12-2003

Perceptual Sketch Interpretation

Markus Wuersch

PERCEPTUAL SKETCH INTERPRETATION

By

Markus Wuersch

Eidg. Dipl.–Ing. HTL, University of Applied Sciences, Rapperswil, 2001

A THESIS

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Master of Science

(in Spatial Information Science and Engineering)

The Graduate School

The University of Maine

December, 2003

Advisory Committee:

Max J. Egenhofer, Professor of Spatial Information Science and Engineering, Advisor

Michael F. Worboys, Professor of Spatial Information Science and Engineering
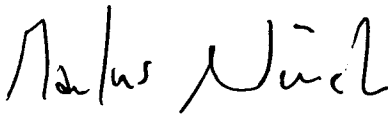
Werner Kuhn, Professor of Geoinformatics, University of Münster, Germany

ii

LIBRARY RIGHTS STATEMENT

In presenting this thesis in partial fulfillment of the requirements for an advanced degree at The University of Maine, I agree that the Library shall make it freely available for inspection. I further agree that permission for "fair use" copying of this thesis for scholarly purposes may be granted by the Librarian. It is understood that any copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Signature:

Date: 12 / 16 / 03

PERCEPTUAL SKETCH INTERPRETATION

By Markus Wuersch

Thesis Advisor: Dr. Max J. Egenhofer

An Abstract of the Thesis Presented
in Partial Fulfillment of the Requirements for the
Degree of Master of Science
(in Spatial Information Science and Engineering)
December, 2003

Sketching is a creative form of describing a spatial scene. People perceive such a scene in a straight forward way and build a mental model of the objects contained in a sketch. Whereas these objects might be regions, a sketch only contains lines and, therefore, developing automated sketch interpretation means outlining a rationale to grouping lines according to the objects they belong to. Automated sketch interpretation allows efficient processing of sketches. Labor intensive manual extraction could be brought to a minimum and, therefore, spatial data in form of sketches and spatial information extracted from sketches would be available more readily. Though spatial data in the form of sketches are less common than collected aerial photographs or satellite images, an automated sketch interpretation could provide valuable findings for researching feature extraction from imagery. This thesis outlines a *perceptual sketch interpretation model* that applies theories from spatial reasoning and gestalt theory. Gestalt theory provides laws of organization, which describe how people organize their visual input. The law of good continuity is incorporated into the sketch interpretation model. It is used to identify regions with a continuous boundary. The model first identifies a set of

regions and then extracts the region with the most continuous boundary. Two alternative sets of regions can be used in the model: (1) the set of all possible regions of a sketch or (2) a subset containing regions with continuous boundaries. The latter is typically significantly smaller than the full set. This thesis outlines a perceptual sketch interpretation model. The assessment of perceptual sketch interpretation model shows that by applying the law of good continuity to identify regions, the correctness of the sketch interpretation is improved. The sketch interpretation process also executed much faster using a subset of regions than when using all possible regions.

ACKNOWLEDGMENTS

Technically, the pages of this thesis document the research work that I have completed here at the University of Maine. On a personal note, there is much more between the book covers; there is an adventure of learning, friendship, maturation, and appreciation. The people that have been part of this adventure deserve my gratitude.

Sincere thanks go to my advisor Dr. Max J. Egenhofer. His advice, guidance, and positive outlook were invaluable.

I would also like to thank my other committee members, Dr. Michael F. Worboys and Dr. Werner Kuhn. Many thanks also to the staff of the SIE department.

Thanks also go to friends here in Orono. In a place far away from home, I am especially grateful for your encouragements and support; for your friendship.

Above all, I owe sincere gratitude to my parents who have supported me at all times. Whereas I learned a great deal here in Maine, the values that you taught me made this adventure a success.

A thank you to the Institutional Review Board for the Protection of Human Subjects, University of Maine, for their review and approval of the survey in this thesis.

A thank you also to the Leica Fond who opened the door to graduate school.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

INTRODUCTION

Every day, large amounts of spatial data are collected. Interpreting these data poses a shear never ending task. Automated feature extraction could reduce manual labor. The objective of this thesis is to investigate feature extraction from sketches. Though spatial data in form of sketches are less common than aerial photographs or satellite images, an automated feature extraction could provide valuable findings for researching feature extraction from imagery. We explore theories from spatial reasoning as well as laws from gestalt theory and outline a perceptual sketch interpretation model. A sketch interpretation prototype is implemented to verify the developed model.

1.1 Feature Extraction from Sketches

The challenge of feature extraction lies in recovering features undamaged and free of breaks, and in successfully grouping them according to the object to which they belong (Blake and Isard 1998). Analogically to the latter part of this definition, this thesis models people's perception to identify regions in a sketch by grouping topologically clean lines that form region boundaries.

Sketching is a direct, creative, and visual form of expression (Blaser 2000a; Goldschmidt 1991). Because sketching is by default spatial, it is well suited for describing spatial scenes (Blaser 2000a; Blaser 2000b). In addition, the way of

1

representing spatial objects is often generalized while little attention is given on details. These characteristics provide a well-suited research base for investigating automated feature extraction from spatial data.

1.1.1 Scope

Only static sketches with regions in R^2 are of concern for this research. Static sketches are line drawings on paper that reveal no information about the drawing sequence of lines. Lines, patches, and regions are valid elements in a sketch and are defined as follows:

- A line is the connection without any self intersection between exactly two points. Such lines are a subset of 1-cells in set theory.

- A patch is an element in a partition of space. The interiors of any two patches are disjoint. Such patches are homeomorphic to 2-discs.

- A region is a perceived object in a sketch. Regions are a single patch or the union of two or more connected patches. Such regions are also homeomorphic to 2-discs.

This thesis excludes such complex cells as regions with holes, separations, or regions with spikes (Figure 1.1). Also excluded are tessellation because they do not allow a distinct interpretation of the drawn objects.

2
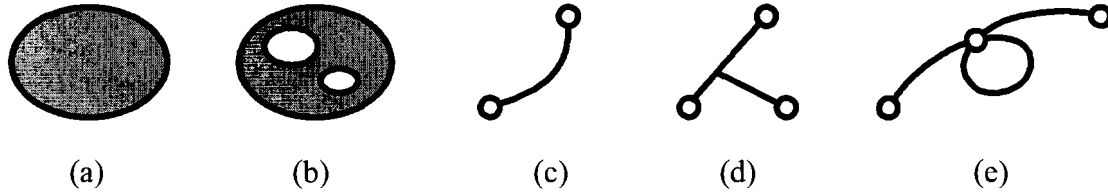
Figure 1.1. (a, c, and e) valid objects in a sketch and (b and d) complex cells and, therefore, invalid objects in a sketch.

1.1.2 Recovering Features

According to Blake and Isard (1998), the first step in feature extraction is to recover any feature undamaged and free of breaks. In the case of sketches, the features to be recovered are lines. Recovering features of a sketch means converting from a raster representation into a simple vector model.

In order to recover these vectors undamaged and free of breaks, it is important that any drawing errors are eliminated. Drawing errors occur in form of overshoots, undershoots, and slivers. If a sketch represents a cognitive map, errors in the sketch are mostly metrical and rarely topological (Lynch 1960). It follows that the topology of a sketch is important in peoples perception whereas metric properties refine: "topology matters, metric refines" (Egenhofer and Mark 1995).

1.1.3 Grouping Features to Objects

Topologically, a sketch contains lines. This view of a sketch is significantly different from the mental model that people build when perceiving a sketch. The regions contained in the people's mental model are represented by a group of lines that form

3

the region's boundary. Knowing the drawing sequence of lines in a sketch could allow to group lines to regions (Blaser 2000a). This thesis, however, outlines a reasoning that is independent of the drawing sequence.

1.1.4 Benefits of Feature Extraction from Sketches

By extracting features from spatial data, information is gained and the data become more meaningful, more available, and, therefore of a higher quality. Manually extracting features from spatial data, however, comes at a high cost in labor. Automated feature extraction reduces the cost of data processing while also gaining additional information.

Though, few data are available in form of sketches, solving feature extraction for sketches will be of a great value for future research. Research in computer vision and feature extraction from medias such as satellite images or aerial photographs could profit from the results of this thesis. In addition, this algorithm can be applied to any data that can be transformed into sketch-like representation. For example, converting a paper sketch into a digital sketch simply requires a vectorization and a clean algorithm. The digital sketch then consists of vectors that form boundaries of regions. Converting an aerial photograph into a vector representation requires one more step before the vectorization, namely that of edge detection. The result is topologically the same as the result from converting a paper sketch into a digital sketch. The latter result, however, is highly dependent on the complexity of the photograph and of the performance of the edge detection and the clean function.

4

Further, the PSI algorithm can also be applied to vector data that have been merged onto the same layer (e.g., "spaghetti data") to recreate layers.

