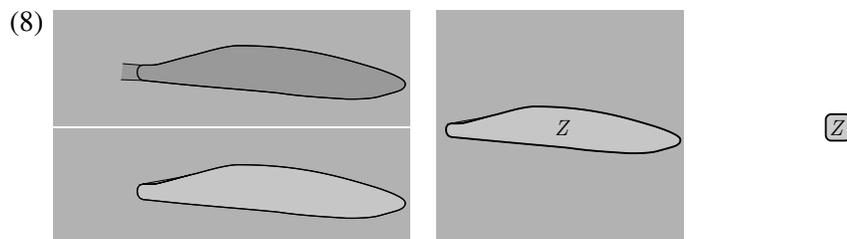
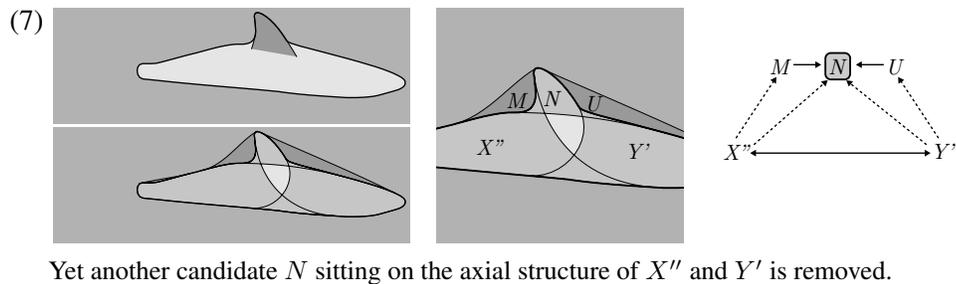
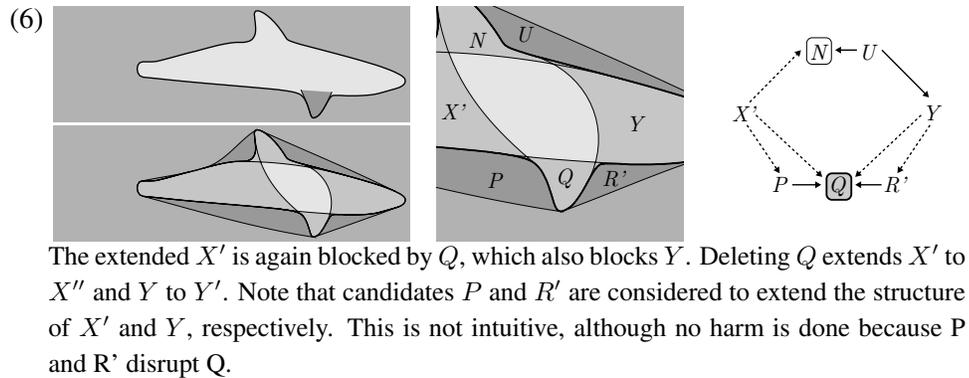
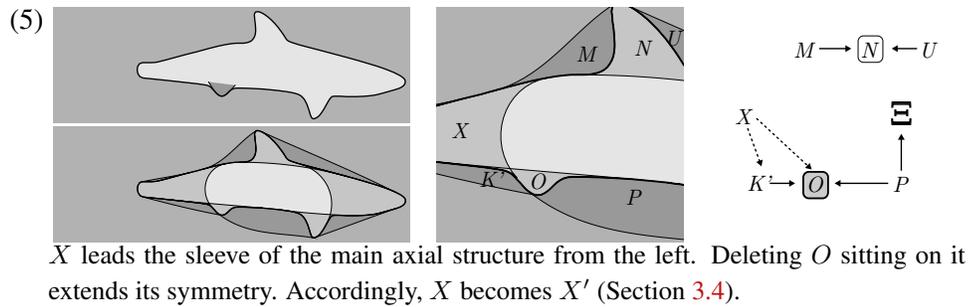
- Each part candidate is a node of the graph;
- If candidate  $P_1$  interrupts the structure of its neighbor  $P_2$ , and after  $P_2$  is deleted,  $P_2$  does not almost disappear, add an edge from  $P_1$  to  $P_2$  (Section 3.3.1);
- If candidate  $P$  intrudes or disrupts a salient axial structure  $\Xi$  that does not belong to any basic part candidates (Section 3.3.2), simply remove  $P$  from the graph, along with all its incident and outgoing edges<sup>9</sup>;
- If two candidates are considered symmetric either locally or across a sleeve they both sit on, we make sure they are deleted together by adding a pair of edges pointing from each candidate to the other (Section 3.3.5);
- If a candidate  $P_2$  sits on a sleeve led by one of its neighboring candidates  $P_1$ , and deleting  $P_2$  results in a more symmetric sleeve, add an edge from  $P_1$  to  $P_2$ . (Section 3.4).

---

<sup>9</sup> Note that this can create a situation where the removal of a neighboring part can distort  $P$ . However, given the choice of allowing the deletion of  $P$  which will distort a salient axial structure  $\Xi$ , or the distortion of  $P$  itself, this compromise solution avoids “deadlock” and opts for the latter—the potential distortion of  $P$ .



**Figure 3.20(a):** Step by step decomposition of the shark.



**Figure 3.20(b):** Step by step decomposition of the shark, continued. See text for explanation.

We compute the strongly connected components of the graph. These components divide the set of candidates into groups. Then we find groups with no outgoing edges since any of such groups is valid for deletion as long as all the candidates are deleted together based on the way the graph is constructed. There must be at least one such group since strongly connected components form an acyclic graph. We then remove the group that contains the smallest part.

Figures 3.20(a) and 3.20(b) illustrate the analysis of a shark shape by sequential constrained part pruning. Each step is demonstrated by a set of figures. In particular, the upper left figure of each step highlights the next part to be deleted in a dark color, and the choice is further explained by the remaining figures and the notes below. The lower left figure shows the entire set of part candidates; the figure in the middle is a close-up version that shows the region around the part to be deleted, each letter denoting a candidate; the graph is shown on the right, with the directed edges explained below the figures. The parts that distort  $\Xi$  and their outgoing edge are also included, although they are not counted when solving for deletable parts using the graph algorithm. In the graph for each step, the strongly connected components with zero outdegree are labeled with a rounded rectangle; in particular, the rectangle of the component to be deleted next is shown in gray. In our approach, we simply delete the strongly connected component that contains the part with the smallest scale, which is the scheme explained in Section 3.1.2. The difference here is that this time our choice also conforms to heuristics that model perceptual organization.

### 3.6 User guided shape analysis and performance assessment

Thus far, at each step in the shape parsing, the system automatically chooses the next set of parts that comply with the heuristics. No human interaction is necessary or possible. It is often desirable, however, to allow users to interactively express their “explanation” or “understanding” of the shape. We provide a scribble interface that lets users do so by drawing strokes of specified types on top of the shape using a tablet. Two types of strokes are available which correspond to two different explanations for the area of the shape that the strokes cover:

1. **Keep whole:** This stroke tells the parser there should be no part that contains only a portion of the stroke. The top row of Figure 3.21 shows such a stroke on the map in blue of the Fife area of Scotland, UK. The stroke covers the whole *Firth of Tay* as well as *St. Andrews Bay*. By default, as shown in the middle figure, our shape analysis system explains the eastern part of Fife as a single peninsula with an indentation<sup>10</sup>. The Firth of

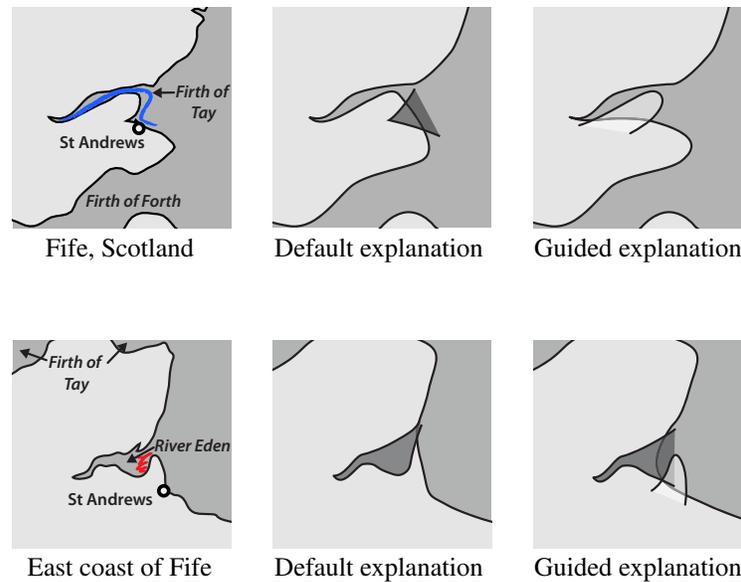
---

<sup>10</sup> Geographically, it is the *Estuary of the River Eden* and *St. Andrews Bay*.

Tay and the bay are considered two separated parts. However, this stroke prohibits this explanation. Instead, it tells the parser to analyze it in a such way that a single part covers both the firth and the bay. The only compatible analysis is to consider northern Fife as a part protruding into a very wide version of the Firth of Tay, as the top right section of Figure 3.21 shows.

2. **Not a part:** Briefly speaking, this stroke informs the system that there is no part present on this side of the shape boundary. More precisely, any locally part-like elements that contain any portion of the stroke are considered as resulting from the formation of a neighboring *real* part, which is typically on the other side of the shape boundary. If the stroke is on the foreground, then there should be no foreground part whose body is formed by the convex boundary curves around the stroke. Likewise, if it is on the background, the nearby concave shape boundary should not form the body of any part on the background. The bottom left section of Figure 3.21 is a closed-up version of the top left map, zoomed in to show details of east Fife around St. Andrews and the River Eden. By scribbling strokes (shown in red) to the left of the protrusion above St. Andrews, the system is informed that the concave curves to the left do not form any part on the background. Thus, it knows the protrusion on the right should be deleted so that it becomes part of the shape description. After this deletion, as the bottom right figure shows, the river is still recognized as a salient part, but this time, its form does not conflict with the stroke.

This “understanding” could be used in two different ways. We might consider it as guidance for shape parsing. The automatic shape parsing system can be modified so that the user guidance is respected. Alternatively, the user strokes could be used as a way to measure how well the system parses the shape, by counting how many part deletions are inconsistent with the strokes that the user has input. The system is more effective if it gives fewer explanations inconsistent with user strokes.



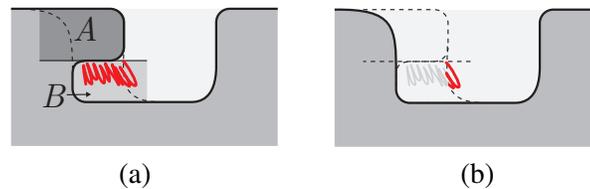
**Figure 3.21:** Top row: the “keep whole” stroke merges St. Andrews Bay and the Firth of Tay into a single background part, and the northeastern Fife is interpreted as a peninsula into this part; Bottom row: the “not a part” stroke makes certain that the peninsula above the town of St. Andrews is recognized as a part.

### 3.6.1 User guidance

To guide part pruning using strokes, we can simply modify the candidate graph constructed by Section 3.5: remove all the conflicting candidates, together with their incident and outgoing edges. This treatment is similar to how we treat candidates that disrupt the global axial structure  $\Xi$  (section 3.3.2).

Strokes of type “not a part” need slightly more care. First, after the actual enforcement of such a stroke, i.e., a neighboring candidate on the other side of the shape boundary is removed, and the portion of the stroke that overlaps the conflicting candidate should be removed too. Figure 3.22 shows a common case that necessitates user guidance: the automatic analysis opts to remove  $B$ , but the user can simply cross out the region of  $B$  using the red stroke. This way, deleting candidate  $A$  becomes the best next step. After  $A$  is removed, only the portion of this stroke outside the region of  $B$  is still active. Imagine what would happen if we did not deactivate strokes this way: the larger indentation shown in Figure 3.22(b) would never be removed. Another issue is that the user may have difficulty in deciding what the region of a candidate actually covers. Also observe that in Figure 3.22(a) the computed completion

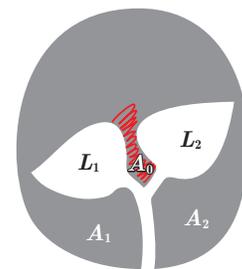
curve for  $B$  does not really enclose the whole stroke. To overcome this problem, and also to tolerate strokes that are not placed very carefully, we judge the *size* of the stroke and that of the conflicting candidate. We measure the effective area of a stroke as the area of its smallest bounding square  $A_s$ . Only candidates with area  $A_p < nA_s$  are tested to see whether they comply with the “not a part” constraints. In our implementation, we set  $n = 9$  and find it works effectively.



**Figure 3.22:** (a) A “Not a part” stroke disables interpreting  $B$  as a part; (b) After the deletion of part  $A$ , the portion of the stroke inside the region of part candidate  $B$  is disabled (in gray), as the acknowledgement of this stroke.

By definition, it does not make sense to have any stroke that crosses the shape boundary. The user will get immediate warning upon such a mistake. Strokes could also be inconsistent or contradictory, which could lead to deadlocks in the shape parsing if too many strokes are specified. However, the mechanism is not currently available to effectively detects such inconsistencies. The user needs to see the actual decomposition results before making adjustments accordingly.

We also allow the user to interfere with the shape parsing by specifying a weight  $w_l$  to adjust the the distance space of the foreground. In computing the part scale (to order the deletion) and symmetry breaking, the distance of the foreground is scaled by  $w_l$ . This is often a more effective way to globally bias towards deleting foreground or background part candidates. Figure 3.23 uses a “not a part” stroke to avoid wrongly classifying the gap  $A_0$  between two leaves  $L_1$  and  $L_2$  as a part.



**Figure 3.23:** The “not a part” stroke to avoid wrong figure-ground discrimination

For biological shapes like this, we can simply bias towards removing parts on the figure’s side. In this case, since the figure is on the background, we set  $w_l \gg 1$ . This way, the effective scales of foreground candidates, like that of  $A_0$ , are enlarged so that deleting them has lower priority.

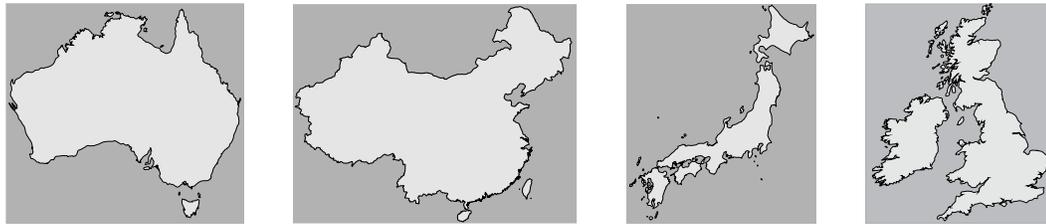
### 3.6.2 Assessing thresholds

Our shape analysis system involves two types of heuristics that drive the decision-making in choosing the right set of candidates. These heuristics are parameterized by a set of thresholds; the system behaves differently with varying values of these thresholds. We gather a set of user strokes that we consider to be "ground truth" for a particular shape. The strokes provide an opportunity to quantitatively measure the system's performance with a given set of thresholds. To assess our approach, we let the system finish its analysis on a given shape *without* enforcing the constraints of the strokes. Instead, we simply count the number of strokes that each step of deletion conflicts with. The scale factor  $w_l$  described in Section 3.6.1 still applies for more intuitive correspondence between strokes and parts.

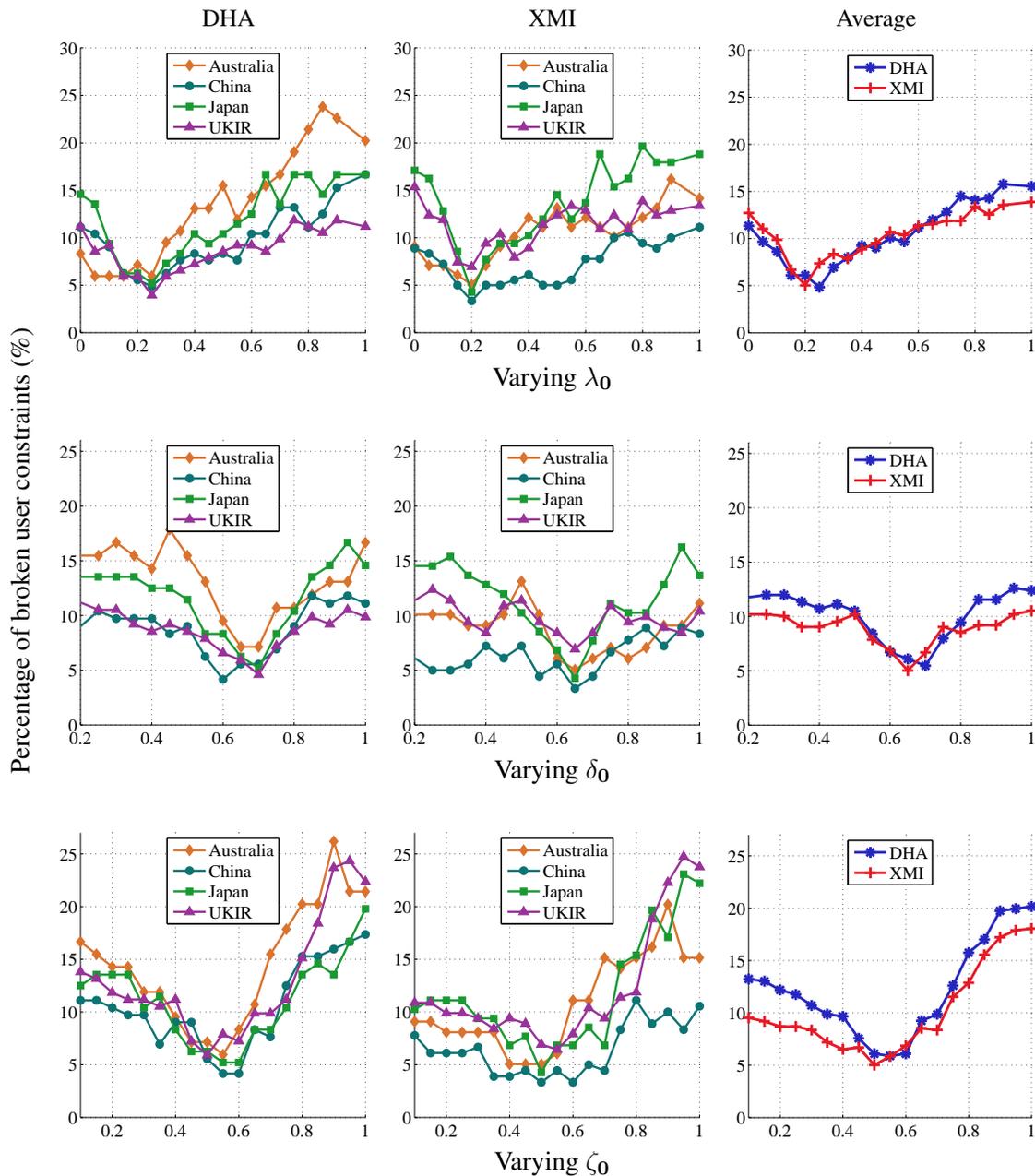
In particular, we are interested in the following thresholds:

- $\lambda_0$  The threshold to determine whether the symmetry of a part with a salient axial structure is disrupted. A larger value of  $\lambda_0$  tolerates more symmetry breaking.
- $\delta_0$  The threshold to test whether a candidate disappears after a neighboring part has been deleted. A value of 0 virtually disables the heuristics of symmetry breaking, and a value of 1 ignores the exception of allowing part deletion that causes neighboring candidates to disappear.
- $\zeta_0$  The threshold to determine the main sleeves of the CDT (see Appendix 3.B). A smaller value of  $\zeta_0$  allows more sleeves to be connected. This certainly encourages the extension of the main symmetry and makes it less likely that a salient axial structure would be disrupted in the middle, which helps to avoid breaking the "keep whole" constraints. However, this also makes the system too aggressive in extending the axial structures of part candidates at the cost of deleting some of the wrong candidates that sit on them.