1.1.5 Problem Statement

The perceived regions in a sketch are not contained in the sketch as topological elements. Only lines are present. A group of lines that form a closed loop implicitly form regions (Figure 1.2). People are typically very good at perceiving such sequences of lines as regions. What seems to be such a simple task for people, however, has proven to be a complex task to be formalized so that it can be implemented and carried out by a machine.

(a)   (b)
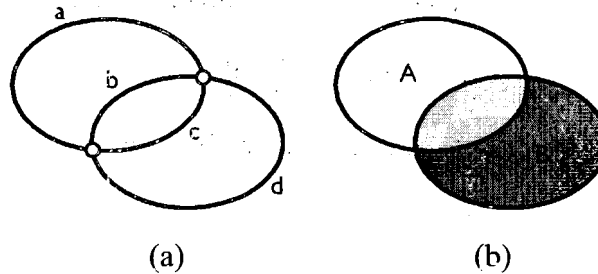
Figure 1.2. Problem statement: (a) a sketch containing lines and (b) the regions that people perceive in the sketch.

1.1.6 Related Work

Identifying objects in images or sketches is a broad research topic involving image processing, computer vision, and machine intelligence. The most related work is briefly reviewed here.

1.1.6.1 Perceptually Closed Path Finding

An algorithm of identifying closed or nearly closed regions is outlined by Saund (2003). This algorithm uses two alternatives to trace along lines in a sketch: (1) *maximal turning* path and (2) *smooth continuation*. The maximal turning path traces the smallest figure possible while the smooth continuation path prefers smooth continuation traces through junctions. The latter makes use of perceptual theory.

For each seed point, a node, a search for figures is performed. This search is applied to a *search tree* that contains all possible paths in the sketch. On this tree, the algorithm traces along branches depending on the maximal turning path or a smooth continuation, and turn direction. The found figures are either accepted or rejected based on a measure for a *good gestalt*.

A drawback of this algorithm is that it requires domain knowledge about the ways in which curve fragments tend to compose into larger closed paths.

1.1.6.2 PerSketch

PerSketch is a perceptually-supported sketch editor Saund and Moran (1995). This interactive sketch editing tool gives the user suggestions when editing an object in a sketch. In doing so, PerSketch tries to read the user's mind. The class of introduced tools is called *WYPIWYG* (What You *Perceive* Is What You Get).

PerSketch defines *prime objects* and *composite objects*. Prime objects are lines in a sketch, whereas composite objects are a set of grouped strokes. Any possible object in the sketch is contained in an object lattice. From this lattice, the algorithm picks objects

based on a set of rules that identify closure, parallelism, corners, and T-junctions. Literature on building these rules can be found in the computer-vision literature, e.g., Mohan and Nevatia (1989) and Sarkar and Boyer (1993).

1.1.6.3 CANC2

CANC2 (Mohan and Nevatia 1992) is a computer vision system that is able to identify object edges from a vectorized image. The set of vectorized edges of an image is reduced to actual object edges by applying gestalt laws (e.g., proximity, continuity, symmetry, closure, and familiarity) thus eliminating noise. Identified edges are grouped to non-overlapping object surfaces. CANC2 is used in a 3D vision system to model people's perception.

1.1.6.4 Sketching Spatial Queries

Sketching Spatial Queries (Blaser 2000a) aims to build a query from a sketch input. The query returns other sketches that are similar to the input sketch. The input sketch has to be drawn in real-time. This way, Sketching Spatial Queries is able to record temporal information about the drawn elements (e.g., drawing sequence). The model then groups the drawn lines into objects based on geometric and temporal attributes. Similarity between sketches is computed based on completeness, geometry, topology, metric, and directions of objects and topological relations.

7

1.2 A Perceptual Sketch Interpretation Model

Grouping lines to form regions is a perceptual task. A formalization of this task is investigated and outlined in a PSI model. The following sections describe the goal, the approach, and the hypothesis of this thesis.

1.2.1 Goal

The goal of this thesis is to formalize the task of interpreting regions that are contained in a sketch. For this purpose, we investigate theories from spatial reasoning as well as the use of the laws of organization from gestalt theory. We want to demonstrate the suitability of the law of continuity as a rationale for identifying regions in a sketch. Finally, we propose a perceptual sketch interpretation model.

1.2.2 Approach

This thesis is concerned with feature extraction from a paper sketch. In order to automatically execute this task in a digital environment, analog-to-digital conversion has to be addressed. This conversion creates a digital sketch from which the PSI algorithm first creates the partitions of space. Partitions subdivide space into patches, which are used as building blocks for any region to be identified. Theories from spatial reasoning as well as laws from gestalt theory are incorporated into the PSI algorithm.
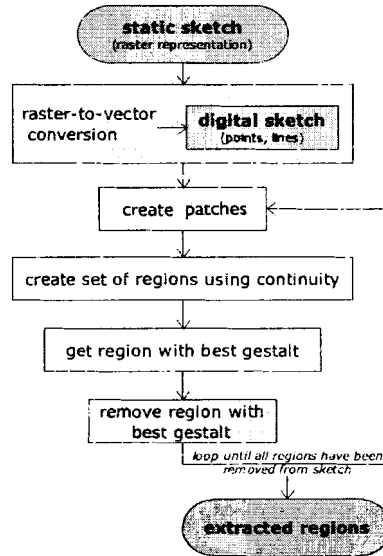
Figure 1.3. Algorithm of extracting regions from a sketch.

1.2.2.1 Digital Sketch

Any static sketch has to be converted into a digital sketch before the PSI algorithm can be applied to it. This process includes three steps: (1) *scanning*, (2) *vectorizing*, and (3) *cleaning*. Scanning can be accomplished with a common scanner that fits the desired paper sketch. Typically, sketches are drawn with the same pen or pencil for the entire sketch and, therefore, no distinction between colors is necessary. The raster representation of such a sketch is a binary image. Vectorization, detects all lines in the sketch. It is assumed that the sketch only contains lines (representing boundaries of regions) rather than filled regions. Vectorization has to return a set of lines that do not cross each other. Instead, at every intersection the lines are split so that they meet at the intersection point. This data model is chosen because crossing lines would limit the possible grouping of lines into regions and could not be relied on. The third step cleans drawing errors such as overshoots, undershoots, and slivers. A fuzzy tolerance is used

to detect slivers and a dangle length is used to detect overshoots and undershoots. The result of these three steps is a digital sketch with a topologically clean vector representation of the original paper sketch. Throughout the rest of this thesis, the term *sketch* is used to describe a digital sketch.

1.2.2.2 Identifying Patches

Geographic information must be embedded in a reference system for *time*, *space*, and *attribute* (Chrisman 2001). Feature extraction from a static sketch, however, can only use information about space since no temporal information about the drawing's sequence is available. Also attribute information, such as line colors is uniform throughout the entire sketch so that no further distinctions can be made. Based on information about space, a tracking algorithm that traces along lines and continues in the same direction when reaching an intersection point (e.g., always turn left, or always turn right) identifies boundaries of patches. Patches, however, do not necessarily coincide with objects that people perceive (e.g., two overlapping regions are interpreted as three patches).

1.2.2.3 Identifying Regions

The patches in a sketch can be used as building blocks for any region. Such a region is built by the union of two or more patches, or a patch is itself a region. In order to identify regions that are to be extracted from the sketch, a set of regions that are candidate to be extracted is identified. This set is identified by a 'set module', which

either returns the full set of possible regions or a subset. The smaller the subset the more efficient the algorithm would be.

To identify meaningful objects other than patches, people use intuitive reasoning. Regions are formed by grouping the correct set of lines in a sketch. Theory describing such reasoning has been developed in gestalt theory, a branch of psychology, and has been published as the *laws of organization* (Wertheimer 1923).

A component of the laws of organization is the law of *good continuation*. This law states that two lines are more likely to be grouped together if one line is perceived as the continuation of the other. The law of good continuation is of great importance and allows for identifying regions in the sketch: two patches A and B are likely to form a new region if a segment of patch B's boundary appears as the continuation of a segment of patch A's boundary.

The notion of a good continuity is incorporated in a set module that identifies a set of regions that are candidates to be extracted. As a result the set of possible regions is generally much smaller than the set of the full set of all possible regions and, therefore, fewer regions have to be processed. This set does not necessarily contain all the regions to be extracted.

From the above described set of regions, the region with the best gestalt is stored and removed from the sketch. With the remaining lines in the sketch (original sketch – removed region's boundary), patches are newly built on which a new iteration of finding regions is performed. This iterative process is repeated until no patches are left in the sketch at which point, all regions are identified.

1.2.3 Hypothesis

The PSI algorithm creates a set of regions that is identified by tracing along lines under consideration of a good continuation. This task is outlined in the *continuity set module*. From this set, the region with the best gestalt is picked and extracted from the sketch. Alternatively, the *full set module* creates a set of all possible regions in the sketch (Figure 1.4). In doing so, it is guaranteed that all the regions that should be extracted are contained in this set. This set, however, is described with a lattice and the number of elements in such a lattice can be very large (Equation 1.1) as described by Birkhoff (1948).

Reducing the set of regions is beneficial for the performance of the algorithm and the hypothesis is:

Using the continuity set module in the Perceptual Sketch Interpretation algorithm yields comparable results of interpreted sketches to using the full set module.

$$H*(n+1) = \sum_{h=0}^{n}\binom{n}{h}H*(h), \ where \ \binom{n}{h} = n!/h!(n-h)! \tag{1.1}$$

H*(n) denotes the number of patches of an aggregate of n elements.
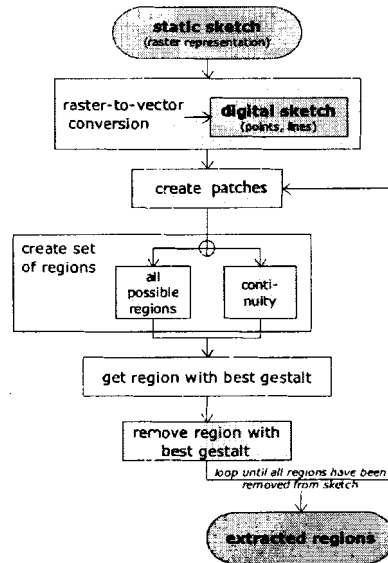
Figure 1.4. The algorithm of extracting regions from a sketch with two set modules.

1.3 Intended Audience

This thesis's intended audience is any researcher interested in computer vision and cognition and any researcher or software developer interested in the design of methods for extracting features from imagery and from raw vector data, as well as methods for advanced vectorization. This thesis may also be of interest to researchers of multimodal human-computer interaction; especially sketch-based interaction and multimodal querying due to its relation to Sketching Spatial Queries (Blaser 2000a). A broader audience including GIS professionals and geographers might also be interested in this thesis, because it is aimed to the development of future, intelligent GISs.

13

1.4 Organization of Remainder of Thesis

This thesis identifies the problem then attempts to provide answers. Existing methods are described as well as their refinement used in this work. The Perceptual Sketch Interpretation algorithm is outlined. A guided tour describes the process of a sketch-to-object conversion, including a sketch interpretation prototype. The PSI model is assessed, conclusions are drawn and future work is suggested.

Chapter two describes two different representations of a spatial scene: a perceptual and a mathematical representation. For the perceptual representation, laws from gestalt theory are analyzed. Spatial reasoning is used to build a mathematical representation that qualitatively describes a spatial scene.

In chapter three, the theories described in chapter two are tailored to the use in this thesis. Refinement of theories of spatial reasoning identify a sketch-specific set of topological relations amongst objects in a sketch. Equivalence classes are formed to simplify the task of forming regions from lines in a sketch. The laws from gestalt theory are formally defined so that they can be implemented in the PSI algorithm.

Chapter four develops the PSI algorithm. Assumptions that the algorithm is based on are stated. A description is given of how the theories form chapter two and three are incorporated into the algorithm. In addition to a description of the algorithm in writing, pseudo code is given for the overall algorithm and for the continuity set module.

Chapter five illustrates a sketch-to-object conversion. Necessary methods from image processing are given and the sketch interpretation prototype is described. The description of the prototype includes project specific classes and a user interface.

Chapter six describes the model evaluation based on a set of sketches that were collected through an online survey. The results of this evaluation are ordered into two aspects: correctness and processing time. Shortcomings of the PSI algorithm are addressed.

Chapter seven concludes this thesis with a review of objectives, methodology, and results of this thesis. It closes with a discussion of possible future research topics as well as extensions and refinements to the current model.

Chapter 2
SKETCH REPRESENTATIONS

People perceive a spatial scene and build a mental model of it according to their individual perception, which is influenced by assumptions, previous knowledge, and experience. The objects contained in the resulting mental model are what this thesis is aiming to identify automatically and to represent this process with a mathematical model. How people perceive and order their visual input has been described by some laws of gestalt theory. The mathematical model used in this thesis is based on set theory. This theory allows a formal description of a spatial scene and the application of formal operators on objects within the mathematical model. Spatial scenes are described qualitatively using theories from spatial reasoning. This chapter describes different representations of a sketch: (1) how it is perceived by people and (2) how it is represented mathematically.

2.1 Perception of a Sketched Scene

People's perception imposes order onto sometimes chaotic visual input (Ballard and Brown 1982) in a way that meaningful objects are identified that build a mental model of the perceived subject. It does so by using intrinsic information that may reliably be extracted from the input, through assumptions, and by applying previous knowledge (Arm 1997; Ballard and Brown 1982). It is a challenge of this work to extract the

intrinsic information and to make the correct assumptions. In doing so, any identified object in a sketch corresponds to an object in people's mental model of the same sketch.

2.1.1 Gestalt Theory

Gestalt theory is a theory of perception that was developed by Max Wertheimer (Wertheimer 1923), Wolfgang Köhler, Kurt Koffka, and Kurt Lewin in Germany during the first part of the twentieth century. Gestalt is the German word for a unified whole, with properties which are more than the sum of its parts. In gestalt theory, people are viewed as open systems in active interaction with their environments. Gestalt theory hypothesizes that there is cognitive processing and that an individual's perception of stimuli has an effect on their response. Gestaltists believe that individuals group stimuli in their own perception depending on several factors which can be considered the laws of gestalt theory (Clark 1999). The basic laws of gestalt theory are the *laws of organization* and the *law of pragnanz* (Wertheimer 1923).

An important notion of gestalt theory is that the larger picture is perceived before its component parts. This is equivalent to "the whole is bigger than the sum of its parts." For example, one sees a series of discontinuous dots upon a homogeneous ground not as a sum of dots, but as a figure (Figure 2.1).
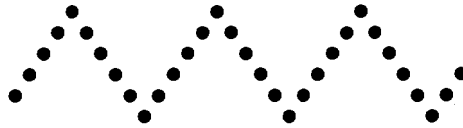
Figure 2.1. A series of dots that are perceived as a figure and not as a sum of dots; after Wertheimer (1923)

2.1.2 Gestalt Laws

How component parts are grouped to form the whole is described in the laws of organization, which are composed of five parts: the laws of *similarity, proximity, closure, symmetry*, and *continuity* (Wertheimer 1923). Co-linearity, co-circularity, parallelism, and symmetry have also been identified for grouping parts to a whole (Koffka 1935; Zhu 1999). The law of continuity (later referred to as the *law of good continuation*) as well as the law of pragnanz are critical for the model development (Chapter 3 and 4).

2.1.2.1 Law of Good Continuation

The law of good continuation states that when two curves intersect, it results in the separation of one from the other. A state of collocation results in conformity with a good, a "curvitally proper" continuation (Petermann 1932). This state is still unequivocally given even when curved lines are used (Wertheimer 1923). For example, a cross composed of four lines that meet at one point is perceived as two intersecting lines, combining the two most similarly oriented, the most continuous lines (Figure 2.2).
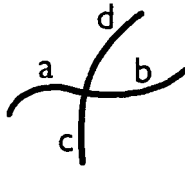
Figure 2.2. Good continuation of two intersecting lines: the line segments a and b are grouped together as well as the segments c and d because they are the most continuous.

2.1.2.2 The Law of Pragnanz – Good Gestalt

"The law of pragnanz implies that if a perceptual field is disorganized when an organism first experiences it, the organism imposes order on the field in a predictable way. This predictable way is in the direction of a good gestalt, a psychological task that does not necessarily involve a change in the physical environment but one which represents a change in how an organism 'sees' its physical environment" (Blosser 1973). In other words, a good gestalt refers to the simplest, most stable figure possible (Zabrodsky and Algom 1994; Zhu 1999).

2.2 Mathematical Model of a Sketched Scene

A mathematical model that formally describes a sketched scene enables automatic processing and analyzing of the scene by a formal system. However, the spatial objects described by the mathematical model often differ from the mental model that people build by perceiving a sketched scene. Essentially the mathematical model has to be able to describe a spatial scene as it corresponds to people's mental model; therefore, it is desired for the mathematical model to capture all the objects in a spatial scene in a way that is closely related to people's thinking. Research has shown that people think

19

in a qualitative manner and that metrical information is secondary – *topology matters, metric refines* (Egenhofer and Mark 1995). Therefore, a qualitative model is chosen for this thesis to closer relate to people's way of thinking.

2.2.1 Spatial Data Model

A spatial data model is a formalization of the spatial concepts that humans employ when they organize and structure their perception of space. A formal model is necessary, because computer systems manipulate symbols according to formal rules (Egenhofer and Herring 1991a;, 1991b; Frank 1992). Here, the formulism also serves as a means to limit the scope of this work to a subset of valid symbols in a sketch.