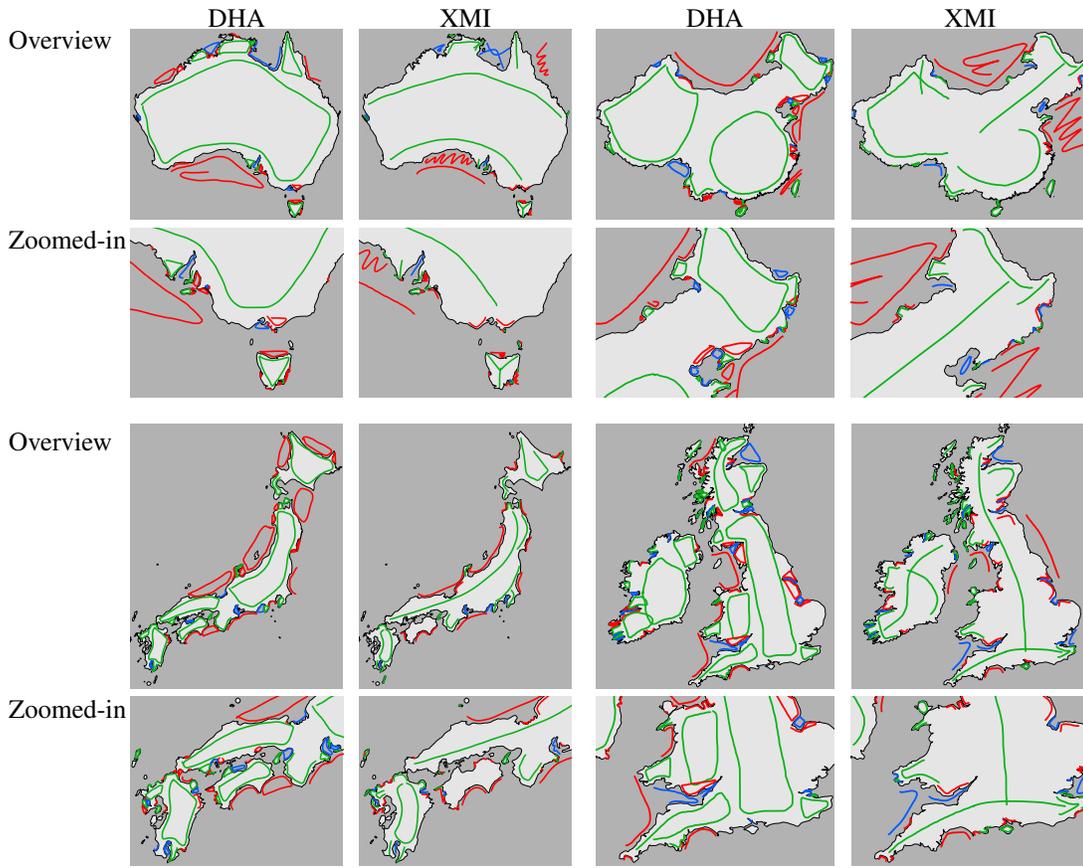
We confirmed the proper establishment of these thresholds by evaluating the performance of the parser. The assessment involved two subjects (XMI, the author; and DHA, a graduate student from Rutgers University), each of whom was asked to draw strokes of both types on four maps: Australia, Mainland China and Taiwan, Japan, and the United Kingdom and Ireland, as



Maps used from left to right: Australia, Mainland China and Taiwan, Japan, UK and Ireland



**Figure 3.24:** The performance of the shape parsing system in terms of the percentage of strokes with which the deletions are inconsistent. With other thresholds set at their default values, each row shows how the performance varies by changing values of  $\lambda_0$ ,  $\delta_0$  and  $\zeta_0$ , respectively. The first and second columns shows results from two users DHA and XMI on of the four maps. The last column gives the system's average performance for all maps with respect to each user's inputs.



**Figure 3.25:** Examples of strokes used for threshold assessment. “Keep whole” strokes are in green (foreground) and blue (background). “Not a part” strokes are in red.

shown in the top row of Figure 3.24. Some of the strokes are shown in Figure 3.25. In particular, each figure of the top row shows the strokes related to the most prominent set of parts. And the bottom rows are the zoomed-in versions with strokes related to parts of finer detail. For many shapes, functional semantics could be attached to specific parts. We exclusively chose geographic shapes in the hope of ruling out the effects of the users’ knowledge of part semantics not related to the shape itself.

As described in the previous sections, our system uses a set of threshold values fixed at  $\lambda_0 = \frac{1}{5}$ ,  $\delta_0 = \frac{2}{3}$ ,  $\zeta_0 = \frac{1}{2}$ . (These threshold values were chosen by hand so that the parser analyzes the Australia shape satisfactorily to us.) It is impractical to find the connections between these values and the number of inconsistencies analytically. In part, it is because whether to enforce a heuristic constraint on deletion is determined by a profound process involving a wide range of geometric factors. Moreover, the determination of a subsequent deletion is based on the

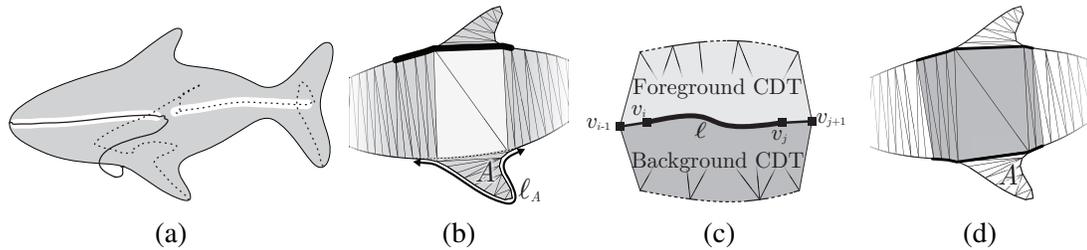
analysis of the resulting shape from previous deletions, which means the “butterfly effect” more or less exists: a small part might affect the analysis of a relatively wide region around its neighborhood. Nevertheless, we are still able to find these connections through experiments and draw conclusions in a statistical sense based on our observations. The remaining three rows of Figure 3.24 show the percentages of user strokes with which our automatic shape analysis is inconsistent. Each row has one varying and two fixed threshold values. As the figures imply, the set of default values is indeed locally optimal. This can further be confirmed by a non-linear optimizer using the Hooke and Jeeves method [Hooke and Jeeves, 1961]. The table below lists the numbers of strokes that the analysis of the four shapes above is inconsistent with. These numbers are grouped into four rows by different combinations of the heuristics (with the thresholds set to their default values). In particular, H1 in the table refers to the heuristics to prevent symmetry breaking; and H2 refers to the heuristics to avoid premature truncation. The total number of strokes that each subject input for each shape is listed on the top row. The bottom two rows are respectively the numbers of deletions and the numbers of resulting parts that the automatic shape parsing produces. As the table shows, the shape analysis governed by both types of heuristics indeed respects the gathered data the best.

	<b>Australia</b>		<b>China</b>		<b>Japan</b>		<b>UKIR</b>	
	DHA	XMI	DHA	XMI	DHA	XMI	DHA	XMI
<b>Total strokes</b>	84	99	144	180	96	117	152	202
<b>H1 and H2</b>	6	5	8	6	6	5	9	14
<b>H1 only</b>	18	15	25	19	19	26	34	48
<b>H2 only</b>	17	14	24	20	16	22	27	27
<b>None</b>	28	21	45	44	29	33	47	58
<b>Total Deletions</b>	505		855		733		887	
<b>Parts</b>	544		908		768		950	

### 3.A Locating global chordal symmetries

When we explained the heuristics in this chapter, we heavily referred to potential symmetries that might be distorted by a part deletion. These symmetries reveal the global axial structure that a deletion should maintain. However, sometimes these symmetries are not available since each SLS tracing is terminated early when the system speculates that continuing the SLS tracing would result in no more potential part candidates. Even worse, a salient axial structure may not necessarily be represented by a single SLS axis, but by several of them. Even for the simple

shapes shown in Figure 3.26(a), the major axial symmetry of the fish body is divided into two continuous pieces, each in a separate SLS axis. For more complicated shapes, it would be very difficult to locate those segments of SLS axes that collectively form the major axial structures of the shape.



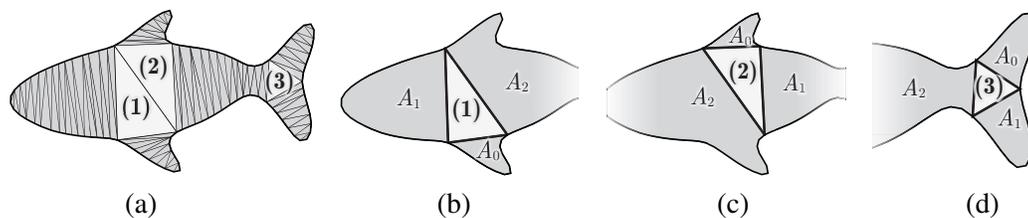
**Figure 3.26:** (a) The primary symmetry, which is a salient structure of the fish shape, is separated into multiple SLS axes; (b) the “other side” of the symmetry with respect to a boundary curve  $\ell_A$  of a part  $A$  is approximated as a path along the constrained Delaunay triangulation; (c) approximate global symmetries across both foreground and background CAT; (d) the foreground symmetries needed to be accounted for when deleting the bottom fin part  $A$ .

The good news is that salient axial symmetries are typically present in the medial axis representation, although each may span multiple medial edges. This inspires us to consider the subset of SLS: the medial symmetry, which is just the CAT representation. In practice, we use the constrained Delaunay triangulation (CDT), the discrete version of CAT, to help locate potential global symmetries. Figure 3.26 explains how CDT is used. For the boundary segment  $\ell_A$  of any given part  $A$ , spanning from  $v_i$  to  $v_j$ , we find its one-ring neighborhood in the CAT. By definition, this first-ring must include the previous and following vertices  $v_{i-1}$  and  $v_{j+1}$ . As Figure 3.26(c) shows, the segment from  $v_{i-1}$  to  $v_{j+1}$  divides the first-ring neighborhood into the background and foreground regions of triangles. The boundary of the first-ring is also divided into two sequences of vertices. Neighboring vertices of each sequence are also connected by a constrained Delaunay edge. More importantly, observe that one sequence of vertices is adjacent to  $\ell_A$  by internal Delaunay edges on the foreground and the other by edges on the background. Those Delaunay edges are used in our system as the set of the local symmetries that need to be accounted for when the part  $A$  is deleted. Figure 3.26(d) shows the symmetries across the foreground in gray. Note that on the foreground side, not all vertices of  $\ell_A$  are involved in these symmetries.

### 3.B Discovering salient axial structures with CDT

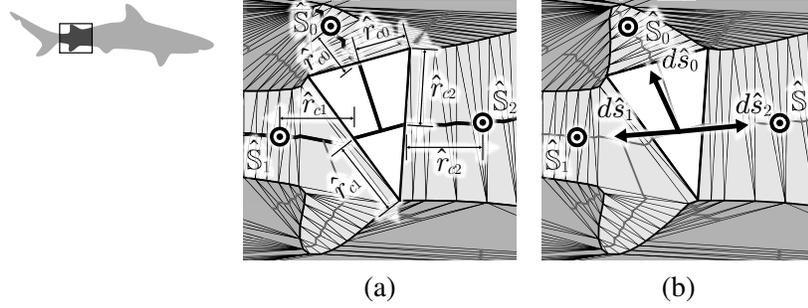
This section describes a method for finding long salient axial structures in the CAT. In particular, we join up separate sleeves of the CAT if the junction that joins them together is only present due to a small attached part, and also if the two sleeves line up well.

Recall that there are three types of triangles in a given CDT: terminal, sleeve and junction triangles which respectively have one, two and three internal (unconstrained) Delaunay edges. Here, we extend this classification by noting that a junction can be considered as a generalized sleeve triangle. Observe that a junction divides a shape into three subregions; each of its edges is adjacent to a sleeve of the CDT. A sleeve triangle can be considered as a junction with one subregion empty. Thus, from the opposite perspective, by measuring how “empty” the most empty side of a junction is, we are able to evaluate the “sleeviness” of that junction. Denoting the areas of the three regions divided by a junction as  $A_0$ ,  $A_1$  and  $A_2$  such that  $A_0 \leq A_1 \leq A_2$ . Figure 3.27(a) shows a fish shape with three junctions identified as (1), (2) and (3), among which the subregions  $A_0$ ,  $A_1$  and  $A_2$  corresponding to each junction are shown in (b), (c) and (d) respectively. We define the *sleeviness* as  $\zeta = 1 - \frac{A_0}{A_1}$ . Obviously, for a sleeve triangle,  $\zeta = 1$ . This value is not defined for a terminal triangle, where  $A_1 = 0$ . In our system, we introduce a threshold  $\zeta_0$  such that all triangles with  $\zeta > \zeta_0$  are *sleevy*. In a sleevy triangle, we call the two edges that are adjacent to  $A_1$  and  $A_2$  *sleevy edges*. In the discrete CAT, each Delaunay edge forms a local symmetry; thus, sleevy edges connected by sleevy triangles form a sequence of more salient and consistent local symmetries.



**Figure 3.27:** (a) The CDT of the foreground of the fish shape; (b, c) more “sleevy” junctions; (d) A less “sleevy” junction.

By generalizing the sleeve triangle to junctions, it is possible for our system to realize the continuation of the main body from the head on the left to the tail on the right, which stops at



**Figure 3.28:** The axial direction of each sleeve  $i$ , ( $i = 0, 1, 2$ ) at a sleevey junction is estimated as the vector pointing from the center of the junction axis to the axial point that is  $\hat{r}_{c_i}$  in distance from the Delaunay edge adjacent to region  $i$ . (a) Shows how distances are measured; (b) shows the estimated axial directions at the junction triangle.

the left of junction 3.

We do not permit every sleevey junction triangle to connect the sleeve of its region 1 with that of its region 2. Instead, we take into account how smoothly one sleeve aligns with the other. Here, we define the inter-sleeve turning angle  $\Psi_{\hat{S}_{i,j}}$  ( $i, j = 0, 1, 2$  and  $i \neq j$ ), as the turning angle from the direction  $d\hat{s}_i$  to  $d\hat{s}_j$ , where  $\hat{S}_i$  and  $\hat{S}_j$  are the discrete chordal axes of the  $i^{th}$  and  $j^{th}$  regions, respectively;  $d\hat{s}_i$  and  $d\hat{s}_j$  are their axial directions at the junction. To handle noise [Prasad, 2007], the direction of the chordal axis is averaged along a distance of  $\min(\hat{r}_c, \|\hat{S}\|)$ , where  $\hat{r}_c$  is the chordal radius of the sleeve at the junction and  $\|\hat{S}\|$  is the length of the sleeve. Figure 3.28 explains how each chordal axis direction  $d\hat{s}_i$  at the junction is estimated. When evaluating  $\Psi_{\hat{S}_{i,j}}$ , we care about the amount of turning angle only, not the turning direction. Thus,

$$\Psi_{\hat{S}_{i,j}} = \arccos \left( -\frac{d\hat{s}_i \cdot d\hat{s}_j}{\|d\hat{s}_i\| \|d\hat{s}_j\|} \right)$$

Sleeves  $\hat{S}_1$  and  $\hat{S}_2$  are connected by the sleevey junction only if they line up significantly better than  $\hat{S}_0$  and  $\hat{S}_2$ . Formally, we require that

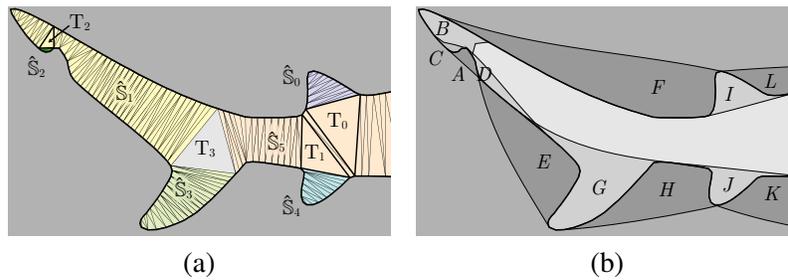
$$\Psi_{\hat{S}_{0,2}} - \Psi_{\hat{S}_{1,2}} > \Psi_0 \frac{1 - \zeta}{1 - \zeta_0}, \quad (3.1)$$

where  $\Psi_0 > 0$  is a constant that is applied as a weight to the smoothness factor: a larger value of  $\Psi_0$  means a higher requirement in how well  $\hat{S}_1$  should line up with  $\hat{S}_2$ . In our system, it is fixed to be  $\frac{\pi}{3}$ . By multiplying  $\Psi_0$  with  $\frac{1 - \zeta}{1 - \zeta_0}$ , we purposely alleviate the condition for very sleevey

junctions. For example, if  $\zeta = 1$ , then (3.1) becomes equivalent to  $\Psi_{\widehat{S}_{0,2}} - \Psi_{\widehat{S}_{1,2}} > 0$ . which means as long as  $\widehat{S}_1$  connects to  $\widehat{S}_2$  more smoothly than  $\widehat{S}_0$  does, the junction is considered to be a valid sleevy triangle that can also connect neighboring sleevy edges to form a more salient sleeve structure. As  $\frac{1-\zeta}{1-\zeta_0}$  becomes closer to 1, the value of the right hand side of (3.1) increases. Accordingly, the connection from  $\widehat{S}_1$  to  $\widehat{S}_2$  needs to be significantly smoother than that from  $\widehat{S}_0$  to  $\widehat{S}_2$ .

Junctions that meet the conditions described above join isolated sleeves into longer ones. These enhanced sleeves are the salient axial symmetries that we would like to preserve. Since sleevy triangles include some of the junctions, the axial symmetries are the generalized version of the discrete chordal axis transformation described by Prasad [1997]. For a junction, a value of  $\zeta$  closer to 1 implies a better proximity to a sleeve triangle in the Delaunay triangulation and more dominant continuation from  $A_1$  to  $A_2$ . Note that the continuation of the sleeve is uniquely defined at each triangle as long as  $\zeta_0 > 0$ . Section 3.6.2 examines how different choices of  $\zeta_0$  affect the results of shape analysis.