2.2.1.1 Cells and Cell Complexes

A spatial data model for sketch analysis is similar to the data model that has been used as a base to build upon the definitions of topological relations (Egenhofer and Herring 1991b). It uses algebraic topology, which is based on primitive geometric objects, called *cells*, that are defined for different spatial dimensions: a *0-cell* is a node, a *1-cell* is the link between two distinct 0-cells, and a *2-cell* is the area described by a closed sequence of non-intersecting 1-cells. The topological primitives relevant for this work are the *closure, interior, boundary,* and *exterior* of a cell and are defined by Egenhofer and Herring (1991b) (Equation 2.1-2.4, Figure 2.3).

The closure of an *n-cell A*, denoted by, $\overline{A}$ is the set of all faces r-f of A, where $0 \leq r \leq n$ (Equation 2.1).

$$\overline{A} = \bigcup_{r=0}^{n} r - f \in A \qquad (2.1)$$

The set-theoretic boundary of an *n-cell A*, denoted by ∂A, is the union of all r-faces *r-f*, where *$0 \leq r \leq (n-1)$*, that are contained in *A* (Equation 2.2).

$$\partial A = \bigcup_{r=0}^{n-1} r - f \in A \qquad (2.2)$$

The interior of a cell *A*, denoted by A°, is the set difference between *A 's* closure and *A 's* boundary (Equation 2.3).

$$A^\circ = \overline{A} - \partial A \qquad (2.3)$$

The exterior of a cell *A*, denoted by A^-, is the set of all cells in the universe $\mathcal{U}$ that are not elements of the closure (Equation 2.4).

$$A^- = U - \overline{A} \qquad (2.4)$$

21
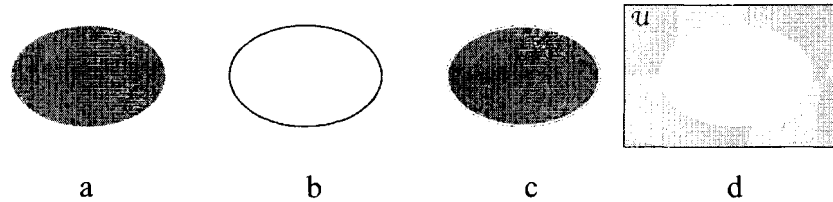
Figure 2.3. Cell primitives: (a) closure, (b) boundary, (c) interior, and (d) exterior.

These primitives can be aggregated to form *cell complexes*. From further definitions by Egenhofer and Herring (1991b) it concludes that (1) interior, boundary, and exterior of a cell or a cell complex are mutually exclusive and (2) their union coincides with the universe.

Any cell is embedded into a universe with its dimension n. The difference between the dimension of the embedding space and the dimension of the cell is defined as the *codimension*.

2.2.1.2 Operations on Cells and Cell Complexes

There is a set of operations that can be carried out on cells or on cell complexes. From this set *union* and *set difference* are important for this thesis, and, therefore are described here.

The *union* of two sets, A and B, forms a third set, $A \cup B$, which contains all the elements of both sets A and B.

The *intersection* of two sets, A and B, forms a third set $C = A \cap B$, which contains only the elements common to both A and B.

2.2.2 Patches

In a sketch that contains regions, the patches of the partition formed by the lines in the sketch are the building blocks for all possible objects in the sketch. Any region either is a patch or can be expressed as the union of several patches. In set theory, a partition of a set A is defined as a collection of subsets of A, such that each element in A belongs to exactly one of these subsets (Preparate and Yeh 1973). Expressed as cells, partitions are subdivisions of space and consist of cells in the most general case, where any two distinct cells do not have a common interior (Egenhofer and Herring 1991a). Partitions may be complete subdivisions of space where the set of patches covers the entire embedding space. Sketches, however, mostly represent incomplete subdivisions of space.

2.2.3 Qualitative Description of a Sketched Scene

Qualitative binary topological relations between two objects A and B have been described by the 4-intersection model (Egenhofer and Herring 1991a) and the 9-intersection model (Egenhofer and Herring 1991b). Objects are described as point sets, where each point set can be a point, line, or area. A point set A has an interior (A°), a boundary (∂A), and an exterior (A^-) as defined in set theory.

In the case of the 9-intersection, a topological relation between two point sets A and B is described by the intersection of A's interior, boundary, and exterior with B's interior, boundary, and exterior. In the case of the 4-intersection, only intersections between A's interior and boundary with B's interior and boundary are used (Equation

23

2.5). Each of the intersections can be either empty ($\varnothing$) or non-empty ($\neg\,\varnothing$) and, therefore, the number of possible intersection matrices are 2^9 for the 9-intersection and 2^4 for the 4-intersection.

$$R(A,B) = \begin{pmatrix} A° \cap B° & A° \cap \partial B \\ \partial A \cap B° & \partial A \cap \partial B \end{pmatrix} \tag{2.5}$$

2.2.3.1 Topological Relations Between Regions in R^2

In the case of two simple regions without holes embedded in R^2 (i.e., the codimension is zero) only a subset of the possible intersection matrices has a meaningful geometric representation (Egenhofer and Franzosa 1991; Egenhofer and Herring 1991a;, 1991b). Both the 4-intersection and the 9-intersection reveal the same set of eight topological relations (Figure 2.4). Since the 4-intersection model requires less computations it is used in this work.

$$\begin{pmatrix} \varnothing & \varnothing \\ \varnothing & \varnothing \end{pmatrix} \quad \begin{pmatrix} \neg\varnothing & \neg\varnothing \\ \varnothing & \varnothing \end{pmatrix} \quad \begin{pmatrix} \neg\varnothing & \varnothing \\ \neg\varnothing & \varnothing \end{pmatrix} \quad \begin{pmatrix} \neg\varnothing & \varnothing \\ \varnothing & \neg\varnothing \end{pmatrix}$$

disjoint   *contains*   *inside*   *equal*

$$\begin{pmatrix} \varnothing & \varnothing \\ \varnothing & \neg\varnothing \end{pmatrix} \quad \begin{pmatrix} \neg\varnothing & \neg\varnothing \\ \varnothing & \neg\varnothing \end{pmatrix} \quad \begin{pmatrix} \neg\varnothing & \varnothing \\ \neg\varnothing & \neg\varnothing \end{pmatrix} \quad \begin{pmatrix} \neg\varnothing & \neg\varnothing \\ \neg\varnothing & \neg\varnothing \end{pmatrix}$$

meet   *covers*   *coveredBy*   *overlap*
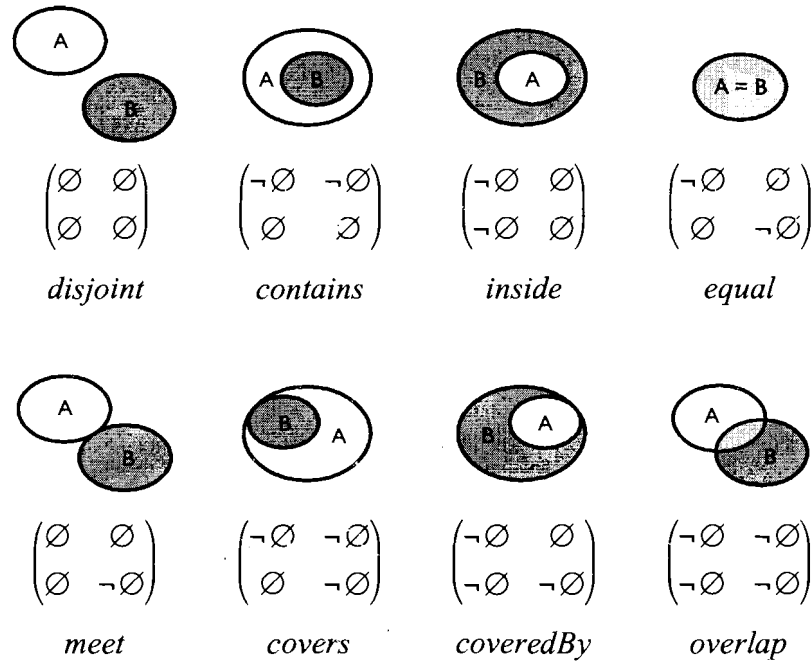
Figure 2.4. The eight topological relations between two regions in R^2.

2.2.3.2 Refinement of Relations Between Two Regions

Relations between two regions are further distinguished by the dimension of their intersections. Two regions in a sketch have codimension 0 and, therefore, only the non-empty $\partial \cap \partial$ distinguishes different dimensions (Egenhofer 1993) (Equation 2.6).

In the case of a sketch, the embedding space is two-dimensional and two relations are distinguished when $\partial \cap \partial$ is non-empty: (1) where the common boundary element is 0-dimensional, the relation is called a *0-relation* and (2) where the two regions share a 1-dimensional boundary element, a segment of each region's boundary, the relation is called a *1-relation*. A non-empty value of $\partial \cap \partial$ in the intersection matrix is, therefore, replaced with the highest dimension of the intersection (Figure 2.5–6).

25

$$Dim(\partial A_n \cap \partial B_n) = 0...(n - 1) \qquad\qquad (2.6)$$

$$0\text{-}meet(A,B) = \begin{pmatrix} \varnothing & \varnothing \\ \varnothing & 0 \end{pmatrix} \qquad 1\text{-}meet(A,B) = \begin{pmatrix} \varnothing & \varnothing \\ \varnothing & 1 \end{pmatrix}$$
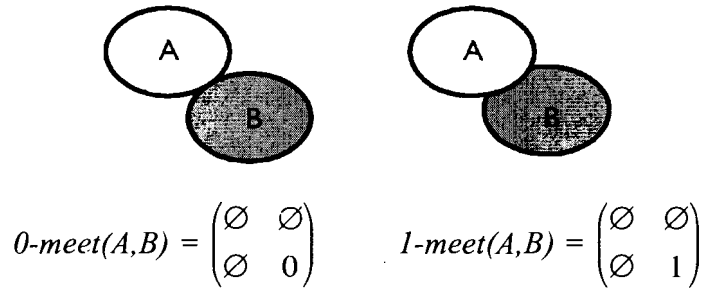
Figure 2.5. Topological relation 0-meet (left) and 1-meet (right).

$$0\text{-}covers(A,B) = \begin{pmatrix} \neg\varnothing & \neg\varnothing \\ \varnothing & 0 \end{pmatrix} \qquad 1\text{-}covers(A,B) = \begin{pmatrix} \neg\varnothing & \neg\varnothing \\ \varnothing & 1 \end{pmatrix}$$

$$0\text{-}coveredBy(B,A) = \begin{pmatrix} \neg\varnothing & \varnothing \\ \neg\varnothing & 0 \end{pmatrix} \qquad 1\text{-}coveredBy(B,A) = \begin{pmatrix} \neg\varnothing & \varnothing \\ \neg\varnothing & 1 \end{pmatrix}$$
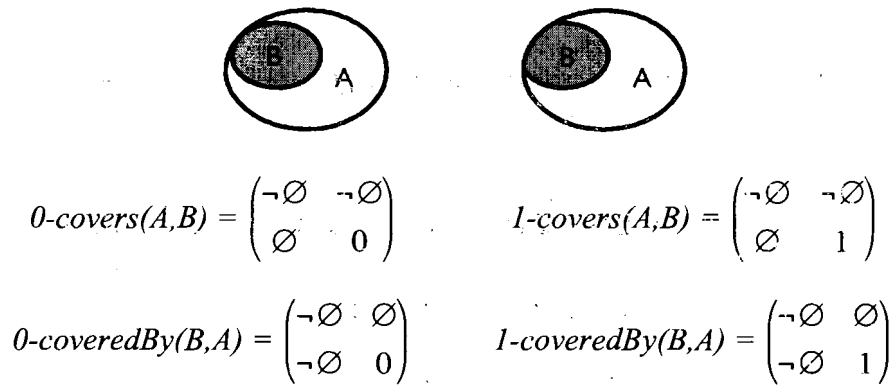
Figure 2.6. Topological relations 0-covers/0-coveredBy (left) and 1-covers/1-coveredBy (right).

$$0\text{-}overlaps(A,B) = \begin{pmatrix} \neg\varnothing & \neg\varnothing \\ \neg\varnothing & 0 \end{pmatrix} \qquad 1\text{-}overlaps(A,B) = \begin{pmatrix} \neg\varnothing & \neg\varnothing \\ \neg\varnothing & 1 \end{pmatrix}$$
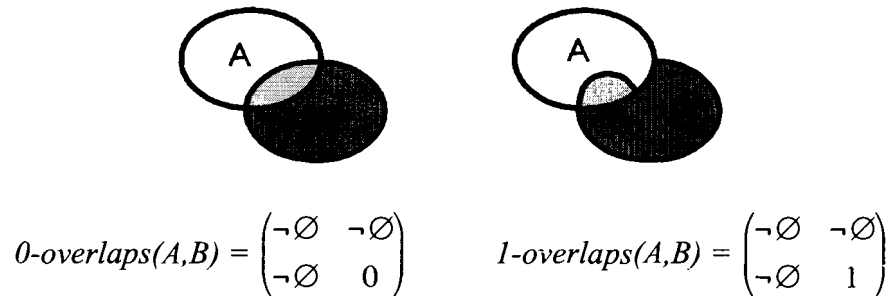
Figure 2.7. Topological relations 0-overlap (left) and 1-overlap (right).

2.2.3.3 Relation Matrix for a Sketched Scene of Regions in R^2

For a sketch with n regions an $n \times n$ matrix can be created that contains the binary topological relations between regions in the sketch. This matrix qualitatively describes a sketched scene (Table 2.1). A full relation matrix contains the following consistencies: (1) along the diagonal of the matrix, the relation between each object and itself must be *equal* (*node consistency*), and (2) the relation between A and B must be equal to the converse relation between B and A (*arc consistency*) (Egenhofer and Sharma 1993; Mackworth 1977). It concludes that only one part of the relation matrix (upper or lower), not including the diagonal elements, need to be calculated to fully describe a spatial scene. The number of necessary relations to be calculated is $(n^2-n)/2$, where n is the number of regions (Egenhofer and Sharma 1992).
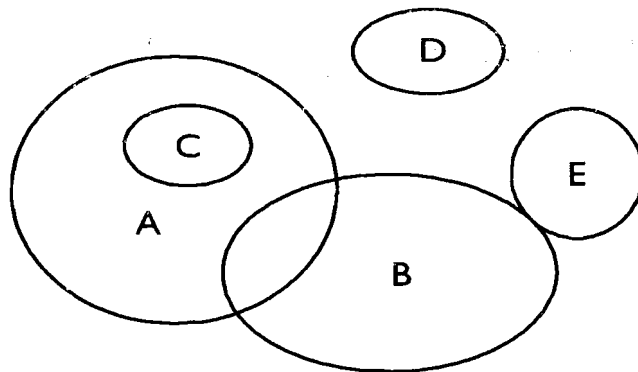
Figure 2.8. An example of a spatial scene with regions A-E.

	A	B	C	D	E
A	equal	0-overlap	contains	disjoint	disjoint
B	0-overlap	equal	disjoint	disjoint	0-meet
C	containedBy	disjoint	equal	disjoint	disjoint
D	disjoint	disjoint	disjoint	equal	disjoint
E	disjoint	0-meet	disjoint	disjoint	equal

Table 2.1. The relation matrix describing the topological relations in Figure 2.8.

2.2.3.4 Topological Relations Between Lines in R^2

Lines in R^2 have codimension 1. Because of this higher degree of freedom, more topological relations can be distinguished between two lines. In order to do so, the 9-intersection model has to be used. because it also considers intersections between the exterior of a line with the interior, boundary, and the exterior of a second line. For two simple lines embedded in R^2, the 9-intersection shows 33 distinct topological relations (Egenhofer *et al.* 1994).

2.3 Summary

Two representations of a sketched scene are described in this chapter: (1) how people perceive a sketch and (2) a mathematical representation.

For the perceptual representation, it was shown that when people perceive a sketched scene they order their visual input according to assumptions, previous knowledge and experience, but also according to the laws of organization and the law of pragnanz. In doing so, people build their own mental model of a perceived sketched

28

scene; therefore, this chapter investigated in the use of the law of good continuation to identify meaningful objects of the sketched scene.

A mathematical model describes the same sketched scene in a formal way but allowing automated analysis of a sketched scene. A qualitative model was chosen because it is closely related to people's way of thinking about geographic objects. Set theory was used to describe objects in a sketch as cells and cell complexes and spatial reasoning was used to describe the topological relations between two objects. Finally, a relation matrix containing any topological relation between any two objects was given as a qualitative, topologically distinct description of a sketched scene.

Chapter 3

REFINED SKETCH REPRESENTATIONS

In this chapter, the theories from Chapter 2 are refined or formalized so that they can be used in this work. The spatial data model is reduced to a subset of possible cell complexes in a sketch, which in turn results in a subset of topological relations that are possible in a sketch. In addition, the possible topological relations between two patches and the relations between lines in a sketch are described. Equivalence classes are defined to reduce the computational effort when grouping lines to form a patch's or a region's boundary. In order to use the law of good continuity from gestalt theory and in order to being able to describe a good gestalt, these notions are formalized.

3.1 Cell Complexes in a Sketch

From the spatial data model described in chapter 2, lines and regions are defined as cell complexes that are "homogeneously n-dimensional" and not partitioned into non-empty, disjoint parts (Egenhofer and Herring 1991b):

- A line is a sequence of connected 1-complexes in R^2, such that they neither cross each other nor form closed loops and have exactly two disconnected boundaries.

- A region is a 2-complex in R^2 with a non-empty, connected interior, a connected exterior, and a connected boundary.

30

Any line or region that is excluded from these definitions (e.g., a region with a disconnected boundary or a line with more than two disconnected boundaries) is referred to respectively, as a complex line or complex region (Figure 3.1).