Similar to the sleeve originally defined by Prasad [1997], the enhanced sleeve also consists of a series of adjacent Delaunay edges. Figure 3.29(a) offers examples of such generalized



**Figure 3.29:** (a) Generalized sleeves of the shark shape around the tail; (b) basic part candidates for this region

sleeves, with the associated parts shown in (b). Since there are no salient axial structures on the background, we show the constraint Delaunay triangulation of the foreground only. Of all the foreground junction triangles, the only non-sleevy triangle is  $T_3$ .  $T_0$ ,  $T_1$  and  $T_2$  are examples of sleevy triangles. They connect intermittent sleeves into longer ones, revealing salient axial structures in the shape. Among these sleeves,  $\widehat{S}_1$  is led by the part candidate  $B$ .  $C$ ,  $G$ ,  $I$  and  $J$  contain the entire sleeves of  $\widehat{S}_2$ ,  $\widehat{S}_3$ ,  $\widehat{S}_0$  and  $\widehat{S}_4$ , respectively, but are not considered to lead

them. Instead,  $C$ , together with  $A$  and  $D$ , sits on  $\widehat{\mathbb{S}}_1$ . Candidates  $I$  and  $J$  sit on the sleeve  $\widehat{\mathbb{S}}_5$ , which is part of the main shark body.

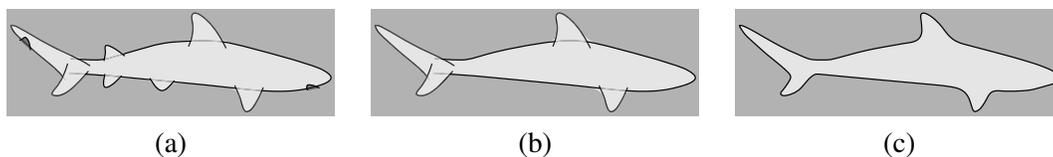
In describing these examples, we say that a part candidate contains a Delaunay edge if its boundary includes both end vertices. A candidate *leads* a sleeve if it contains either of the two end edges but not both. Also, we say that a candidate *sits* on a sleeve if its two boundary vertices are on the same side of a sleeve. A part candidate could be sitting on the sleeve from either the same or the other side of the shape boundary. For example,  $A$  sits on  $\widehat{\mathbb{S}}_1$  from the other side of the boundary since  $A$  is a background candidate, while  $I$  and  $J$  sit on  $\widehat{\mathbb{S}}_5$  from the same side. Either case can lead to a boundary section along the symmetry that the sleeve represents, which stops the tracing of the SLS from continuing. In the case of Figure 3.29,  $A$ 's body forms the transitional region which causes the tracing of  $B$ 's SLS to stop. The SLS tracing of the main body will also be stopped at around  $I$  and  $J$ , since these two protrusions forms negative curvatures on the boundary segments. In other words, it is the parts sitting on the sleeve that cause the transitional region of the leading candidate. Thus, instead of accepting  $B$  as a real part, in order to correctly analyze the upper tail fin, the first step is to remove the candidate  $A$  and put it in the shape description.

## Chapter 4

### Abstracted Depiction in terms of Parts

A complete sequence of deletions driven by heuristic constraints and user guidance produces a description of the shape in terms of its deleted parts. This set of parts serves as a structural representation that describes the distinct components of the shape. This makes it possible for each stroke in the depiction to articulate a certain shape component. This articulation is particularly important for an abstracted depiction where the economical placement of strokes is required. Fused-part artifacts, which are often seen in children’s drawings, can be effectively avoided. Recall that in Section 1.1.2, we pointed out that these fused parts are possibly caused by the lack of a clear compositional description of the shape.

The key to maintaining component articulation using only a few strokes is that not all components are needed for an expressive shape depiction. Only a small subset of them is essential for recognition [Hoffman and Richards, 1984]. For a successful visual presentation, the clear distinction between the depicted components is much more important than a complete view of the entire set of shape components [Biederman, 1987]. Thus, a natural approach to an effective system would be to retain the subset of parts that serves as the invariant for the recognition of the object and removing the rest. Figure 4.1 follows this procedure and presents an abstracted shark shape in (c) by removing irrelevant parts from (a), including the indentation on the upper



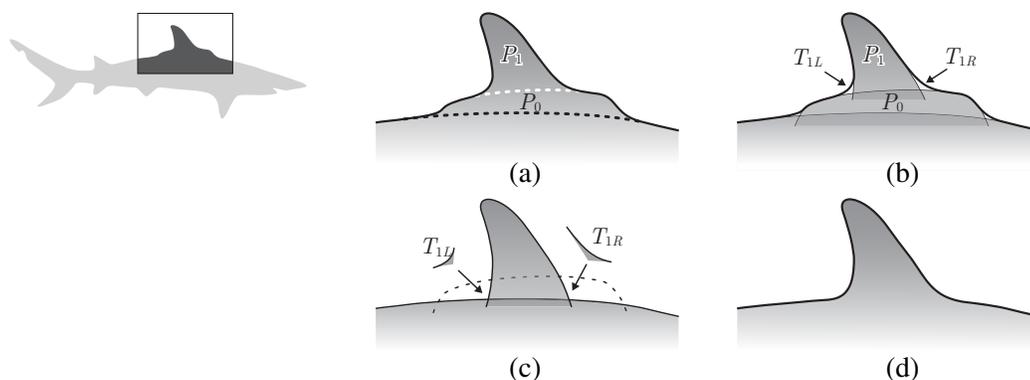
**Figure 4.1:** (a) The description of the shark shape in terms of parts, which follows naturally from the decomposition sequence illustrated by Figures 3.20(a) and 3.20(b) on pages 86–87. (b) The set of parts that is sufficient to express the shape of the shark; other parts are considered irrelevant. (c) The resulting shape with irrelevant parts removed.

tail fin, the mouth and three little fins on the body.

Without knowing the functions of the parts, the best we can do is to infer from each part's geometric properties its importance to the recognition process. Before we describe our approach to this problem, we first address a potential issue related to the topological constraints that constrain when some parts can be eliminated when making the abstraction.

#### 4.1 Part attachment and shape reconstruction

The success of abstraction by simply reversing the deletion order, such as shown in Figure 4.1(c), relies on luck. Recall that deleting a part is the process of replacing the boundary with the completion curve. If any portion of the boundary of a part to be deleted (say,  $P_0$ ) is the completion curve of another part ( $P_1$ ) that needs to be preserved, that portion of  $P_0$ 's boundary does not exist unless  $P_1$  is deleted first. This topological dependency between parts is caused by the sequential deletion in the shape analysis stage, which in turn constrains how parts can be eliminated for the abstraction. It is also possible to obtain the abstracted shape in Figure 4.1(c) by reattaching parts to be retained, starting from the main body. If the reattachment of a part is done by replacing the completion curve with its boundary, however, the topological dependency is still an issue.

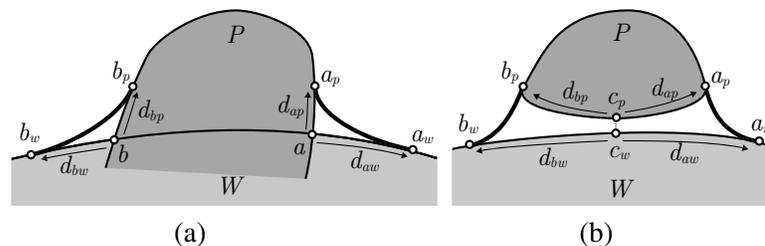


**Figure 4.2:** (a) The top fin,  $P_1$ , grows from the bumpy back  $P_0$ . By means of part attachment or removal based on curve replacement, it is not possible allow the shape to retain  $P_1$  without also retaining  $P_0$ . (b) Explicit separation of the extended part body from the transition is an alternative scheme for part attachment. (c) This scheme does not require the existence of the completion curve of the part being attached, thus,  $P_1$  can be attached to the shark without  $P_0$ . (d) The resulting abstraction.

Figure 4.2 shows a shark shape similar to that of Figure 4.1, but this shark has a slight bump between the back and the fin on the top. This bump is recognized as a separate part, as shown in Figure 4.2(a). The bump is not crucial for the shape to be recognized as a shark—the shape in Figure 4.1 does not have one—but the fin that grows from it is essential for a successful abstraction. This is exactly the situation described above. It is not straightforward to remove the bump simply by replacing its boundary with its completion curve without also removing the fin. Likewise, it is not possible to attach the fin on the top without attaching the bump first, since the completion curve of the fin comes from the bump.

#### 4.1.1 Parametrization of part transition

Our system computes the abstracted shape by the sequence of part attachments in reverse order of the shape deletions previously applied for shape analysis, except that non-essential parts are skipped. Inspired by the paradigm of “modeling by example” [Funkhouser et al., 2004], when attaching a part, we simply disregard the completion curve to overcome the dependency problem. Instead, we locate the intersection points between the base that the part is attached to and the extended part body, and then glue them together smoothly. Since the topology involved in the 2D shape attachment is considerably simpler than the situation Funkhouser et al. [2004] needed to handle, we go a little further than simply smoothing the intersections out. We hope to ensure that the new transitions from the fin into the shark body resemble its original transitions to the slight bump, which are labeled as  $T_{1L}$  and  $T_{1R}$  in Figures 4.2 (b) and (c).



**Figure 4.3:** Attaching a part  $P$  to the whole shape  $W$ .

To ensure the seamless, smooth connection of parts, each of a part’s two transitional boundary curves is described by a set of parameters that fits the original attachment. When attaching

the part to a new base, the same set of parameters is used to *reconstruct* the transitions. The parameters of each transition indicate both *where* to build the transition, and *how* the transitional boundary should appear. As demonstrated by Figure 4.3(a), the first transitional boundary of a part  $P$  (in counterclockwise order along the boundary curve) is the original curve segment from  $a_w$  to  $a_p$ , where  $a_w$  is also the starting point of both the boundary curve and the completion curve of  $P$ , and  $a_p$  is the location where the part body starts extrapolation on the right side. This transition smoothes out the first intersection  $a$  between the part body and the completion curve. We use two offset lengths  $d_{aw}$  and  $d_{ap}$  to describe the location of  $a_w$  and  $a_p$ , where  $d_{aw}$  is the offset from  $a$  to  $a_w$  along the completion curve and  $d_{ap}$  measures the offset from  $a$  to  $a_p$  along the extrapolated part body. The location of the second transitional curve between  $b_p$  and  $b_w$ , which corresponds to the second intersection point  $b$ , can be described by the two offsets  $d_{bw}$  and  $d_{bp}$  similarly.

Not all part bodies can be extrapolated. As Section 2.6 has described, when extrapolated radii along the the extrapolated axis become negative, the part has a closed body only and this body may not intersect with the completion curve. However, the transitional curves can still be described following the scheme illustrated by Figure 4.3 (b). In this case,  $d_{ap} = d_{bp}$  is defined as the distance from half way point  $c_p$  between  $a_p$  and  $b_p$  along  $P$ 's body curve;  $d_{aw}$  and  $d_{bw}$  are the offset lengths from  $c_w$  to  $a_w$  and  $b_w$ , respectively, where  $c_w$  is the closest point on the completion curve to  $c_p$ .

To parameterize the form of the transitions, we use cubic Hermite splines to fit the transitional boundaries, such that

$$\begin{cases} (m_{aw}, m_{ap}) = \arg \min_{m_{aw}, m_{ap}} \sum_i \left\| \alpha(s_{ai}) - \mathbf{h}_\alpha(a_w, m_{aw}, a_p, m_{ap}, \frac{s_{ai} - a_w}{a_p - a_w}) \right\|^2 \\ (m_{bw}, m_{bp}) = \arg \min_{m_{bw}, m_{bp}} \sum_j \left\| \alpha(s_{bj}) - \mathbf{h}_\alpha(b_p, m_{ap}, b_w, m_{bw}, \frac{s_{bj} - b_p}{b_w - b_p}) \right\|^2 \end{cases} \quad (4.1)$$

where  $\{s_{ai}\}$ ,  $\{s_{bj}\}$  are vertices on the two transitional curves, respectively, and  $\mathbf{h}$  is the hermite

curve given by:

$$\mathbf{h}_{\alpha}(s_0, m_0, s_1, m_1, s) = (s^3, s^2, s, 1) \begin{bmatrix} 2 & -2 & 1 & 1 \\ -3 & 3 & -2 & -1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{pmatrix} \alpha(s_0) \\ \alpha(s_1) \\ \alpha(s_0) + m_0\alpha'(s_0) \\ \alpha(s_1) + m_1\alpha'(s_1) \end{pmatrix}$$

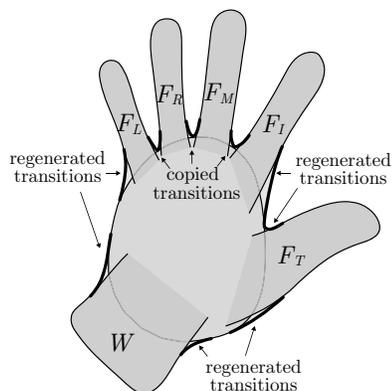
Note that the  $4 \times 4$  matrix in the middle is the standard Hermite geometry matrix. The values of  $m_0$  and  $m_1$  are simply set to zero in the rare event that they are negative from the fit of (4.1), which ensures the tangent directions don't flip along the curve.

The composition of the abstracted shape from the subset of parts proceeds in reverse order of the deletion in the shape analysis stage. For each connected shape, the set of parts deleted in the last step is not attached to anything. Accordingly, when we add them back onto the abstraction, we simply build a new shape consisting of these parts. This shape becomes the base for the subsequent attachment of parts which were originally deleted from it in the decomposition.

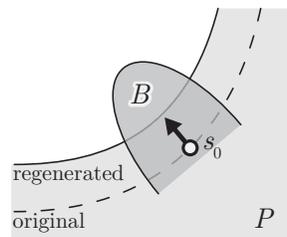
#### 4.1.2 Part attachment

Section 4.1.1 shows how the transition curves of a given part can be parameterized by an 8-tuple  $(d_{ap}, d_{bp}, d_{aw}, d_{bw}, m_{ap}, m_{bp}, m_{aw}, m_{bw})$ . To add transitional curves to glue the part and the base shape, intersections  $a$  and  $b$  are first located and the start and end points of each transitional curve are recovered from the first four offset parameters. For a part whose body is not extendable,  $d_{ap}$  and  $d_{bp}$  can respectively represent the parameters of points  $a_p$  and  $b_p$  on the curve of the part body. Also, note that since we are now attaching a new part  $P$  to an existing base shape  $W$ ,  $d_{aw}$  and  $d_{bw}$  are no longer the offsets along the completion curve, but rather along the boundary of  $W$ . The other four parameters determine the lengths of the tangent vectors at these end points. With these tangents and the positions of the end points as the boundary conditions for the Hermite splines, the geometry of each new transition is completely defined.

Note that when attaching onto the shape a group of neighboring parts, which were deleted together during the analysis step, they are attached together as well. For group attachments like



**Figure 4.4:** When attaching the fingers and wrist, it is not necessary to regenerate transitions between grouped fingers.



**Figure 4.5:** A displacement is given to  $B$  so that its prominence is maintained in the resulting abstraction

this, instead of rebuilding the transitions for all the parts, only the first transition (counterclockwise along the boundary, transition  $a$  in Figure 4.3) of the first part and the second transition of the last part (transition  $b$  in Figure 4.3, the parts are also ordered counterclockwise along the shape boundary) are regenerated. It is not necessary, and somewhat unwise, to regenerate the intermediate transitions since they are readily available from the original shape. Such an example is shown in Figure 4.4. The index, middle, ring and little fingers ( $F_I$ ,  $F_M$ ,  $F_R$  and  $F_L$  in the figure, respectively) are to be attached together. When we do this, we simply copy the transitional boundaries from the original model.

In rare cases a cubic curve does not quite fit the transitional curve of a part  $P$  well enough. Typically this is not a problem for the visual presentation of  $P$  itself since the human visual system is not sensitive to this level of geometric accuracy. However the inaccurate transition occasionally “buries” a smaller part that grows from it, as Figure 4.5 shows, where part  $B$  is almost buried by the reconstructed transition of part  $P$ . Consequently, in the resulting abstraction,  $B$  is considerably less prominent.

One solution to this problem is to remember the intersection point ( $s_0$ ) between the symmetry axis of part  $B$  and its completion curve. When attaching  $B$  back to the shape in the reconstruction of the abstraction, we find the closest point on the boundary of the base shape to  $s_0$ , and translate  $B$  accordingly so that  $s_0$  stays on the boundary. All the subparts of  $B$  are also translated accordingly to maintain their positions relative to  $B$ . In all the cases we have studied,

this simple solution suffices. However, the amount of such a translation is completely determined locally. For a more general solution, especially when displacements of different parts are involved, we would expect that special care needs to be taken to avoid boundary overlap or other global inconsistencies.

## 4.2 Selecting the right set of parts

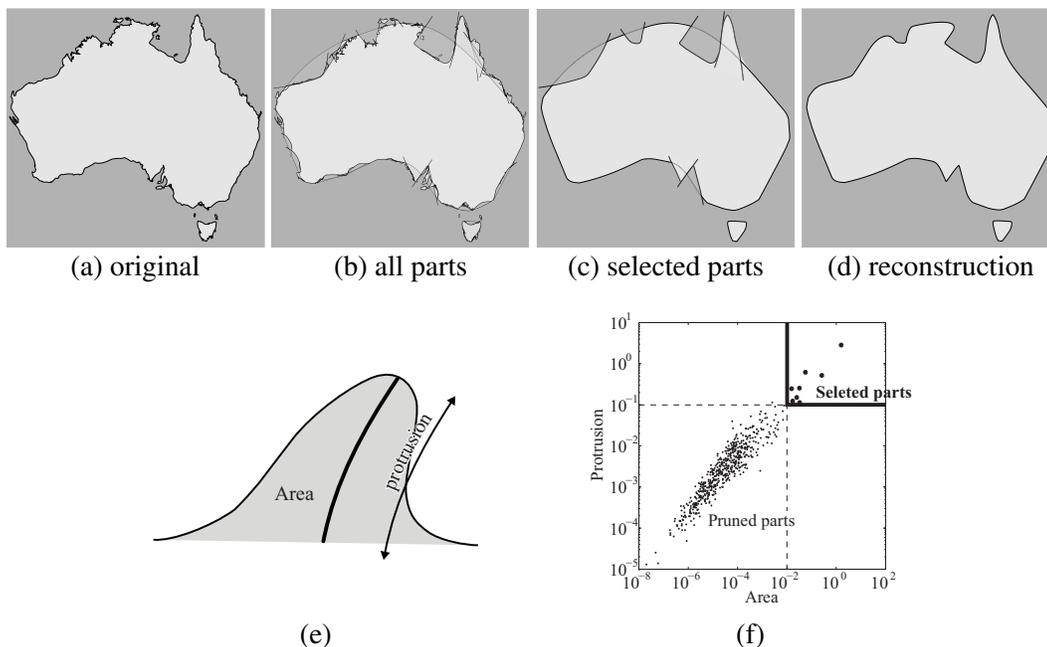
The part regeneration scheme gives us the maximal freedom to choose any reasonable set of parts for abstraction. Thanks to this flexibility, we no longer need to concern ourselves with topological dependency. The only criterion for deciding whether a part should be retained in the final abstraction is its importance. However, as explained in the beginning of this chapter, an obvious measurement of such importance does not exist. Thus, we design a set of tools that allows the user to select the subset of parts interactively.

First, we rely on the the visual *prominence* of a part to prioritize the part. The prominence measures how easily a part grabs visual attention. This differs from part *salience* [Hoffman and Singh, 1997] in that the latter measures not only how prominent a part is, but also how convincingly a fragment of shape contour is perceived as a part, as opposed to a mere subset of a part. Chapter 2 and Chapter 3 described our model of determining if a fragment of shape contour is a part. Here, we only use the components related to the prominence: the area and the protrusion<sup>1</sup>. We measure the area and the protrusion of all parts, which are the two obvious geometric properties related to a part’s prominence. A box selection tool is provided to allow the user to select the most prominent subset of parts based on these two measurements. (Recall that the protrusion of a part is measured as the perpendicular component of the length of the part’s SLS axis—see Section 3.3.5—and the area is the region between the part boundary and the completion curve).

We apply this thresholding scheme to the Australia shape, with key steps that produce the abstraction illustrated by Figure 4.6. In particular, Figure 4.6 (b) is the set of parts collected by

---

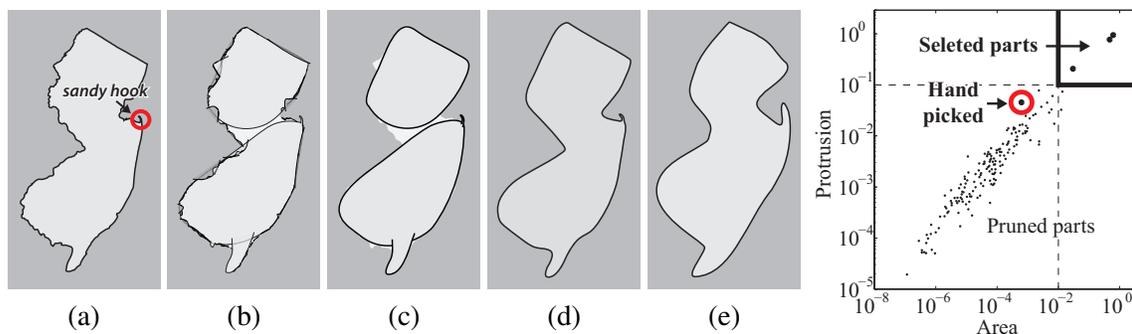
<sup>1</sup> In [Hoffman and Singh, 1997], the protrusion is defined as the ratio of the perimeter the part to the length of the part line, which is a scale independent part distinctiveness measure. As Hoffman and Singh have hypothesized, a bigger protrusion implies a more distinctive visual part



**Figure 4.6:** Selecting a subset of 8 parts for the abstraction of the Australia coastline by their visual prominence. The decomposition in (b) contains 544 parts.

our automatic shape parser described in Chapter 3. These parts provide a complete description of the whole shape, but we do not need all of them. (c) is the set of parts selected using the box selector and the resulting abstraction is shown in (d), which is reconstructed as described in the previous section. Figure 4.6 (e) illustrates the two geometric measurements we used. Each dot in Figure 4.6 (f) represents a pair of measurements in the “area-protrusion” parameter space of each part. In this space, the user can effectively select the most important components of the shape using a box selector.

We also allow the user to pick parts manually for reconstruction because the box selector is not always flexible enough, for two reasons. First, not all the most prominent parts of any shape would be chosen by the box selector since other factors might also affect the part prominence. More importantly, a part might happen to be very important to the user despite its lesser geometric prominence. Figure 4.7(a) shows a map of the state of New Jersey, USA, together with the part decomposition in (b). For New Jersey residents, it is often important for them to have Sandy Hook included in their state map, as circled in Figure 4.7(a), which features a very



**Figure 4.7:** To match the part composition of the artist-drawn New Jersey map (e), the abstraction of the New Jersey map by our system requires a manual selection of the part corresponding to Sandy Hook.

nice beach and the oldest lighthouse of the US<sup>2</sup>. Observe that in the drawing of Figure 4.7(e) by a New Jersey artist, this area is purposely included with very exaggerated prominence. The actual prominence parameters of this part, however, correspond to the circled dot in the “area-protrusion” space shown at right in Figure 4.7. Unfortunately, no simple box selector would include this dot without also including some others that are less critical. Instead of making up an arbitrary prominence function, we simply allow users to override the prominence-based thresholding by clicking the parts they think are important to include. Thus, the user can select the three major parts with the default box thresholding, but can also click the extra part that corresponds to Sandy Hook in order to include it in the reconstruction. This gives a total of four parts as shown in (c). Figure 4.7(d) displays the assembled abstraction. Compared with the artist’s drawing in (e), our abstraction manages to capture the “basic shape” [Robinson et al., 1995] of the map but lacks the ability to depict important features with proper exaggeration and idealization of visual forms.

### 4.3 Summary of our approach

We now have all the elements of our system in place. We summarize our approach, including the heuristics-driven shape analysis and the user-assisted shape abstraction, as follows.

#### Part candidate detection

<sup>2</sup> <http://www.sandy-hook.com/>

1. Trace local symmetries from each curvature extremum including those with negative curvatures. Stop tracing when no more potential basic parts are expected (Section 3.1.3). This step gives the transitional sections of SLS.
2. Locate transitional sections of SLS with significant strength  $\mathcal{T}_n > t_0$  (Section 2.5.4). Build the data structures related to the part candidate based on all transitional sections, including necessary part extensions (Section 2.6), when possible. By calling the elements produced by this step “part candidates”, we emphasize that each of these elements might be a part in certain shape contexts, which we have yet to examine in order to determine whether or not the candidate is indeed a part of the shape.

Compared to the “naive” morphological shape analysis described on page 69, the two steps above explore both the foreground and the background for part candidates. The tracings of SLS are also stopped early for better efficiency.

### **Part decomposition**

3. Compute the CAT, which we currently use to infer the global structure of the shape context (Appendices 3.A and 3.B).
4. Follow the steps described in Section 3.5 on page 85 to construct a graph  $\mathcal{D}$  that infers the formation of each part candidate and constraints how parts should be deleted (Sections 3.3 – 3.5).
5. Accommodate optional user guidance expressed by the strokes. Modify  $\mathcal{D}$  accordingly (Section 3.6.1).
6. Find the strongly connected components of  $\mathcal{D}$  with no outgoing edges. The construction of the graph guarantees that removing any of these components respects both our models of perceptual organization and the user constraints (Section 3.5).
7. Find the group of parts with the smallest scale. Replace its boundary with the completion curve. Each deleted part becomes a part of our shape representation.
8. If the shape is not empty, go back to step 1.

With the entire set of the deleted parts, we are able to recover the original shape. To clarify

structures and to emphasize important components, however, it is often better to show only a more simplified shape with only a small subset of these parts.

### **Abstraction**

9. The user selects a subset of the most prominent parts from the decomposition using the box selector that imposes a set of geometric thresholds. Alternatively, this thresholding can be overridden by clicking and including parts that are important to the user (Section 4.2).
10. Reconstruct the shape by adding back the selected subset of parts, proceeding in the reverse order of the deletion in the decomposition stage (Section 4.1.2).

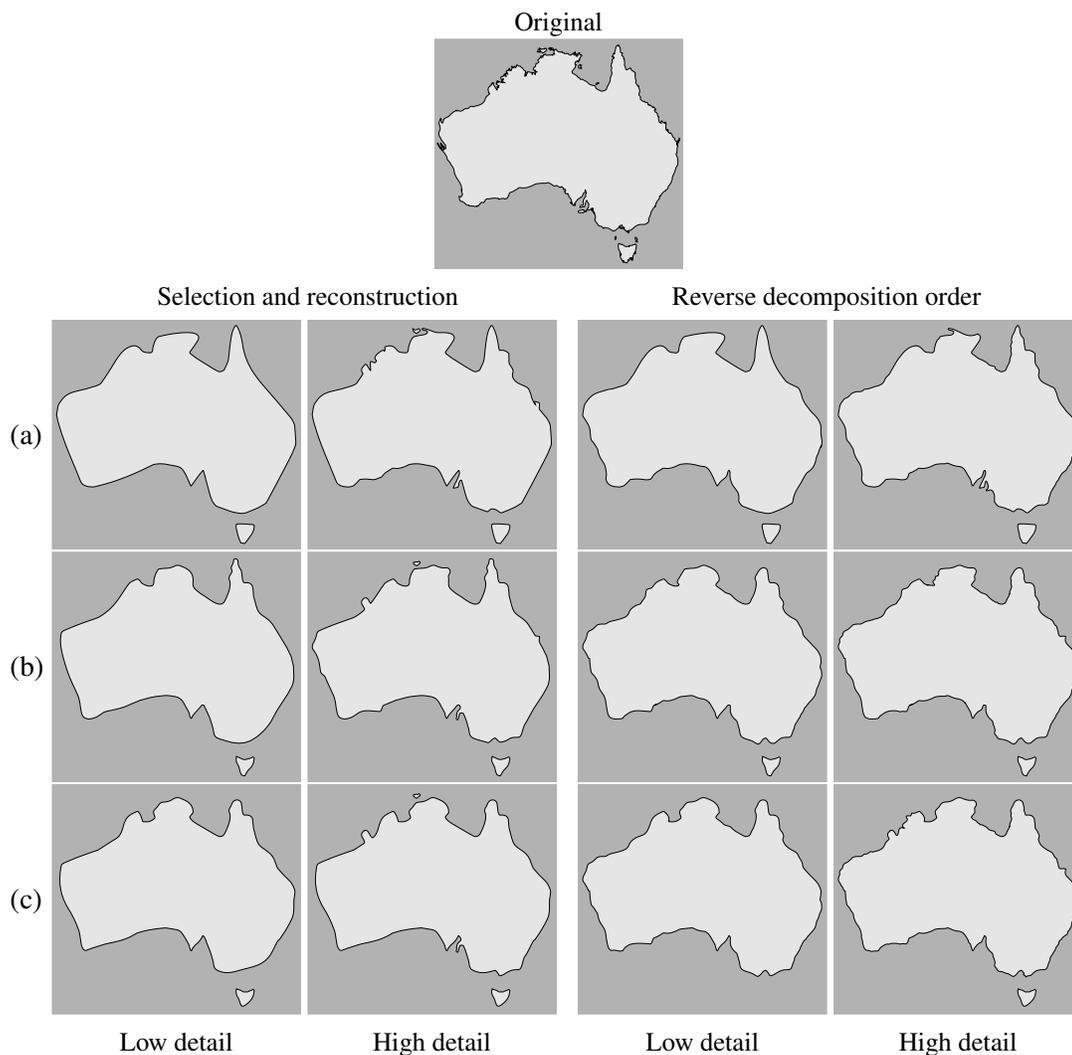
## **4.4 Results and comparisons**

### **4.4.1 Comparing to abstractions with degraded features**

Section 4.3 summarizes the ingredients of our approach to shape analysis and abstracted depiction that are new to the fields of computer vision and graphics. These ingredients are: (1) the shape analysis driven by a set of heuristics that models human perceptual organization, as discussed in Chapter 3; (2) consideration of not only parts from the foreground, but also those from the background of the shape, which better accommodates the possible explanation of subtractive composition; (3) a shape reconstruction scheme that effectively eliminates the topological dependency among parts introduced in the decomposition step, which makes it possible to retain the essential parts in the abstracted depiction and omit any less prominent components.

Figure 4.8 shows abstractions of the Australia map with different combinations of ingredients. In particular, the abstractions on the first row are generated based on the description (in terms of parts) by our shape analyzer with all the heuristics turned on. The set of parts that leads to the abstractions on the second row is produced by our shape parser without heuristics. To generate the third row, we allow the parser to consider foreground parts only. As indicated by the results, the abstraction derived from a better description of the shape in terms of parts (row (a)) also produces a simpler shape composition with cleaner strokes. The components presented

also resemble the components of the original shape better. Observe that in the abstractions on the second and the third rows, many parts are distorted as a result of the original distortion in the decomposition stage, which is not surprising since the analyzer would not preserve the shape structure if the heuristics were not enforced.



**Figure 4.8:** Disabling features: (a) with heuristics; (b) no heuristics; (c) no heuristics, foreground only. The left two columns show reconstructions from selected parts. The right two columns use reverse decomposition order.

Moreover, Figure 4.8 provides a strong testimony to the importance of freedom in choosing the correct set of parts. The rightmost two columns of Figure 4.8 show the results created by simply reversing the deletion to a point that all the important parts that are present in the corresponding reconstructed abstraction are reattached. As the figures show, there are wiggles on the boundaries caused by the attachment of shallow, less prominent parts. Instead of expressing

any important features, they are reattached only to maintain the topological consistency, so it is possible for more prominent parts to be attached. Unfortunately, this is done at the cost of the clarity of the lines and the expressiveness of the overall abstraction.

#### 4.4.2 Comparing to traditional geometric simplifications

We further compare our shape abstraction to three traditional geometric simplification techniques, which were introduced briefly in Chapter 1:

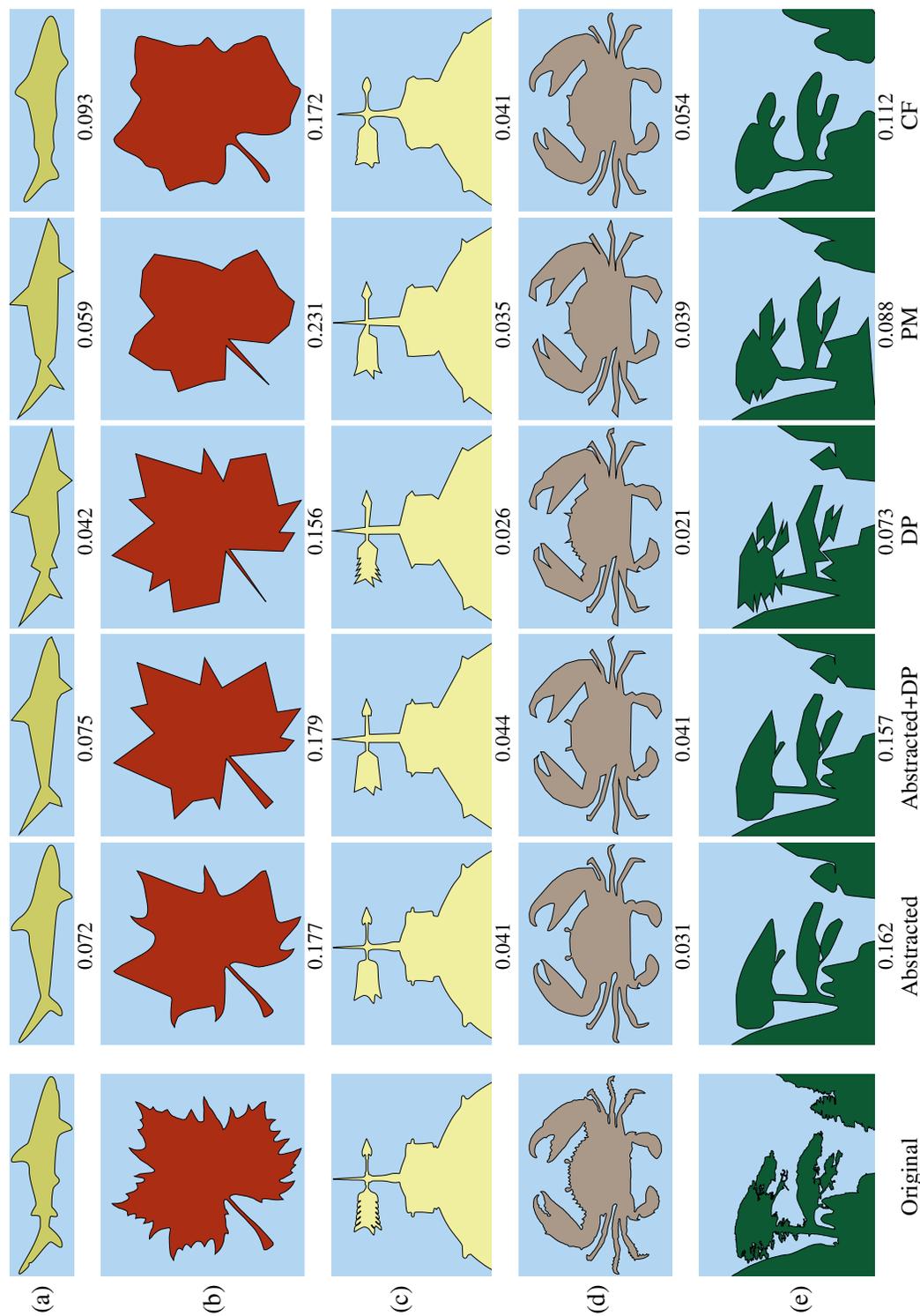
**Douglas-Peucker (DP) algorithm** [Douglas and Peucker, 1973], which approximates the original shape by top-down greedy vertex sampling;

**Progressive Mesh (PM) representation** [Garland and Heckbert, 1997; Hoppe, 1996], which is a paradigm of bottom-up approximation that simplifies the shape by collapsing edges and adjusting vertices to locally minimize the area errors;

**Curvature Flow (CF) smoothing** [Desbrun et al., 1999], which simplifies the shape in the frequency domain and produces results with smoothed boundary curves.

The abstractions and simplifications of five shapes are shown in Figure 4.9. The set of parts retained for the abstracted shark shape is selected using the area and protrusion thresholds explained in Section 4.2. The other simplifications are hand-tuned to the degree that the same set of features are preserved. To allow polygon-to-polygon comparisons to DP and PM, we also decimate our abstracted shape by DP such that the decimated polygons of each shape have the same number of vertices.