(a)　　　(b)　　　(c)　　　(d)
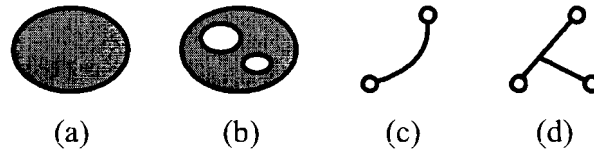
Figure 3.1. Examples of simple cells (a, c) and complex cells (a, b); after (Egenhofer and Herring 1991b).

3.2 Topological Relations Between Objects in a Sketch

The topological relations and their refinement described in chapter 2 apply to 2-discs in R^2. Because patches are a subset of 2-discs with distinct attributes, some topological relations between patches are impossible. In addition, because of the chosen representation of lines, only a small set of the 33 topological relations between two lines are possible in a sketch. For further processing of a sketch, lines are classified by intersection types. These types describe how many lines meet each end of a line.

3.2.1 Topological Relations Between Patches in a Sketch

This work does not deal with true subdivisions of space. Holes are treated as a separate region, inside or contained by another region. This limited set of relations is sufficient, because without semantic information it is impossible to decide if a patch inside another patch is representing a hole or if it is in fact inside a patch. For example, the

31

same representation of two regions is topologically different depending on semantic knowledge about the regions. In Figure 3.2 a lake with an island is drawn and, therefore, the inner region represents a hole in the outer region. On the other hand, Figure 3.2 shows a building on a land parcel. In this case, the inner region is inside the outer region, because the land parcel's area is continuous even where the building stands.
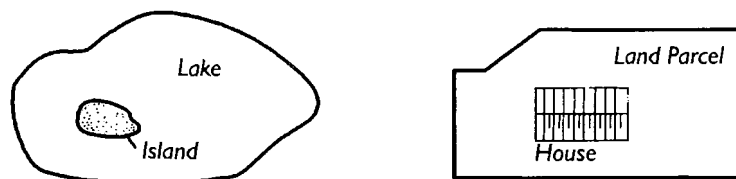
Figure 3.2. Using semantic information: Semantic information reveals that the inner region on the left is a hole in the other region (island in a lake), whereas the inner region on the right is contained in the other region (house on land parcel).

When strictly dealing with partitions of space, only the relations *disjoint, meet (0-meet, 1-meet)*, and *equal* are possible because patches of a partition do not overlap. Since holes are treated as separate regions inside another patch, the topological relations *inside/contains* are possible as well. Relations *1-covers/1-coveredBy* are *1-meet* relations amongst patches of a partiton. Relations *0-covers/0-coveredBy* are valid relations between elements in a partition, because if treated as *meet* relations, the outer regions would be a complex cell. *Overlap* relations are always *1-meet* relations between patches (Table 3.1).

Relation	2-Discs	Partition	Patches in Sketch
disjoint	√	√	√
0-meet	√	√	√
1-meet	√	√	√
0-overlap	√	1-meet	1-meet
1-overlap	√	1-meet	1-meet
0-covers/0-coveredBy	√	1-meet [1]	0-covers/0-coveredBy
1-covers/1-coveredBy	√	1-meet	1-meet
inside/contains	√	1-meet [2]	inside/contains
equal	√	equal [3]	equal [3]

Table 3.1. The relations between two two-discs, partitions, and patches in a sketch: ([1] results in a complex region, [2] results in region with disjoint boundary, [3] only to itself).

3.2.2 Topological Relations Between Lines in a Sketch

Because no line crosses another line, the intersections with the interior of a line $(\partial \cap °, ° \cap \partial, ° \cap °)$ are always empty. As a result, from the 33 topological relations between two lines in R^2, only three relations are possible: *disjoint, meet once,* and *meet twice* (Figure 3.3).
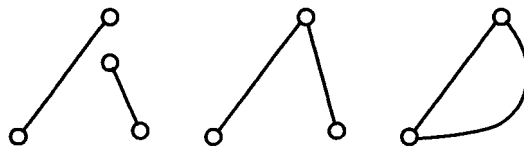
Figure 3.3. Three possible relations with metric refinement between two lines (disjoint, meet once, meet twice).

3.2.3 Intersection Types of Meet Relations Between Lines

The meet relations are further distinguished by the number of line ends that meet at one boundary point. The different relations are called intersection types and are abbreviated with m1, m2, m3, etc. (Figure 3.4). Because the number of strokes at one point is relative to each point and not to a line, the intersection type has to be recorder for each end point of a line.

(a)　　(b)　　(c)　　(d)　　(e)
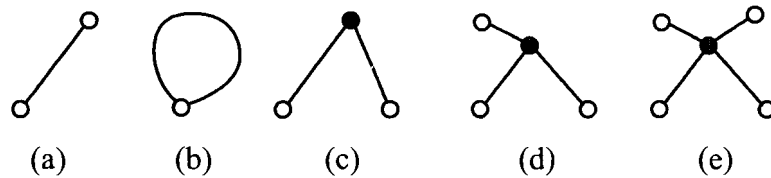
Figure 3.4. Intersection types of lines: with metric information about the number of lines meeting at one point: (a) m1, (b, c) m2, (d) m3, (e) m4.

3.3 Equivalence Classes

A line or a region can only be constructed in one way without changing its topology (e.g., a line *AB* is topologically different from a line *BA*). Different definitions of lines, however, can represent the same graphical object even though the two objects are not topologically equal. Such definitions are described as equivalence classes and simplify the task of composing a region's boundary from a set of lines.

34

3.3.1 Equivalence Classes for Lines

This section describes two equivalence classes. From these classes follows that a polyline composed of line segments can have an arbitrary order of line segments and each line segments direction is arbitrary as well.

3.3.1.1 Order of Boundary Points

A line segment with the boundary points A and B, in that order, is equal to a line with the boundary points B and A ($AB \neq BA$). The graphical representations of these two lines are the same as well. It follows that the order of a line segment's boundary is irrelevant for the purpose of this work.

3.3.1.2 Order of Line Segments that Form a Polyline

A polyline with the vertices A, B, and C, in that order is not equal to a polyline with the boundary points C, B, and A ($ABC \neq CBA$). If the polyline is broken into line segments, however, it follows that the order of boundary points does not matter. In any case, the graphical representation is the same.

3.3.2 Equivalence Classes for Regions

A patch is composed of one closed polyline or several polylines that form a closed boundaries. If the patch is composed of several polylines, then the equivalence classes for lines apply to patches's boundary. The equivalence classes for lines give great freedom in creating a patch's boundary by grouping together several lines. In fact, without the equivalence classes, the number of possible combinations of lines of a

patch's boundary is $n!*2^n$ where n is the number of lines. Not having to deal with all these possibilities simplifies the task of creating patches. Lines do not have to be ordered and the directions of the lines of a boundary do not have to match amongst each other.

3.4 A Definition of Continuity

The law of continuity is of great importance for grouping lines in a sketch to form a region's boundary. Gestalt psychology only provides a descriptive theory but no specific computational process (Zhu 1999). This section formally defines continuity in order to use it in this work.

3.4.1 Continuity Between Two Lines

In this work, a simple definition of continuity is used (Figure 3.5). Continuity is expressed by the angle γ formed by two meeting lines (a and b). This angle is then compared to a threshold. A disadvantage of this continuity measurement is that it results in a local continuity angle. This angle, between the last line segments of two boundaries, can be significantly different from the overall perceived direction of a boundary line (Figure 3.6). In this research stage, however, a local continuity angle is sufficient for developing a formal model of extracting regions.
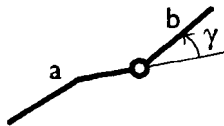
Figure 3.5. The continuity angle γ from line a to line b.

36

local direction
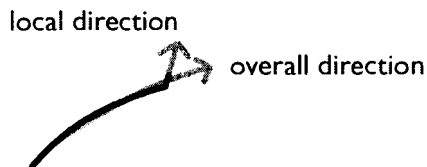
overall direction

Figure 3.6. The overall direction is perceived differently than the local direction (direction of the last line segment).

3.4.2 Continuity Between more than Two Lines

In a case where more than two lines meet at one point, the continuity is first found from the line of interest (line *a*) to any of the other lines. The line (*b*) with the best continuity angle γ is the continuation to line *a*. In this case, we see that continuity is symmetric (Equation 3.1) and, therefore, it has to be examined in both directions, from *a*, with *b* as the continuing line and from *b*, with *a* as the continuing line. It is possible that the best continuity is different when checked from both ways (Figure 3.7). Such a one-way continuity is invalid as it contains uncertainty.

$$\forall \ a, b: \ a \ continuity \ b \Rightarrow b \ continuity \ a \qquad (3.1)$$
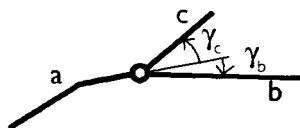
Figure 3.7. Symmetry of continuity: continuity angle γ_b from a to b and γ_c from a to c. The best continuity from a to b is symmetric. The continuous line to a, however, is b and, therefore, there is no continuity between a and c.

37

3.5 A Definition of a Good Gestalt

The notion of a *good gestalt* was first mentioned by gestaltists (Koffka 1935; Wertheimer 1923), but has not been described in great detail. In order to use this notion in any formal model, a formalization is necessary. Properties that contribute to the description of a good gestalt are continuity, regularity, and symmetry. This thesis uses a simple gestalt measure based on continuity described in this chapter.

For a qualitative gestalt value, each absolute continuity angle is compared with a continuity threshold. If the angle is lower than the threshold, it contributes to the overall gestalt value with a plus, otherwise with a minus. The sum of all pluses and minuses describes the gestalt value of a region. This gestalt value is an approach to describe a gestalt in 'good' (plus) or 'poor' (minus) and, therefore, it is more closely related to people's thinking than a possible quantitative description.

3.6 Summary

This chapter refined the methods derived in chapter 2 so that they can be applied to a sketch and its spatial data model. Cell complexes in sketches were discussed as well as topological relations amongst patches and amongst lines. Equivalence classes were defined in order to simplify the task of composing a region's boundary from lines in a sketch. In order to implement the law of good continuity and the law of pragnanz from gestalt theory, a formal definition of these laws was introduced.

Chapter 4

PERCEPTUAL SKETCH INTERPRETATION ALGORITHM

The Perceptual Sketch Interpretation (PSI) algorithm of extracting regions from sketches uses the theories from spatial reasoning and from gestalt theory. Laws from gestalt theory are used to make assumptions that people otherwise intuitively apply based on previous knowledge and experience. In doing so, correct groups of lines that form regions that correspond to people's mental model are found.

In the further writing of this thesis, the term *region* refers to an identified object in a sketch, usually the union of two or more patches.

4.1 Scope

The success of feature extraction methods depends to a large extent on the scope of the geometric objects that may be handled (Bennamoun and Mamic 2002). The scope of the PSI algorithm is limited to regions in a sketch.

Even within a sketch that only contains regions, one cannot always identify the correct set of regions with certainty. This is the case when a space is partitioned with a highly patterned texture, referred to as tessellations. Tessellations are either regular (where only one kind of regular polygon is used), semi-regular (where two kinds of

regular polygons are used) or irregular (where a variety of regular or irregular polygons are used). A typical example of a regular tessellation is a chessboard. Many possible combinations of squares and rectangles can lead to the same graphical representation of a chessboard, especially when the different colors are disregarded (Figure 4.1). In these cases, additional knowledge is necessary to identify the correct set of regions; therefore, resolving puzzles with tessellations is beyond the scope of this thesis.

(a) (b) (c) (d)
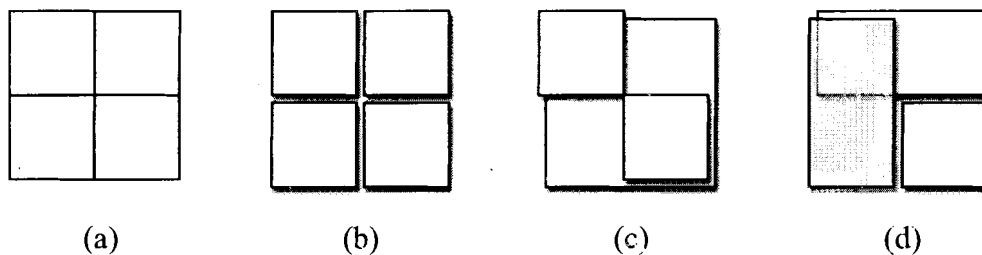
Figure 4.1. A tessellation and different interpretations: (a) a tessellation composed of four squares and (a–c) three different possible interpretations that all lead to the same graphical representation as in (a).

4.2 Algorithm Design

The PSI algorithm finds regions with a good continuing boundary by checking a line in the sketch for other adjacent lines that form a good continuity. Repeating this step systematically for every line in a sketch and in both directions of each line, results in a set of regions; in most cases this result is a subset of all the possible regions in a sketch. The process of identifying this set is outlined in the *continuity set module*. Alternatively, a set of regions ca be created by the *full set module*. This set is described by a lattice and contains all possible regions in a sketch. From the set of regions, the

region with the best gestalt is stored and removed from the sketch, leaving all the other elements in the sketch. Determining the region with the best gestalt is carried out by the *gestalt module*. The process is repeated until no patches are left.

This algorithm makes three assumptions derived from gestalt theory, which are vital for the result returned by the algorithm.

Assumption 1:   *Good continuity is a major factor used in people's perception to organize visual input into meaningful objects.*

Assumption 2:   *By using the notion of good continuity to identify regions in a sketch (continuity set module), the set contains at least one region that corresponds to people's mental model of the same sketch.*

Assumption 3:   *From the set of identified regions, the region with the best gestalt corresponds to people's mental model of the same sketch.*

This algorithm is formally described in the following pseudo code and the subparagraphs describe parts of the pseudo code in more detail.

Perceptual Sketch Interpretation Algorithm:

newSketch = empty sketch

newRegions = list of regions

continuity threshold = minimum threshold

loop

    remove patches that are not 1-meet to any other patch in sketch and add them to newSketch (see 4.2.1)

    find set of all possible regions in sketch (see 4.2.2)

    or find set of possible regions using continuity (see 4.2.3):

```
for each line in sketch

    for each patch containing line

        find region using continuity at start of line and add it to newRegions

        find region using continuity at end of line and add it to newRegions

    end for

    end for

end or

if no region was found (newRegions is empty): increment continuity threshold

else

    remove region of newRegions with best gestalt and add it to newSketch (see 4.2.4)

    build patches with remaining lines in sketch (see 4.2.5)

end if

loop until no patches left in sketch or until continuity threshold > maximum threshold
(see 4.2.6)

if there are patches left in sketch: add patches to newSketch (see 4.2.7)

end if

add unused patches to newSketch (see 4.2.8)
```

4.2.1   Patches with Topological Relation other than 1-Meet

In chapter 3 the spatial data model was described. From the definitions of a line not forming a closed loop and a region with a connected boundary it can be inferred that if a closed loop is detected in a sketch, it always represents a region's boundary. A line that forms a closed loop represents the boundary of a patch that has a topological relation other than 1-meet to any other patch in the sketch. Any such patch is itself a region and is stored and removed from the sketch before any further processing.

4.2.2 Finding All Possible Regions (Full Set Module)

The set of all possible regions in a sketch is described by a lattice (Equation 1.1). The elements of the lattice are found by combining patches that have a topological relation of 1-meet and that, when combined, do not violate the constraints of the specified data model. The algorithm adds all patches to the set of all possible regions and then systematically checks each pair of patches. If they form a valid region, then that region is added to the set. Each region is then again checked with each patch whether or not their union forms yet another region

This is a computationally expensive task because of the generally high number of possible combinations of patches in a full lattice.

4.2.3 Finding Regions using Continuity (Continuity Set Module)

In this section, the continuity set module identifies a set of regions by applying the notion of a good continuation rather than creating a set of all possible regions in a sketch. Good continuity serves as a means of limiting the set of regions to a smaller number than the set of all possible regions. In this thesis, the set module is in relation to the model of describing a good gestalt of a region (gestalt module) and it is able to identify a set of regions without having to create the full lattice. The latter characteristic is expected to significantly improve the efficiency of the algorithm.

Regions are identified by finding two lines that form a good continuation, starting at any segment of any patch's boundary, here called the *starting line*. The two patches

containing the two lines that form the good continuation are combined to build a new region.

Continuity Set Module:

for stroke (thisLine) and region (thisRegion) containing stroke

loop

    find continuing line (otherLine) so that

- otherLine is contained in region (otherRegion) that is 1-meet to thisRegion (see 4.2.3.1)

- thisLine and otherLine are contained in thisRegion∪otherRegion (see 4.2.3.1)

    if continuing line was found:

        thisRegion = thisRegion∪otherRegion (see 4.2.3.1)

        thisLine = otherLine

    else if intersection is > m3: **end** (see 4.2.3.2)

        else

            thisLine = next line in thisRegion

        or end

        end if

    end if

loop until a closed boundary is found

if thisRegion is not complex (see 4.2.3.1): **return** thisRegion

end if

4.2.3.1   Conditions for Continuing Lines and New Regions

The continuing line and the starting line have to form a good continuation as described in chapter 3. The continuing line has to satisfy the following conditions: first, the patches that contain the continuing line (patch *A*) and the starting line (patch *B*) will

44

later be combined to form a new region that contains both of these lines in its boundary. Because continuity can only be determined if the two lines meet and because regions are defined as cell complexes with a connected interior, it follows that only unions of patches that have a 1-meet relation are valid regions (Figure 4.2). Accordingly, the topological relation between A and B has to be 1-meet. Second, $A \cup B$ has to contain both, the starting line and the continuing line, in its boundary; therefore, both lines cannot be the intersecting boundary segment that determines the 1-meet relation between the two unioned patches. Third, the resulting new region must be simple as described in the spatial data model.