We have already seen the decomposition of the shark shape in Figures 3.20(a) and 3.20(b) on pages 86–87. Here we present the abstraction in Figure 4.9(a). Since the processing units are the parts, the integrity of their forms is naturally preserved in the resulting abstraction. Compared to the clearly spelled-out part structure in our abstraction, the decimated shapes using either DP or PM contain edges that span multiple neighboring fins due to the lack of a perceptually based structural representation—one of the factors that also causes the “fused” parts seen in children’s drawings [Arnheim, 2004a, also introduced briefly in Chapter 1]. Similar problems with DP or PM can be found in their simplifications of the rest of the shapes. For



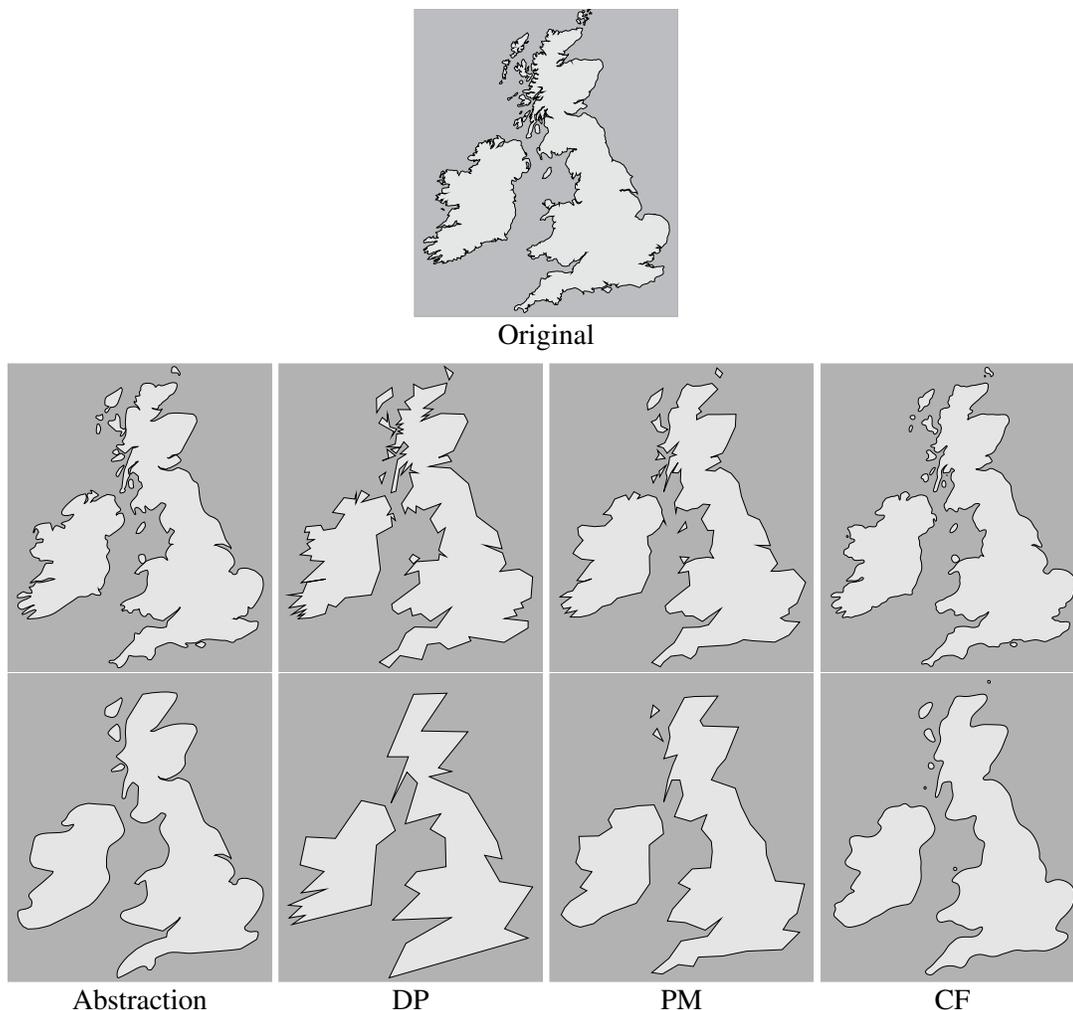
**Figure 4.9:** Automatic abstractions and simplifications of various shapes: (a) shark (423 vertices, 12 parts); (b) maple leaf (576 vertices, 73 parts); (c) lighthouse (1898 vertices, 57 parts); (d) crab (1620 vertices, 288 parts); (e) pine (4298 vertices, 1035 parts) The number below each simplified shape is the Hausdorff distances from the original. All original shapes are normalized in size so that the radii of their bounding boxes are 1.

example, the maple leaf simplified by DP has its lower lobes distorted due to the existence of a few longer teeth nearby. The feathering of the lighthouse weathervane and the teeth of the crab are all fused together, and both approaches distort the pine’s crown in an irregular manner. In contrast, our approach omits some of the minor features (such as the “teeth” on top of the carapace and cheliped claws of the crab shape) and instead dedicates vertices to more important parts. Our abstraction is the only one to fully preserves the forms of the crab eyes, and it ignores the minor teeth of the maple leaf so that the composition of the lobes can be expressed more clearly. CF does smooth the shapes in a pleasant way. However, it not only fails to eliminate some of the minor parts cleanly, it also distorts other salient parts. The two fins on the front of the shark shape are shortened. Sharp features have all disappeared, though they were salient on the maple lobes, as well as on the arrow head of the weathervane. The smoothed pine crown exhibits a very different structure, something like a tortoise shape.

Figure 4.9 also reveals that a simplification using DP always carries the smallest error in terms of the Hausdorff distance to the original shape; and errors in our abstractions in terms of this distance metric are typically 50% – 100% larger. However, the DP method does not seem to produce the most preferable results that best express the original forms—we will further confirm this in Section 4.5. Indeed, a geometrically defined distance metric may not necessarily align well with the distance as perceived by the human visual system.

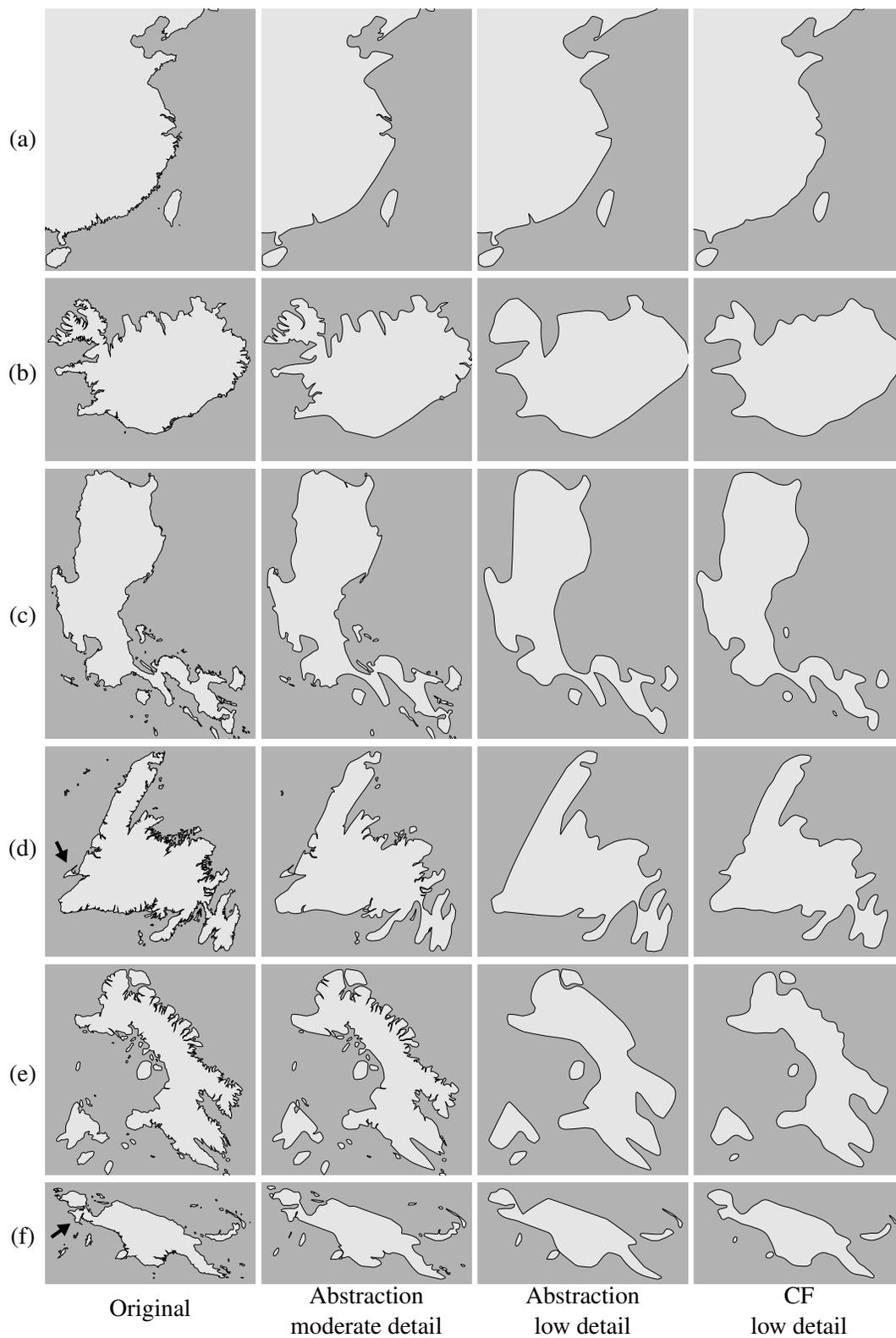
With its default settings, our shape analysis system does not always do a good job of parsing all the shapes of Figure 4.9. In particular, the automatic analyses of the maple leaf and the crab fail to produce shape descriptions that respect human perception. However, since most of the parts of both shapes appear in the foreground, it makes sense to inform the system of this. Following the trick described at the end of Section 3.6.1, we bias the analysis procedure towards deleting foreground parts by setting the foreground distance weight  $w_l = \frac{1}{8}$ .

Figure 4.10 shows a maps of the UK and Ireland, including the abstractions and simplifications at two different levels of detail, using four different approaches. The abstractions created using our approach are based on the set of parts generated by our automatic parser with the default parameter settings. (For shapes like coastline maps, parts are formed from both the

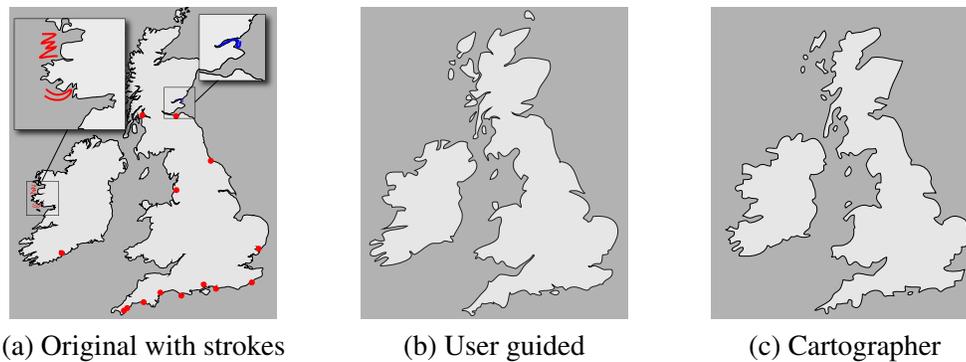


**Figure 4.10:** Automatic abstractions and simplifications of the UK and Ireland (7600 vertices, decomposed into 950 parts; abstractions use 114 (top) and 31 parts (bottom)).

foreground and the background. The default parameters are thus more appropriate.) All approaches produce results that are reasonable but have different styles. In particular, features retained in our abstraction are presented literally, while omitted features have no effect on the geometries of the result. Each of the other approaches tries to include all the components by spreading the error, which is caused by the limits in the complexity of the representation, across the entire shape. Consequently, minor details can affect the final shape more or less. They can form meaningless minor features along the boundary, which cannot be mapped to any real feature in the original shape, but which may also reflect certain geometric textures on the boundary. In particular, such textures produce a result with a “map-like” style which is appealing, especially when the viewer does not care what the individual features actually are.



**Figure 4.11:** More maps: (a) eastern China; (b) Iceland; (c) Luzon Island (Philippines); (d) Newfoundland; (e) Baffin Island (Canada); (f) the island of New Guinea.



**Figure 4.12:** User guided results from the strokes in (a), compared to one produced by a cartographer (after [Robinson et al. \[1995\]](#)).

Figure 4.11 shows more maps together with their abstracted depiction in both low and moderate levels of detail both using our approach. For comparison purposes, it also includes on the rightmost column the results produced by curvature flow. Again, compared to the results of CF our approach produces crystal clear imagery. Every section of the boundary curves can be mapped directly to a perceptual part of the original shape. The smoothing-based approach removes the high frequency geometry coherently. Neighboring minor features are averaged together, resulting in shallow wiggles along the shape boundary. The differences between these two approaches are particularly evident at places such as the coastline of southeast China and northeast of Baffin Island. Figure 4.11 also shows a potential risk that our approach takes: if an eliminated part happens to be important to the user, the resulting abstraction could be very distracting. It is very possible that the viewer would feel that the distorted presentation of these features is at least better than nothing, even though they might not resemble the original forms at all. The *Port au Port Peninsula* of Newfoundland Island, and the south portion of the *Bomberai Peninsula* of western New Guinea (marked by arrows in (d) and (f), respectively) are two such cases<sup>3</sup>.

Of course, the users can manually include in the abstraction whatever parts they feel necessary just by clicking them, as described in Section 4.2. Figure 4.12 shows an example with a moderate amount of user interactions involved, matching our abstraction to a cartographer’s

<sup>3</sup> The overall shape of New Guinea is famous for its resemblance to a bird. The Bomberai Peninsula, together with the *Bird’s Head Peninsula* to its north, form the head of the bird. By removing the south part of the Bomberai Peninsula, our low-detail abstraction drops the “jaw” of the head, which is particularly distracting to people familiar with the geography of New Guinea.

drawing obtained from Robinson et al. [1995], which was also shown by Figure 1.1 in the very beginning of this thesis. To create a similar set of “basic shapes”, we interact with the system in both the shape analysis stage and the abstraction stage. In particular, 2 “keep whole” and 3 “not a part” strokes (3 of which can be seen in the insets of (a)) were used to modify the default interpretation of the part composition. In addition, we manually selected 14 tiny parts (red dots), which would have been eliminated by the geometric thresholds. The abstracted shape is shown in Figure 4.12(b), side by side with Robinson et al.’s in (c). We feel that our abstraction produces a part composition similar to what the cartographer has drawn. However, the cartographer rearranges the configurations of the parts, and distorts their forms so that they look more distinctive while maintaining the coherence of the whole.

## 4.5 Assessment

To assess our approach to shape abstraction in comparison to the existing geometric simplification methods, a perceptual preference study is conducted. In particular, we compare our approach to the Douglas-Peucker (DP) algorithm [Douglas and Peucker, 1973] as well as Progressive Meshes (PM) [Hoppe, 1996] with a quadratic error metric [Garland and Heckbert, 1997].

The difficulty here is that our approach has a different goal from DP or PM. Our abstraction aims to clarify the shape structure by eliminating irrelevant component parts, which we hypothesize reduces the complexity of the representation for the human visual system. DP or PM, on the other hand, tries to reduce the number of vertices in the shape. Their goal is to reduce the complexity of the shape representation for computer storage. This difference in goals leads to differences in the visual styles. Ours produces results with smooth shape boundaries since down-sampling for smaller spatial complexity is not our concern; the results produced by DP or PM are decimated polygons, which are often pointy.

To reduce the effect of such differences as much as possible, we avoid direct comparisons between simplifications from the two different styles. Instead, when compared to the result from the DP or PM decimation in each test case, the abstracted shape is further simplified with the

same decimation technique. As a result, each original shape has a total of four simplified versions, grouped in two pairs based on the decimation technique applied. For Douglas-Peucker, the comparison is made between:

**DPO:** Douglas-Peucker decimation of the original shape

**DPA:** Douglas-Peucker decimation of the abstracted shape

And for the Progressive Mesh decimation, it is between

**PMO:** Progressive Mesh decimation of the original shape

**PMA:** Progressive Mesh decimation of the abstracted shape

Thus, the difference between the members of each pair is whether our abstraction technique is applied as an intermediate step. Note that both decimation methods simplify the shape such that all the vertices or edges are taken into account. With the extra abstraction step in DPA and PMA, less prominent parts are eliminated. The subsequent decimation is thus not “distracted” by the presence of these parts. We thus expect that, if the numbers of vertices are the same, the decimation from the abstraction contains fewer parts than the corresponding decimation from the original, while the structure of the preserved parts in the decimated abstraction are better preserved. Now we use a perceptual preference study to answer this question: whether, and when, does this extra abstraction step help in creating more expressive depictions. In the study, we control the amount of part distortion by decimating the shape to a varying numbers of retained vertices. We are particularly interested in cases when the abstracted shapes are decimated to a level such that the preserved parts are about to be distorted. We hope to find out, in this case, how do these results compare to the decimation from the original shapes with the same number of vertices. At this point, the corresponding parts in DPO and PMO will likely be distorted, but we also expect that they contain more (distorted) detail. If the study shows that people prefer DPA to DPO (and PMA to PMO) when the abstraction is decimated to such a level, we can conclude that the abstraction with fewer details but clearer parts is a more preferable form of shape depiction.

### 4.5.1 The settings of the preference study

We collected 96 coastline maps from the NOAA/National Geophysical Data Center<sup>4</sup>. All maps were rendered by Azimuthal orthographic projection [Snyder, 1987], centered at the point of interest of the original shape.

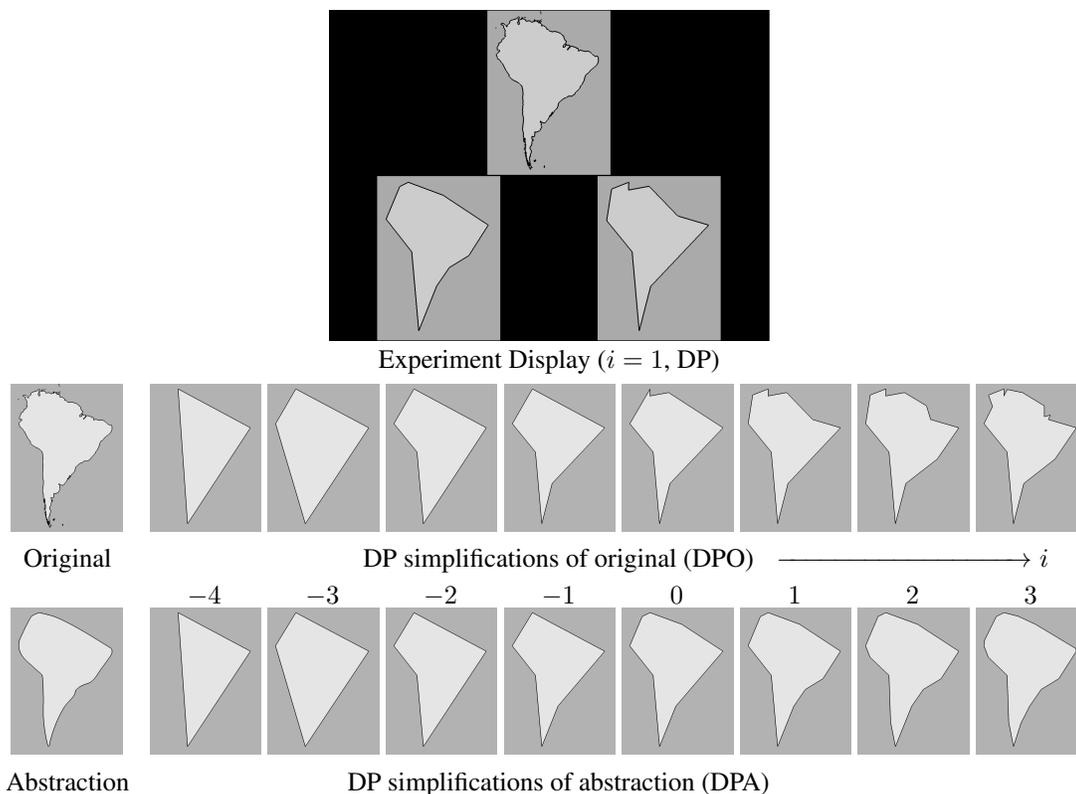
The abstraction of each map was based on the set of automatically computed parts using the default parameter settings described in Section 3.6.2. Each map was abstracted with a manually set protrusion threshold  $\mathbf{p}_0$ . The area threshold was set accordingly:  $\mathbf{a}_0 = \mathbf{p}_0^2$ . We used this constraint in order to make a fair comparison, since both DP and PM use the number of vertices in the decimated shape as the only threshold.

The decimation level decided how many vertices would be shown in both viewports after decimation. We used 8 different levels for each shape:  $2^{\frac{i}{3}}N_0, i = -4, -2, \dots, 3$  where  $N_0$  was an estimate of the minimal number of vertices to be retained to avoid introducing fused parts or interrupting salient axial structures of any parts retained in the abstracted shape. We prepared 25 maps and determined  $N_0$  by hand for each of them for both decimation approaches. Then for each decimation, we performed a linear regression on  $N_0$  using a single independent variable  $\Theta = \sum_P \Psi_P$ , where  $\Psi_P$  denotes the total turning angle of the boundary of part  $P$ , and  $\Theta$  is the sum across all the parts in a shape. Here the total turning angle,  $\Psi_P$ , was merely the discrete version of the integral of the absolute values of curvatures along the part boundary. We found it suitable to use their sum across all parts,  $\Theta$ , as a scale-independent measure of the perceptual complexity of shape contours. For Douglas-Peucker and progressive meshes, the values of  $N_0$  were  $0.65\Theta$  and  $0.60\Theta$ , respectively. For DP, the *coefficient of determination* ( $R^2$ ) of the linear relationship between  $N_0$  and  $\Theta$  was 0.81; and for PM, it was 0.79.

Figure 4.13 shows the settings for the comparison of the South American coastline as simplified by DPO with the same map simplified by DPA. In particular, the top row of Figure 4.13 shows the screen display. The middle row shows the original shape, followed by the results of DPO: 8 simplified versions of the original, each decimated to a particular number of vertices using Douglas-Peucker, as described above. The bottom row shows the abstracted shape and

---

<sup>4</sup> <http://rimmer.ngdc.noaa.gov/>



**Figure 4.13:** Example of an experiment display, and pairs of maps used in the study (decimated using Douglas-Peucker).

the results of DPA: 8 different levels of decimation using the same settings as DPO. The only difference is that these are decimated from the abstracted shape.

The study contains 96 maps with 8 levels of decimation, which yields 768 different (map  $\times$  decimation level) combinations. When assembling the sessions for each subject, we randomly select from these combinations, subject to the following constraints:

- Each subject was assigned 192 combinations, 96 pairs of which were decimated using Douglas-Peucker, and the other 96 using progressive meshes.
- In each decimation method (DP or PM), each map (at any decimation level) was shown exactly once to the subject.
- In each decimation and for each decimation level, the subject was shown 12 different maps. The 12 maps decimated by Douglas-Peucker do not overlap the 12 decimated by progressive meshes.

- One particular (map  $\times$  decimation level) combination was shown only once for a group of 8 subjects. That is, every 8 subjects cover the whole grid once.
- Each subject had an equal chance of getting any of the (map  $\times$  decimation level) combination. The order of the trials shown was also randomly assigned.

Each trial was preceded by a centered red circle that was shown for 500ms for visual fixation purpose. Then the original shape, its decimated version, and the decimated abstraction (using the same decimation approach) were displayed simultaneously in three viewports of the same size arranged in two rows. The viewport displaying the original shape was centered on the top row and the other two, showing the decimated shapes, were on the bottom, as shown in Figure 4.13. Whether the simplification from the abstraction was shown on the left or on the right was randomized and counterbalanced across all trials per decimation level per subject. We applied a *two alternative forced choice* (2-AFC) scheme, which asked the subject to directly compare the results of DPO and DPA, or PMO and PMA and choose the simplification they liked better. In making the choice, they were told to achieve a balance between clarity and expressiveness of the presentation, and a choice had to be made for each trial.

The display we used was a Dell Ultrasharp 2005FP (99.5 PPI), and the distance from the back of the chair to the monitor was fixed at 1 meter. The study involved 16 subjects—all graduate students from Rutgers University. All had normal or corrected-to-normal vision. None of the subjects had participated in a similar experiment before. Before the formal experiment began, all the subjects had a practice session to familiarize themselves with the assessment system.

The 16 participants were divided into two groups, each covering the (map  $\times$  decimation level) grid exactly once for each decimation method. Consequently, each combination was judged exactly twice. Figure 4.14 shows the percentage distributions of all three possible outcomes. The green area shows the percentage of both judgements favoring the decimated abstraction, whereas judgements favoring the decimated original both times is shown in blue. The light yellow area in the middle indicates the percentage of ties in judgements. Intuitively, a larger green area means that the abstraction, as the intermediate step prior to the decimation, improves the results more effectively. From Figure 4.14, we find that subjects favored the decimated abstraction when the decimation caused structural destruction only slightly or not at

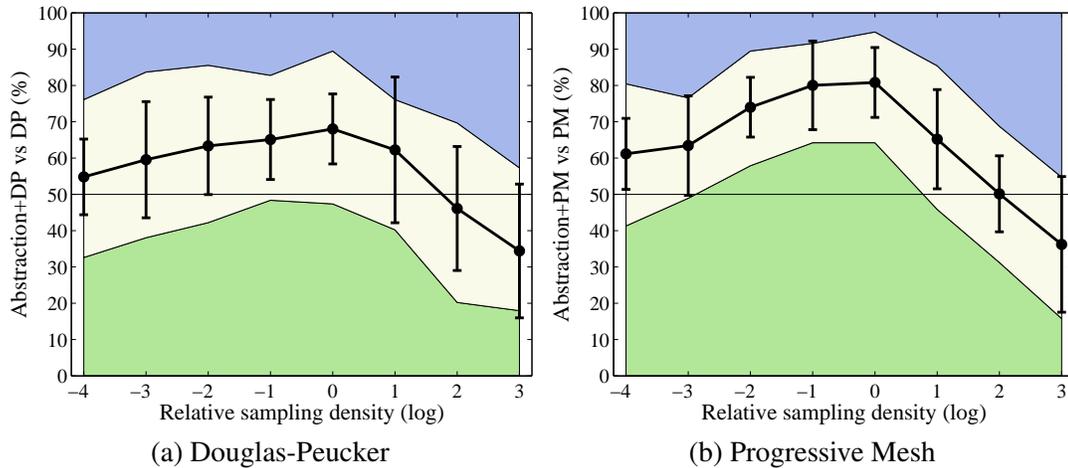
all, which is evident from the high green regions around the centers of both decimation methods. Adding more vertices to the already abstracted shape did not add any shape information because minor parts had been eliminated beforehand, but with the extra vertices a decimation shape from the original was offered a chance to present extra features. Thus, at these levels the abstraction as an intermediate step was not as effective. Going in the opposite direction, when decimated too much the whole shape structure was broken, no matter whether it was from the original or from the abstraction. In such cases (mostly with  $i = -4, -3$ ), the subjects' preferences reverted to chance levels.

#### 4.5.2 Significance of shape abstraction

Special care needs to be taken when formally analyzing the significance of the effectiveness of our approach at each decimation level. As a typical 2-AFC study, it is subjected to a binomial analysis of the proportion of trials in which the abstracted shapes are preferred [Siegel and Castellan, 1988]. However, we could not directly apply this test to the complete set of trials at each decimation level since each level had 16 subjects involved. Each subject had their own unique aesthetic taste and observational habits that could very likely affect their judgements. This adds extra variance to the distribution, which is commonly called the *overdispersion*<sup>5</sup>. Note that the assumption of the binomial distribution is still valid for the set of judgements made by a single subject. Thus, for each subject  $j$ , we assume that the percentage of preferred shapes, where the decimation at level  $i$  from the abstraction is more effective for the subject than the decimation from the original at the same level, is a binomial random variable  $X_{i,j} \sim B(n, p_{i,j})$ , where  $p_{i,j}$ , denoting our approach's success rate in each trial at level  $i$  for subject  $j$ , is another random variable. Our abstraction approach at level  $i$  is significant if we can show that the expectation,  $E(p_{i,j})$ , is significantly larger than  $\frac{1}{2}$ . Here we turn to the family of *Beta* functions

---

<sup>5</sup> We can actually use Tarone's Z statistic to test whether our data can be interpreted by a binomial distribution [Tarone, 1979]. In our study, with 95% confidence, we reject the null hypothesis that the underlying distribution of the decimated abstraction being preferred to the corresponding decimated original is binomial when  $i = -1, 3$  for PM and when  $i = -3, 1, 2, 3$  for DP.



**Figure 4.14:** Results from shape preference study. The green areas indicate the percentage of pairs where PMA or DPA were selected both times, and the blue areas where PMO or DPO were selected both times (the light yellow areas were a tie)

to formalize  $p_{i,j}$ 's distribution, e.g.,  $p_{i,j} \sim \text{Beta}(\alpha_i, \beta_i)$ , i.i.d, where

$$\text{Beta}(p; \alpha, \beta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1} (1-p)^{\beta-1}$$

is the Beta function with  $E(p) = \frac{\alpha}{\alpha+\beta}$  and  $\text{Var}(p) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$

The generality of this setup lies in that the Beta functions can have a wide variety of forms in defining the density over the interval of  $[0, 1]$  by using different combinations of  $\alpha$  and  $\beta$ . With this compound arrangement,  $X_{i,j}$  turns out to be a random variable of the so-called *Beta-Binomial* distribution, typically denoted as  $\mathbf{P}(n, \mu_i, M_i)$ , where  $\mu_i = \frac{\alpha}{\alpha+\beta}$  is the expected mean of the Beta function;  $M_i = \alpha + \beta$ ; and in our case,  $n = 12$ .

At the decimation level  $i$ , each  $X_{i,j}$  is a sample of this random variable. It is possible to estimate the confidence intervals of  $\mu_i = E(p_{i,j})$  as well as other parameters of this distribution, given the set of observations  $\{X_{i,j}\}, j = 1, \dots, N$  using maximum likelihood (ML) [Smith, 1983]. A more implementation-friendly description of this approach is given later by Ennis and Bi. [1998a;b]. Note that the first moment estimate of  $\mu_i$  is simply  $\hat{\mu}_i = \frac{\sum_j X_{i,j}}{N}$ , which is quite close to that of the ML estimates.

We present the estimated parameters in Figure 4.14. The curves in both DP and PM show the values of estimated  $\hat{\mu}_i$  for each of  $i = -4, -3, \dots, 3$ . The error bars show the corresponding

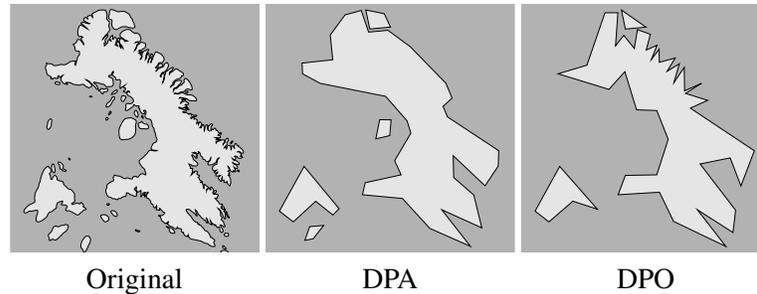
95% confidence intervals at these decimation levels. It is now evident that, for each of DP and PM, there is a wide range of decimation levels, at which the intermediate abstraction step makes a significant improvement and leads to more concise and more expressive results. In particular, at these levels the amount of decimation breaks the part structure of the abstracted shape only slightly or not at all. Subjects preferred DPA significantly over DPO for  $i = -1$  and  $i = 0$  ( $p < 0.02$ ). This preference for PMA over PMO was even stronger for  $i = -2, -1$  and  $0$  ( $p < 0.005$ ).

Finally, we note that the Beta-Binomial model has previously been applied to evaluate the effects of such factors as toxicological treatments [Williams, 1975], policy changes [Gange et al., 1996], and sensory analysis [Anguloa et al., 2007; Ennis and Bi., 1998a]. This model for the analysis of a 2-AFC experiment was first studied by Morrison [1978]. Compared to the traditional binomial analysis, the Beta-Binomial model takes into account differences among subjects and hence provides a more reliable significance test. To the best of our knowledge, our preference study applies this model for the first time to evaluate the effectiveness of certain visual styles. We believe this sets a new paradigm for the evaluation of non-photorealistic rendering research.

#### 4.6 Limitations of abstraction in terms of parts

Besides showing the success of our approach, the assessment also confirms the observation we made in Section 4.4.2 about the abstraction of Newfoundland Island: the removal of Port au Port Peninsula produces a straight coastline on the west, which turns out to be distracting to observers who think the peninsula is important. Leaving out parts that are important to some of the viewers is the risk that the automatic abstraction takes. It was the main reason that the decimated abstraction (either DPA or PMA) lost a trial in the assessment. It could be even more distracting if this happens to the abstraction of a shape with explicit functional semantics attached to its parts, such as those shown in Figure 4.9. For example, our abstraction of the crab shape immediately becomes worse than the result of DP or PM if one of the crab's eyes is missing. However, this potential failure does not contradict the theory that the visual process of recognition is based on the visual process of shape parsing. Instead, it strengthens the theory

that underlies our abstraction: any parts that are used for recognition should be preserved in the abstracted results, because depiction, as a inverse of vision procedure, needs to make sure it retains those elements used for recognition.



**Figure 4.15:** Northeastern coastline of Baffin Island with different visual styles of presentations. Both shapes of DPA and DPO are decimated to 50 vertices.

The assessment also reveals that when a set of parts exhibits a certain pattern on the shape boundary, removing them completely and leaving a smooth, blank contour curve is not the best treatment for some people. For example, our low-detailed abstraction of Baffin Island has this effect on the north east coastline shown in the middle of Figure 4.15. While decimating this area of the shape using Douglas-Peucker results in unfaithful compositions of parts, the existence of shape patterns is somewhat spelled out. During the informal interviews with some of the subjects after the experiment, when asked for the rationale for preferring DPO to DPA, they explained their judgement based on the two different approaches to the treatments of patterns. Obviously, the proper treatment of shape textures is an interesting direction of extending our work. Our abstracted depiction system works the best when a shape can be described as a set of parts. But parts are not the the whole story when it comes to shape representation. Other perceptual elements can also be explored (and sometimes as better candidates) for effective abstraction and structural shape editing. For example, in Mehra et al.'s abstraction system [2009], the abstracted shape is reconstructed from a representation in terms of feature curves.

## Chapter 5

### Discussion and Future Work

This thesis has presented our work on structurally based shape representation and abstraction. A set of perceptually driven heuristics is proposed to guide the search for a hierarchical shape structure that respects human perceptual organization. From the algorithmic point of view, our approach attempts to recover the procedural modeling using reverse engineering [Várady and Martin, 2002], which is a popular paradigm followed by various automatic parametric feature recognition algorithms in CAD systems [Han J.H. and W.C., 2000; W.C., 1995]. However, even for man-made shapes such as mechanical parts, recovering the modeling procedure has been a difficult problem in the field of CAD research. The task is obviously even harder for finding the part structure of a general shape [Pauly et al., 2008]. We feel we have made concrete progress on this problem.

Due to the difficulty of the task, we limit the scope to closed 2D shapes with no self-intersections. Even so, it is still an open-ended problem characterized by multiple plausible explanations of the shape. Further progress on this track is possible, but requires a deeper understanding of human perceptual organization and related computational approaches.

#### 5.1 Contribution

Part-based shape representation is an old idea in computer vision, and it is also not new to use a structural representation for simplification. The main contribution of this thesis is to propose a combination of elements that we show are each important to obtain favorable results within this framework for abstracted depiction, in particular a model of perceptual organization that guides the decomposition process. The computation of the part-based shape structure has made

it possible to create abstracted shapes at different levels of detail, and it opens the door to structurally-based shape editing and sophisticated part manipulation, including stylization and exaggeration of 2D shapes in terms of parts.

The abstraction system we have built shows that simplification and reconstruction based on this part-based representation leads to greater effectiveness in shape depiction by using a set of fewer parts to obtain a clearer shape structure. This is then confirmed by our perceptual preference study, which demonstrates the use of a Beta-Binomial test to handle the problem of overdispersion introduced by the differences in aesthetic tastes of different subjects.

Our formalization of transition region between parts offers the first computational model in vision research that describes the extent and salience of parts and their connections. Its mathematical characterization also nicely unifies previous theories of parts of visual forms. Computationally, it enables the clean removal of parts without leaving traces. It represents a significant improvement over the scheme used by Rom and Médioni [1993] in their morphological shape analysis system.

There are also a few ingredients in our work that we feel are interesting to related fields in part-based shape analysis and computer depiction. We point out that considering parts in both the foreground and the background is justified by ongoing work in perception [Kim and Feldman, 2009]. We also demonstrate through the comparisons in Figure 4.8 that it is important to consider parts from both sides of the shape boundary. Figure 4.8 also shows the importance of a flexible assembly scheme for shape abstraction that allows freedom in choosing the appropriate set of retained parts. Figure 4.8 further demonstrates that an abstracted depiction can be more expressive if it is based on a shape description that respects visual explanations better.

Based on our models of perceptual organization, we have introduced a graph algorithm to disambiguate among multiple explanations of the shape. Our system will certainly benefit from even more flexible approaches to shape assembly and better models of perceptual organization. In particular, we lack a good measure of how symmetric, or organized, a given shape is. Consequently, we lack a way of assessing whether a symmetry is improved. Currently, we avoid this limitation by using a simple (but not general) approach in terms of heuristics (Sections 3.3 and 3.4). We also avoid the difficulties in robustly computing the major approximate axial

structures by simply finding coherent CAT axes (Appendices 3.A and 3.B). As explained in Section 3.4, not all salient structures are present in the CAT. We are optimistic about future work in these directions.

## 5.2 Limitations of the representation

The structural representation in terms of parts offers the opportunity for the abstraction system to effectively allocate strokes to present important features. However, our current implementation of this representation also limits the abstraction systems in various ways.

### 5.2.1 The peanut dilemma

In some cases, the system proceeds by deleting a part merely to avoid breaking a coherent axial structure. However, this is not consistent with artistic drawing as we claimed in the very beginning of the thesis, since a “part” whose only function is to extend the axial structure might not be the most salient among its neighboring candidates when it is deleted.

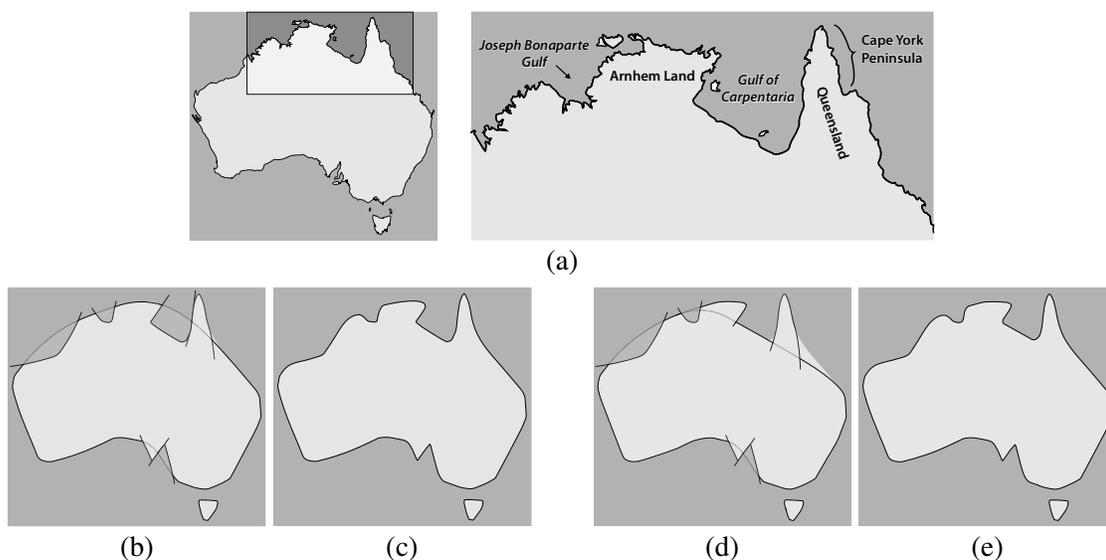
Take the shape of a two-seed peanut as an example (see Figure 2.21 on page 55). In the extreme case, when only one part is allowed to represent the entire peanut, we would choose to approximate both halves with a single part. This can be done by deleting the two indentations between the seeds. However, the indentations should not be perceived as parts. On the other hand, if we choose to describe the peanut with two parts, our system can never produce a reasonable abstraction with a single part that contains both seeds. Unfortunately, our current framework forces the system to choose one representation or the other. A mechanism that effectively accommodates both explanations and allows the flexibility of choosing the representation at any levels of detail is desirable.

### 5.2.2 Effects of figure/ground discrimination on abstraction

A closer study of the experiment described in Section 3.6.2 shows that the system makes a significantly number of mistakes by deleting part candidates which are inconsistent with the “not

a part” constraints. The purpose of these constraints, as described in Section 3.6, is to tell the shape parser that, locally, which side of the shape boundary constitutes the part. Unfortunately, this leads to the still unresolved problem of figure-ground discrimination. Figural assignment in human perception involves complex visual processing. Non-geometric factors could play important roles in discriminating the figure from the ground, e.g., the viewing orientation [Vecera et al., 2002], the type of shape [Ren et al., 2006], and more often, the semantics attached to the part [Peterson and Gibson, 1994]. The figure-ground discrimination problem resides in our shape decomposition due to the ambiguities introduced by considering both foreground and background parts.

On the other hand, two neighboring parts on different sides of the shape boundary could both be represented as figures in the human visual system. In geographic shapes, a peninsula with a neighboring bay is a typical example. Ideally, the structural representations proposed in Section 5.2.1 could be extended to accommodate such ambiguity by including parts from both sides. Our current framework, however, forces the parser to make a binary choice by selecting



**Figure 5.1:** (a) The original Australia coastline with lands and gulfs in northern Australia labeled. (b) Automatic parsing results, with the *Gulf of Carpentaria* parsed as the figure and the *Arnhem Land* as the ground. (c) The resulting abstraction. (d) The reverse explanation achieved by user guidance. (e) The abstraction results from (d), which appears very similar to (c). Note the difference between the tips of Arnhem Lands of the two abstractions. It is less sharp in (c) since it is considered a transitional boundary of the *Gulf of Carpentaria* and is regenerated using a Hermite curve.

parts from one side or the other at each step of the decomposition, instead of both.

Figure 5.1 shows two plausible explanations for the the northern Australia coastline. By default, the system fails to recognize *Arnhem Land* as the figure due to an earlier deletion of the *Joseph Bonaparte Gulf*. The commitment to treat the Joseph Bonaparte Gulf as a part (as the figure) on the background implies considering the neighboring part (with a considerably larger scale) as the ground. Luckily, the reconstruction process still produces a prominent foreground part, thanks to the explicit modeling of the transition—recall that the transitional region of a part is typically also the body of a neighboring part candidate on the other side of the shape boundary. However, the luck resulting from hiding our choice of the figure and the ground is not reliable if any of the wrongly discriminated parts are manipulated, since any dynamics to the parts reveals the figural assignment immediately [Barenholtz and Feldman, 2006].

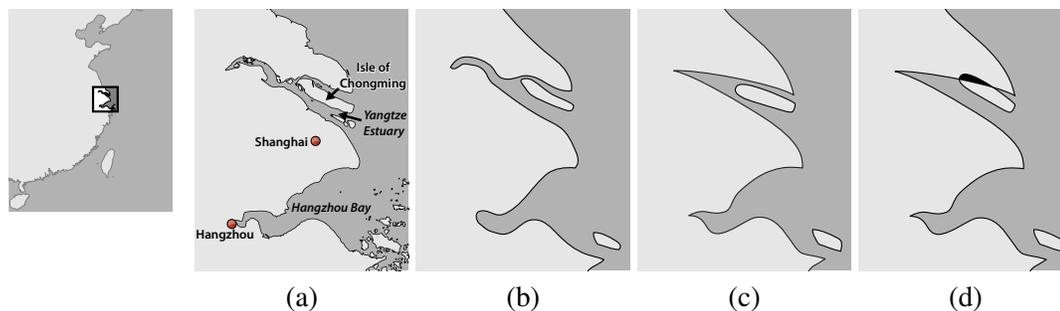
### 5.3 Future work

Art is not what you see but what you make others see.

—Edger Degas

Edger Degas’s teaching and examples from cartographers and artists tell us that we do not need to be too faithful in presenting the forms of shape components (recall the cartographer’s drawing of the UK and Ireland shown in Figure 4.12(c), and the very exaggerated Sandy Hook in the New Jersey map in Figure 4.7(e)). The abstracted depiction does not stop at merely selectively showing or hiding parts of the shape description. Instead, it involves a wide range of sophisticated manipulations and creative distortion of part forms in terms of a denotation system [Willats, 1997, Part II] in order to present the shape structure in a most effective manner. Further more, in terms of a drawing system [Willats, 1997, Part I], an artist should also accordingly adjust the positions and orientations of these parts to accommodate the changes in their forms and to maintain the overall coherence of the shape.

With a given set of parts, our current abstraction system does not necessarily render a shape in its simplest form. This is because some parts can be further simplified. Curve fairing can be used to remove non-critical corners provided that important ones are preserved; also, curvy



**Figure 5.2:** (a) The map of Yangtze River Delta; (b) our mid-level abstraction; (c) a possible depiction that abstracts the part of Yangtze Estuary; (d) intersection of different parts, which necessitates the global adjustments of the positions and orientations of parts.

axial structures are often seen to be straightened out in cartographic practices. Figure 5.2 shows a map of the *Yangtze River Delta*, which includes the metropolitan areas around Shanghai. The entire *Yangtze Estuary* is treated as a single part. In our current abstraction system, we only have two choices: either to consider the Yangtze Estuary an important part and include it in the abstraction as in Figure 5.2(b), or to simply omit it<sup>1</sup>. In cartography, however, this part is often depicted as an indentation with a much simpler axial structure. See Figure 5.2(c).

We call this kind of simplification of individual parts the “per-part” abstraction, which is one possible perceptually driven part operation for shape generalization. We have already seen other interesting part operations such as exaggeration, idealization, etc., all suggesting exciting future work that enhances the function of the denotation system. However, it is not trivial to implement the part operations. Topological consistency needs to be accounted for when geometric operations are applied to the parts separately. Observe that in the example of Figure 5.2(c), the *Isle of Chongming*, which also takes a simpler form, is nudged to the south slightly, and its major axis becomes slightly more horizontal. If were not for this small adjustment, in the resulting abstracted depiction the boundary of the island and the bank of the estuary would have intersected with each other as in Figure 5.2(d).

Recall that our shape analysis has faced difficulties due to the lack of both reliable contour grouping and symmetry measurements, and a general algorithm that computes the approximate axial structure of a given shape. Our current solutions to these problems are temporary and

<sup>1</sup> The second and third figures of 4.11(a) on page 117 adopted these two strategies respectively.

heuristic. Further research in shape analysis, computational vision and perceptual science will certainly improve the performance of systems like ours.

## Bibliography

- Maneesh Agrawala and Chris Stolte. Rendering effective route maps: improving usability through generalization. In *SIGGRAPH '01: Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, pages 241–249, New York, NY, USA, 2001. ACM.
- T. D. Albright and C. G. Gross. Do inferior temporal cortex neurons encode shape by acting as fourier descriptor filters? In *Proceedings of the International Conference on Fuzzy Logic and Neural Networks*, pages 2, 375–378, 1990.
- Helmut Alt, Kurt Mehlhorn, Hubert Wagener, and Emo Welzl. Congruence, similarity and symmetries of geometric objects. *Discrete Comput. Geom.*, 3(3):237–256, 1988a.
- Helmut Alt, Kurt Mehlhorn, Hubert Wagener, and Emo Welzl. Congruence, similarity and symmetries of geometric objects. *Discrete & Computational Geometry*, 3(3):237–256, 1988b.
- Helmut Alt, Bernd Behrends, and Johannes Blömer. Approximate matching of polygonal shapes (extended abstract). In *SCG '91: Proceedings of the seventh annual symposium on Computational geometry*, pages 186–193, New York, NY, USA, 1991. ACM.
- Ofelia Anguloa, Hye-Seong Leeb, and Michael O'Mahony. Sensory difference tests: Overdispersion and warm-up. *Food Quality and Preference*, 18(2):190–195, March 2007.
- M. A. Armstrong. *Basic Topology*, chapter Euler Characteristics, pages 158–161. Springer-Verlag, July 1983.
- Rudolf Arnheim. *Art and Visual Perception: A Psychology of the Creative Eye*. University of California Press, November 2004a.
- Rudolf Arnheim. *Visual Thinking*. University of California Press, April 2004b.

- M. J. Atallah. On symmetry detection. *IEEE Trans. Comput.*, 34(7):663–666, 1985.
- M.J. Atallah. Checking similarity of planar figures. *International journal of computer & information sciences*, 13:279–290, 1984.
- Jonas August, Steven W. Zucker, and Kaleem Siddiqi. Ligature instabilities in the perceptual organization of shape. *Computer Vision and Pattern Recognition, IEEE Computer Society Conference on*, 2:2042, 1999.
- Elan Barenholtz and Jacob Feldman. Determination of visual figure and ground in dynamically deforming shapes. *Cognition*, 101(3):530 – 544, 2006.
- F.S. Beckman. On isometries of euclidean spaces. In *Proceedings of the American Mathematical Society*, volume 4, pages 810–815, 1953.
- Irving Biederman. Recognition-by-components: a theory of human image understanding. *Psychology Review*, 94(2):115–147, Apr 1987.
- H. Blum. *Models for the perception of Speech and Visual Form*, chapter A transformation for extracting new descriptors of shape, pages 363–380. MIT Press, Cambridge, MA, 1967.
- H. Blum. Biological shape and visual science. *Journal of Theoretical Biology*, 38:205–287, 1973.
- H. Blum and R.N. Nagel. Shape description using weighted symmetric axis features. *Pattern Recog.*, 10:167–180, 1978.
- F.L. Bookstein. The line skeleton. 11(2):123–137, October 1979.
- M. Brady and H. Asada. Smoothed local symetries and their implementation. *International Journal of Robotics Research*, 3(3):36–61, 1984.
- Rodney Allen Brooks. *Symbolic reasoning among 3-d models and 2-d images*. PhD thesis, 1981.
- Christina A. Burbeck and Stephen M. Pizer. Object representation by cores: Identifying and representing primitive spatial regions. *Vision Research*, 35(13):1917–1930, 1995.

- Christina A. Burbeck, Stephen M. Pizer, Bryan S. Morse, Dan Ariely, Gal S. Zauberman, and Jannick P. Rolland. Linking object boundaries at scale: A common mechanism for size and shape judgments. *Vision Research*, 36:361–372, 1996.
- L. Paul Chew. Constrained delaunay triangulations. *Algorithmica*, 4:97–108, 1989.
- Sung Woo Choi and Hans-Peter Seidel. Hyperbolic hausdorff distance for medial axis transform. *Graph. Models*, 63(5):369–384, 2001.
- Ondrej Chum and Jiri Matas. Matching with prosac ” progressive sample consensus. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 220–226, 2005.
- T. Cormen, C. Leiserson, R. Rivest, and C. Stein. *Introduction to Algorithms, 2nd edition*. MIT Press, 2001.
- J. De Winter and J. Wagemans. Segmentation of object outlines into parts: A large-scale integrative study. *Cognition*, 99(3):275–325, April 2006.
- Doug DeCarlo and Anthony Santella. Stylization and abstraction of photographs. In *SIGGRAPH '02: Proceedings of the 29th annual conference on Computer graphics and interactive techniques*, pages 769–776, 2002.
- Mathieu Desbrun, Mark Meyer, Peter Schröder, and Alan H. Barr. Implicit fairing of irregular meshes using diffusion and curvature flow. In *SIGGRAPH '99: Proceedings of the 26th annual conference on Computer graphics and interactive techniques*, pages 317–324, New York, NY, USA, 1999. ACM Press/Addison-Wesley Publishing Co.
- George Dickie. *Art and the aesthetic : an institutional analysis*. Ithaca, N.Y. : Cornell University Press, 1974.
- S. J. Dickinson, A. P. Pentland, and A. Rosenfeld. From volumes to views: An approach to 3-d object recognition. In *Proc. IEEE Workshop Directions in Automat. CAD-Based Vision*, pages 85–96, Maui, HI, 1991.

- D.H. Douglas and T.K. Peucker. Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. *The Canadian Cartographer*, 10(2):112–122, 1973.
- Frédo Durand. An invitation to discuss computer depiction. In *NPAR '02: Proceedings of the 2nd international symposium on Non-photorealistic animation and rendering*, pages 111–124, 2002.
- Klaus Ecker. Regularity theory for mean curvature flow. *Progress in nonlinear differential equations and their applications*, (75), 2004.
- J. H. Elder and R. M. Goldberg. Ecological statistics of gestalt laws for the perceptual organization of contours. *Journal of Vision*, 2(4):324–353, 2002.
- Helga Eng. *The Psychology of Children's Drawings*. Routledge, 2nd edition, July 1999.
- Daniel M. Ennis and Jian Bi. The beta-binomial model: Accounting for inter-trial variation in replicated difference and preference tests. *Journal of Sensory Studies*, 13:389–412, 1998a.
- Daniel M. Ennis and Jian Bi. A thurstonian variant of the beta- binomial model for replicated difference tests. *Journal of Sensory Studies*, 13:461466, 1998b.
- Jacob Feldman. *What is Cognitive Science*, chapter Does vision work? Towards a semantics of perception, pages 208–229. Wiley-Blackwell Publisher Inc., Malden, Massachusetts, USA, 1999.
- Jacob Feldman and Manish Singh. Bayesian estimation of the shape skeleton. *proceedings of the national academy of sciences*, 103(47):18014–18019, 2006.
- James A. Ferwerda. Elements of early vision for computer graphics. *IEEE Computer Graphics and Applications*, 21(5):22–33, 2001.
- Adam Finkelstein and David H. Salesin. Multiresolution curves. In *ACM SIGGRAPH 1994: Proceedings of the 21st annual conference on Computer graphics and interactive techniques*, pages 261–268, 1994.

- Martin A. Fischler and Robert C. Bolles. *Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography*, pages 726–740. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1987.
- Shachar Fleishman, Iddo Drori, and Daniel Cohen-Or. Bilateral mesh denoising. *ACM Transaction on Graphics*, 22(3):950–953, 2003.
- Shachar Fleishman, Daniel Cohen-Or, and Cláudio T. Silva. Robust moving least-squares fitting with sharp features. *ACM Transaction on Graphics*, 24(3):544–552, 2005.
- Thomas Funkhouser, Michael Kazhdan, Philip Shilane, Patrick Min, William Kiefer, Ayellet Tal, Szymon Rusinkiewicz, and David Dobkin. Modeling by example. *ACM Transactions on Graphics*, 23(3):652–663, 2004.
- Ran Gal, Olga Sorkine, Niloy Mitra, and Daniel Cohen-Or. iWIRES: An analyze-and-edit approach to shape manipulation. *ACM Transactions on Graphics (proceedings of ACM SIGGRAPH)*, 28(3), 2009.
- Stephen J. Gange, Marc Saez Alvaro Munoz, and Jordi Alonso. Use of the beta-binomial distribution to model the effect of policy changes on appropriateness of hospital stays. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 45(3):371–382, 1996.
- M. R. Garey and David S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman, 1979.
- Michael Garland and Paul S. Heckbert. Surface simplification using quadric error metrics. In *Proceedings of ACM SIGGRAPH 97*, pages 209–216, August 1997.
- Wilson S. Geisler and Jeffrey S. Perry. Natural systems analysis. *Journal of Vision*, 8(17):30–30, 12 2008.
- Wilson S. Geisler, Jeffrey S. Perry, Super B.J., and D.P. Gallogly. Edge co-occurrence in natural images predicts contour grouping performance. *Vision Research*, 41:711–724, 2001.
- Bogdan Georgescu and Peter Meer. Point matching under large image deformations and illumination changes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(6):674–688, 2004.

- A. Gibbons. *Algorithmic Graph Theory*. Cambridge University Press, 1985.
- P.J. Giblin and S.A. Brassett. Local symmetry of plane curves. *The American Mathematical Monthly*, 92(10):689–707, 1985.
- Bruce Gooch, Erik Reinhard, and Amy Gooch. Human facial illustrations: Creation and psychophysical evaluation. *ACM Transactions on Graphics*, 23(1):27–44, 2004.
- A. Ardeshir Goshtasby. *2-D and 3-D Image Registration: for Medical, Remote Sensing, and Industrial Applications*. Wiley-Interscience, 2005.
- Alfred Gray, Elsa Abbena, and Simon Salamon. *Modern Differential Geometry of Curves and Surfaces with Mathematica*, chapter Logarithmic spirals, pages 23–25. Studies in Advanced Mathematics. Chapman & Hall/CRC, 3rd edition, 2006.
- C. G. Gross, R. Desimone, T. D. Albright, and E. L. Schwartz. Inferior temporal cortex and pattern recognition. *Experimental Brain Research*, 11:179–201, 1985.
- Pratt M. Han J.H. and Regli W.C. Manufacturing feature recognition from solid models: A status report. *IEEE Transaction On Robotics and Automation*, 16(6):782–796, 2000.
- Richard Hartley and Andrew Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, New York, NY, USA, 2003.
- D. D. Hoffman. *Representing shapes for visual recognition*. PhD thesis, Massachusetts Institute of Technology, 1983.
- D. D. Hoffman and W. A. Richards. Parts of recognition. *Cognition*, 18:65–96, 1984.
- D. D. Hoffman and M. Singh. Saliency of visual parts. *Cognition*, 63:29–78, 1997.
- Robert Hooke and T. A. Jeeves. “direct search” solution of numerical and statistical problems. *Journal of the ACM*, 8(2):212–229, 1961.
- Hugues Hoppe. Progressive meshes. In *Proceedings of SIGGRAPH 96*, Computer Graphics Proceedings, pages 99–108, August 1996.

- D. H. Hubel and T. N. Wiesel. Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology*, 195(1):215–243, March 1968.
- David Hubel. *Eye, Brain, and Vision*. Scientific American Library, 1988.
- Eduard Imhof. *Cartographic Relief Presentation*. Walter De Gruyter, 1982.
- Hailin Jin. A three-point minimal solution for panoramic stitching with lens distortion. *Computer Vision and Pattern Recognition, IEEE Computer Society Conference on*, 2008.
- Thouis R. Jones, Frédo Durand, and Mathieu Desbrun. Non-iterative, feature-preserving mesh smoothing. In *ACM SIGGRAPH 2003*, pages 943–949, 2003a.
- Thouis R. Jones, Frédo Durand, and Mathieu Desbrun. Non-iterative, feature-preserving mesh smoothing. In *ACM SIGGRAPH 2003*, pages 943–949, 2003b.
- Evangelos Kalogerakis, Patricio Simari, Derek Nowrouzezahrai, and Karan Singh. Robust statistical estimation of curvature on discretized surfaces. In *SGP '07: Proceedings of the fifth Eurographics symposium on Geometry processing*, pages 13–22. Eurographics Association, 2007.
- Henry Kang, Seungyong Lee, and Charles K. Chui. Coherent line drawing. In *NPAR '07: Proceedings of the 5th international symposium on Non-photorealistic animation and rendering*, pages 43–50, 2007.
- Michael Kazhdan, Thomas Funkhouser, and Szymon Rusinkiewicz. Symmetry descriptors and 3D shape matching. In *Proceedings of the 2004 Eurographics/ACM SIGGRAPH symposium on Geometry processing*, pages 115–123, 2004.
- P. J. Kellman and T. F. Shipley. A theory of visual interpolation in object perception. *Cognitive Psychology*, 23:141–221, 1991.
- Sung-Ho Kim and Jacob Feldman. Globally inconsistent figure/ground relations induced by a negative part. *Journal of Vision*, 9(10):1–13, 9 2009.
- B. B. Kimia. On the role of medial geometry in human vision. *Journal of Physiology-Paris*, 97(2):155–190, March 2003.

- B. B. Kimia, A. Tannenbaum, and S. W. Zucker. Toward a computational theory of shape: an overview. In *Proceedings of the first european conference on Computer vision*, pages 402 – 407, 1990.
- B. B. Kimia, I. Frankel, and A. Popescu. Euler spiral for shape completion. *International Journal of Computer Vision*, 54(1/2):157–180, 2003.
- B.B. Kimia, A.R. Tannenbaum, and S.W. Zucker. Shapes, shocks, and deformations. I: The components of two-dimensional shape and the reaction-diffusion space. *International Journal of Computer Vision*, 15:189–224, 1995.
- J.J. Koenderink and A.J. van Doorn. The shape of smooth objects and the way contours end. *Perception*, 11:129–137, 1982.
- Ilona Kovács and Bela Julesz. Perceptual sensitivity maps within globally defined visual shapes. *Nature*, 370:644–646, 1994.
- Ilona Kovács, Ákos Fehér, and Bela Julesz. Medial-point description of shape: a representation for action coding and its psychophysical correlates. *Vision Research*, 38:2323–2333, 1998.
- A. Kuijper and O.F. Olsen. The structure of shapes scale space aspects of the (pre-) symmetry set. In *the 5th International Conference on Scale-Space and PDE Methods in Computer Vision (Scale Space 2005)*, pages 291–302, Hofgeismar, Germany, 2005.
- Arjan Kuijper, Ole Fogh Olsen, Peter Giblin, and Mads Nielsen. Alternative 2D shape representations using the symmetry set. *Journal of Mathematical Imaging and Vision*, 26(1-2): 127–147, 2006.
- Jan Eric Kyprianidis, Henry Kang, and Jrgen Dllner. Image and video abstraction by anisotropic kuwahara filtering. *Computer Graphics Forum*, 28(7):1955–1963, 2009.
- Longin Jan Latecki and Rolf Lakämper. Convexity rule for shape decomposition based on discrete contour evolution. *Computer Vision and Image Understanding*, 73(3):441–454, 1999.
- Seungkyu Lee and Yanxi Liu. Curved glide-reflection symmetry detection. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, March 2009.

- Tai Sing Lee, David Mumford, Richard Romero, and Victor A.F. Lamme. The role of the primary visual cortex in higher level vision. *Vision Research*, 38:2429–2454, 1998.
- M. Leyton. Symmetry-curvature duality. *Computer vision, graphics and image processing*, 38: 327–341, 1987.
- M. Leyton. *Symmetry, causality, mind*. MIT Press, Cambridge, MA, 1992.
- Michael Leyton. *A Generative Theory of Shape*. Springer, November 2001.
- William E. Lorensen and Harvey E. Cline. Marching cubes: A high resolution 3D surface construction algorithm. *Computer Graphics*, 21(4):163–169, 1987.
- Adrian Ludwin. Let’s get drawing. The first chapter ”Draw what you See”, from the animation learner’s site: <http://www.geocities.com/Hollywood/Studio/9830/drawing/draw.htm>, 1998.
- Diego Macrini, Kaleem Siddiqi, and Sven Dickinson. From skeletons to bone graphs: Medial abstraction for object recognition. In *International Conference on Computer Vision and Pattern Recognition*, Anchorage, Alaska, June 2008.
- Siddharth Manay, Daniel Cremers, Byung woo Hong, Anthony J. Yezzi, and Stefano Soatto. Integral invariants for shape matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28:1602–1618, 2006.
- D. Marr. Analysis of occluding contour. *Proc. of Royal Society*, 197:441–C475, 1977.
- D. Marr and K.H. Nishihara. Representation and recognition of the spatial organization of three dimensional structure. In *Proceedings of the Royal Society of London*, volume B, pages 269–294, 1978.
- Scott McCloud. *Understanding Comics: The Invisible Art*. Harper Paperbacks, 1994.
- James McCrae and Karan Singh. Sketching piecewise clothoid curves. *Computers & Graphics*, June 2009.
- Peter Meer. *Robust Techniques for Computer Vision*, chapter 4. Prentice Hall, 1st edition, July 2004.

- Ravish Mehra, Qingnan Zhou, Jeremy Long, Alla Sheffer, Amy Gooch, and Niloy J. Mitra. Abstraction of man-made shapes. *ACM Trans. Graph.*, 28(5):1–10, 2009.
- Xiaofeng Mi and Doug DeCarlo. Separating parts from 2D shapes using relatability. *IEEE International Conference on Computer Vision*, 2007.
- Xiaofeng Mi, Doug DeCarlo, and Matthew Stone. Abstraction of 2D shapes in terms of parts. In *Proceedings of the 7th International Symposium on Non-Photorealistic Animation and Rendering*, pages 15–24, 2009.
- Niloy J. Mitra, Leonidas J. Guibas, and Mark Pauly. Partial and approximate symmetry detection for 3D geometry. *ACM Transactions on Graphics*, 25(3):560–568, July 2006.
- Niloy J. Mitra, Leonidas J. Guibas, and Mark Pauly. Symmetrization. *ACM Transactions on Graphics*, 26(3):63:1–63:8, July 2007.
- U. Montanari. A method for obtaining skeletons using a quasi-euclidean distance. *Journal of the ACM*, 15(4):600–624, 1968.
- Ugo Montanari. Continuous skeletons from digitized images. *Journal of the ACM*, 16(4):534–549, 1969.
- Donald G. Morrison. A probability model for forced binary choices. *The American Statistician*, 23(1):23–25, Feb. 1978.
- Glen Mullineux and Sebastian T. Robinson. Fairing point sets using curvature. *Computer-Aided Design*, 39(1):27–34, 2007.
- D. Mumford. Elastica and computer vision. In C. L. Bajaj, editor, *Algebraic geometry and its applications*, pages 491–506. Springer-Verlag, New York, 1994.
- David Mumford. Mathematical theories of shape: do they model perception? *Geometric Methods in Computer Vision*, 1570(1):2–10, 1991.
- James Munkres. *Topology*. Prentice Hall, 2nd edition, December 1999.

- R. L. Ogniewicz. Skeleton-space: a multiscale shape description combining region and boundary information. In *Proceedings of Computer Vision and Pattern Recognition*, pages 746–751, 1994.
- R. L. Ogniewicz and O. Kübler. Hierarchic voronoi skeletons. *Pattern recognition*, 28(3): 343–359, 1995.
- Francesco Orabona, Giorgio Metta, and Giulio Sandini. Learning association fields from natural images. In *POCV 2006: Proceedings of the 2006 Conference on Computer Vision and Pattern Recognition Workshop*, 2006.
- Alexandrina Orzan, Adrien Bousseau, Pascal Barla, and Joëlle Thollot. Structure-preserving manipulation of photographs. In *NPAR '07: Proceedings of the 5th international symposium on Non-photorealistic animation and rendering*, pages 103–110, 2007.
- Sylvain Paris, Pierre Kornprobst, Jack Tumblin, and Frédo Durand. A gentle introduction to bilateral filtering and its applications. In *ACM SIGGRAPH 2008 classes*, pages 1–50, 2008.
- Anitha Pasupathy and Charles E. Connor. Population coding of shape in area v4. *Natural Neuroscience*, 5(12):1332–1338, 2002.
- M. Pauly, N. J. Mitra, J. Wallner, H. Pottmann, and L. Guibas. Discovering structural regularity in 3D geometry. *ACM Transactions on Graphics*, 27(3):#43, 1–11, 2008.
- Marcello Pelillo, Kaleem Siddiqi, and Steven W. Zucker. Matching hierarchical structures using association graphs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21:1105–1120, 1998.
- Mary A. Peterson and Bradley S. Gibson. Must figured-ground organization precede object recognition? an assumption in peril. *Psychological Science*, 5(5):253–259, Sep. 1994.
- Joshua Podolak, Philip Shilane, Aleksey Golovinskiy, Szymon Rusinkiewicz, and Thomas Funkhouser. A planar-reflective symmetry transform for 3D shapes. In *SIGGRAPH '06: ACM SIGGRAPH 2006 Papers*, pages 549–559, New York, NY, USA, 2006. ACM.
- L. Prasad and R Rao. A geometric transform for shape feature extraction. In *Proceedings of SPIE*, pages 222–233, 2000.

- Lakshman Prasad. Morphological analysis of shapes. Technical Report CNLS Newsletter, No. 139, LALP-97-010-139, Center for Nonlinear Studies, T-DOT, Theoretical Division, Los Alamos National Laboratory, July 1997.
- Lakshman Prasad. Rectification of the chordal axis transform skeleton and criteria for shape decomposition. *Image and Vision Computing*, 25(10):1557–1571, 2007.
- Xiaofeng Ren, Charless C. Fowlkes, and Jitendra Malik. Figure/ground assignment in natural images. In *Proc. 9th Europ. Conf. Comput. Vision*, volume 2, pages 614–627, 2006.
- W. A. Richards and D. D. Hoffman. Codon constraints on closed 2D shapes. *Computer Vision, Graphics, and Image Processing*, 31:156–177, 1985.
- W. A. Richards, Jan J. Koenderink, and D. D. Hoffman. Inferring three-dimensional shapes from two-dimensional silhouettes. *Journal of Optical Society of America. A*, 4(7):1168–1175, 1987.
- Arthur H. Robinson, Joel L. Morrison, Phillip C. Muehrcke, A. Jon Kimerling, and Stephen C. Guptill. *Elements of Cartography, 6th edition*. Wiley, 1995.
- Hillel Rom and Gerard Médioni. Hierarchical decomposition and axial shape description. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(10):973–981, 1993.
- Joe Rosen. *Symmetry discovered: concepts and applications in nature and science*. Dover Publications Inc., Mineola, NY, USA, 1998.
- Joe Rosen. *Symmetry Rules: How Science and Nature Are Founded on Symmetry*. Springer, Rockville, MD, USA, 2008.
- Azriel Rosenfeld. Axial representations of shape. *Computer Vision, Graphics, and Image Processing*, 33(2):156–173, 1986.
- Paul L. Rosin. Shape partitioning by convexity. *IEEE Trans. Systems, Man, and Cybernetics, part A*, 30:202–210, 1999.

- Szymon Rusinkiewicz, Michael Burns, and Doug DeCarlo. Exaggerated shading for depicting shape and detail. In *SIGGRAPH '06: ACM SIGGRAPH 2006 Papers*, pages 1199–1205, New York, NY, USA, 2006. ACM.
- 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.
- T. Sebastian, P. Klein, and B. Kimia. Recognition of shapes by editing shock graphs. *Computer Vision, IEEE International Conference on*, 1:755, 2001.
- T. Sebastian, P. Klein, and B. Kimia. On aligning curves. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(1):116–124, 2003.
- J. Serra. *Image Analysis and Mathematical Morphology*. Academic Press, London, 1982.
- D. Shaked and A. M. Bruckstein. Pruning medial axes. *Computer Vision and Image Understanding*, 69:156–169, 1998.
- K. Siddiqi and B. Kimia. Parts of visual form: Computational aspects. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(3):239–251, March 1995.
- K. Siddiqi and B.B. Kimia. A shock grammar for recognition. *IEEE CVPR*, page 507, 1996.
- K. Siddiqi, K. J. Tresness, and B. B. Kimia. Parts of visual form: computational aspects. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17:239–251, 1995.
- K. Siddiqi, K. J. Tresness, and B. B. Kimia. Parts of visual form: Psychophysical aspects. *Perception*, 25:399–424, 1996.
- Kaleem Siddiqi, Ali Shokoufandeh, Sven J. Dickinson, and Steven W. Zucker. Shock graphs and shape matching. In *ICCV '98: Proceedings of the Sixth International Conference on Computer Vision*, page 222, Washington, DC, USA, 1998. IEEE Computer Society.
- Kaleem Siddiqi, Benjamin B. Kimia, Allen Tannenbaum, and Steven W. Zucker. On the psychophysics of the shape triangle. *Vision Research*, 41(9):1153–1178, 2001a.

- Kaleem Siddiqi, Benjamin B. Kimia, Allen Tannenbaum, and Steven W. Zucker. On the psychophysics of the shape triangle. *Vision Research*, 41:1153–1178, 2001b.
- S. Siegel and N.J. Castellan. *Nonparametric statistics for the behavioral sciences*. McGraw–Hill, Inc., second edition, 1988.
- M. Singh and J. M. Fulvio. Visual extrapolation of contour geometry. *Proceedings of the National Academy of Sciences, USA*, 102(3):939–944, 2005.
- M. Singh and D. D. Hoffman. Completing visual contours: The relationship between relatability and minimizing inflections. *Perception and Psychophysics*, 61:636–660, 1999.
- M. Singh, G.D. Seyraninan, and D.D. Hoffman. Parsing silhouettes: The short-cut rule. *Perception and Psychophysics*, (61):636–660, 1999.
- D. M. Smith. Maximum likelihood estimation of the parameters of the beta binomial distribution. *Applied Statistics*, 32:196–204, 1983.
- J. P. Snyder. *Map Projections - A Working Manual*. U.S. Geological Survey Professional Paper 1395. U.S. Geological Survey, 1987.
- Harold Speed. *The Practice and Science of Drawing*, chapter Vision, pages 43–46. Dover Publications, 3 edition, June 1972.
- Charles V. Stewart. Robust parameter estimation in computer vision. *SIAM Reviews*, 41:513–537, 1999.
- H. Sundar, D. Silver, N. Gagvani, and S. Dickinson. Skeleton based shape matching and retrieval. In *SMI '03: Proceedings of the Shape Modeling International 2003*, page 130, Washington, DC, USA, 2003. IEEE Computer Society.
- Shigeko Takahashi. Aesthetic properties of pictorial perception. *Psychological Review*, 102(4):671–683, 1995.
- R.E. Tarone. Testing the goodness of fit of the binomial distribution. *Biometrika*, 66:585–590, 1979.

- Joshua Charles Taylor. *Nineteenth-century Theories of Art*. University of California Press, 1987.
- Hseyin Tek and Benjamin B. Kimia. Perceptual organization via the symmetry map and symmetry transforms. In *In Proc. of CVPR, Forth Collins*, pages 471–476, 1999.
- C. Tomasi and R. Manduchi. Bilateral filtering for gray and color images. In *ICCV '98: Proceedings of the Sixth International Conference on Computer Vision*, page 839, Washington, DC, USA, 1998. IEEE Computer Society.
- R.M. Tookey and A.A. Ball. *The mathematics of surfaces VII, Information Geometers*, chapter Estimation of curvatures from point data, pages 131–144. Winchester, 1997.
- K. Tsunoda, Y. Yamane, M. Nishizaki, and M. Tanifuji. Complex objects are represented in macaque inferotemporal cortex by the combination of feature columns. *Nature Neuroscience*, 4:832–838, 2001.
- P.M. van der Poorten and Christopher B. Jones. Characterisation and generalisation of cartographic lines using delaunay triangulation. *International Journal of Geographical Information science*, (8):773–794, 2002.
- T. Várady and R. R. Martin. *Handbook of Computer Aided Geometric Design*, chapter Reverse Engineering, pages 651–681. Springer, 2002.
- Shaun P. Vecera, Edward K. Vogel, and Geoffrey F. Woodman. Lower region: A new cue for figure-ground assignment. *Journal of Experimental Psychology: General*, 131:194–205, 2002.
- Jue Wang, Yingqing Xu, Heung-Yeung Shum, and Michael F. Cohen. Video tooning. In *SIGGRAPH '04: ACM SIGGRAPH 2004 Papers*, pages 574–583, 2004.
- Regli W.C. *Geometric algorithms for recognition of features from solid models*. Phd dissertation, Univ. Maryland, College Park, MD, 1995.
- Ben Weiss. Fast median and bilateral filtering. *ACM Transaction on Graphics*, 25(3):519–526, 2006.

- E. P. Wigner. Violations of symmetry in physics. *Scientific American*, 213(6):28, December 1965.
- John Willats. *Art and Representation: New Principles in the Analysis of Pictures*. Princeton University Press, 1997.
- D. A. Williams. The analysis of binary responses from toxicological experiments involving reproduction and teratogenicity. *Biometrics*, 31(4):949–952, Dec. 1975.
- Holger Winnemöller, Sven C. Olsen, and Bruce Gooch. Real-time video abstraction. *ACM Transactions on Graphics*, 25(3):1221–1226, 2006.
- Dengsheng Zhang and Guojun Lu. Review of shape representation and description techniques. *Pattern Recognition*, 37:1–19, 1 2004.
- Song Chun Zhu and A. L. Yuille. Forms: A flexible object recognition and modeling system. *International Journal of Computer Vision*, 20:187–212, 1995.

## Vita

### Xiaofeng Mi

- 2010** Ph. D. in Computer Science, Rutgers University
- 2010** Certificate in Cognitive Science, Rutgers University
- 2003** M. Sc. in Computer Science from Zhejiang University, China
- 2000** B. Sc. in Computer Science from Zhejiang University, China
- 2003-2009** Teaching and research assistant, Department of Computer Science, Rutgers University