(a)   (b)   (c)
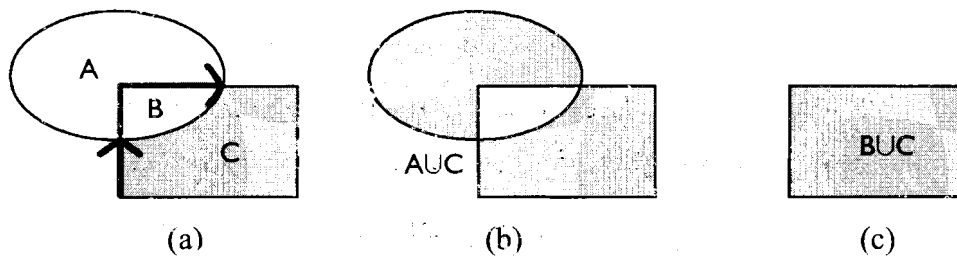
Figure 4.2. Combining patches with continuous boundary segments: (a) a continuous line was found to a boundary segment of patch C; (b) patch A and C have a 0-meet topological relation and A∪C is a complex region; (c) patch B and C have a 1-meet topological relation and B∪C is a simple region.

4.2.3.2 Continuity at M3-Intersections

When a continuation to a line at an m3-intersection is not found, one could argue that the next line in the current patch should serve as the continuing line (Figure 4.3). A reason for doing so is that m3-intersections are points where patches or regions meet. In this case, it is well possible that a region's boundary should continue along the

current patch's boundary even if the continuity angle is higher than the continuity threshold. This reasoning results in an alternative to stopping at m3-intersections if no continuing line was found: to continue with the next boundary segment of the current patch. For the purpose of evaluating this model (Chapter 6), however, the first option is chosen.
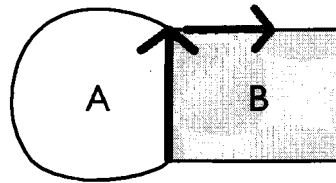
Figure 4.3. Two options for continuity at m3-intersections. If no continuing line is found, the next line of the current patch's boundary can be chosen as the continuing line.

4.2.4 Removing a Region from a Sketch

Removing only the region with the best gestalt value in each iteration of the algorithm can lead to a more accurate identification of any region left in the sketch. Because the patches are newly built after each time a region is removed from the sketch, the number of patches left in the sketch decreases. With a smaller number of patches, the possible combinations of patches to form new regions are fewer as well and the accuracy of the region extraction increases.

Removing an identified region from the sketch is crucial for a successful interpretation of all the regions in a sketch. Because a region is represented by its boundary, removing a region from a sketch is done by removing its boundary.

4.2.4.1 Line Types

In some cases, only a part of a region's boundary can be removed from the sketch because some of its boundary segments are still used to build other patches (Figure 4.4). In order to outline a rationale on deciding which boundary segments can safely be removed, the segments are classified into *line types*. Each line type is described by the number of patches that the segment is part of and by the intersection type of each end of the segment (Table 4.1).
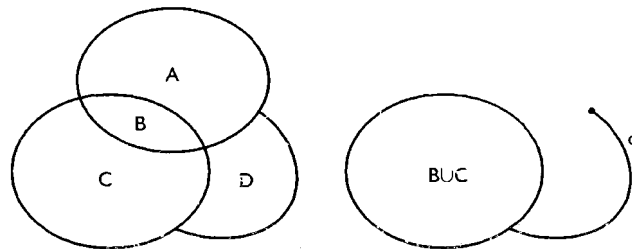
Figure 4.4. Removing a region from a sketch: after removing the region A∪B an open line d is left in the sketch.

	A	B	C	D	E	F	G
Intersection Type at End 1	2	3	3	3	3	4+	4+
Intersection Type at End 2	2	3	3	4+	4+	4+	4+
Number of patches	1	1	2	1	2	1	2

Table 4.1. Classification of line types: the different line types classified by the intersection types and by the number of patches the stroke is contained in.

4.2.4.2 Removing Lines

Some of the line types require additional information to decide whether or not to remove a line. These cases can be checked for by finding two examples of a sketch:

one where the specific line can be removed and one where the line cannot be removed. Finding two such sketches proves that additional information is necessary. For the line types that this proof cannot be found, a rationale is found whether or not to remove the line.

It turns out that for any line types except for type A and type C two representations can be found, one where a line can be removed and one where a line cannot be removed.

Type A: this segment-type should not appear in the sketch at this point of the extraction algorithm. Any line with meet relations m2 at each end can only be part of one patch. These patches are not 1-meet to any other patch (Figure 4.5) and were removed from the sketch at a previous point in the algorithm (Section 4.2.1).
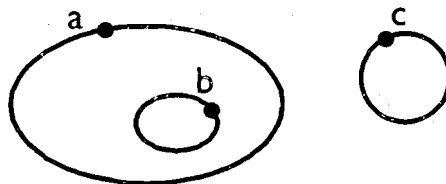
Figure 4.5. A spatial scene with lines a-c of type A.

Type C: when two patches meet, the common boundary segment is of type C. Removing such a line when removing the boundary of one of the patches will always result in a semi-open set in the sketch (Figure 4.6). It follows that a boundary segment of type C cannot be removed from the sketch in any case.

48
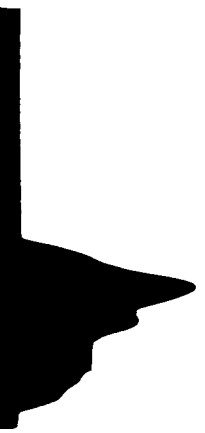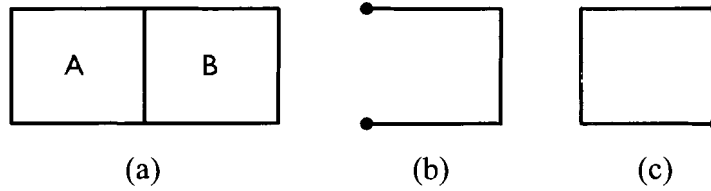
Figure 4.6. Removing a line of type C: (a) a spatial scene with patches A and B, (b) after removing patch A, and (c) after removing patch B.

Type B, D, E, F, and G: for these types it is uncertain whether or not to remove the specific line. The difference between the situations where a line can be removed and the situations where a line cannot be removed is in the number of regions that the line is part of. Where a line cannot be removed, it is because that line is part of one or more regions in the sketch independent of how many patches the line is part of. This information can only be gained by knowing the final result of the region extraction process. Knowing the final result is clearly impossible before the end of the process and, therefore, a different approach is chosen: first, all segments of a region's boundary are removed from the sketch except for the segments of type C. Second, with the remaining lines in the sketch, new patches are built and checked if there are any semi-open sets. A semi-open set appears where a line's end has an intersection type less than m2, which occurs where a line is open. This condition can easily be checked and where this is the case, one or more lines that close the semi-open set (i.e., increase the meet relation to m2 or more) have to be brought back to the sketch.

49

with a less continuous boundary are accepted by applying the increased minimum continuity threshold. The maximum threshold can be set to any value higher than the minimum threshold and less than 180 degrees (Equation 4.1).

$$0 <= minThreshold < maxThreshold < 180\ degrees \qquad (4.1)$$

4.2.7 Remaining Patches

The region extraction algorithm does not always extract regions until no patches are left in the sketch. Sometimes no continuous boundary can be found so that patches are

4.2.4.3 Finding Closing Segment(s)

More than one line called the *closing line* can be brought back into a sketch so that no semi-open sets are left in the spatial scene. A rationale which closing line should be brought back into a sketch is outlined here:

Any closing line must be part of the removed region's boundary, it must meet one or more open lines at their open end, and it must not be an open line itself. These constraints reduce the set of possible closing lines. If there is only one line that meets all the constraints, then this line is clearly the closing line and should be brought back into the sketch. It is possible, however, that several lines are candidates to be the closing line. In this case the line that introduces back the least elements of the removed region is chosen. For example, in the case where there are several lines that meet these conditions, any line that connects more than one open line is preferred over lines that only connect to one line. In another example, a closing line introduces back into the sketch a part of the removed region's boundary, whereas another closing line
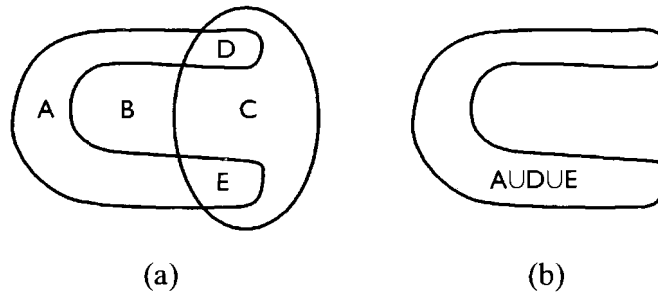
Figure 4.8. Removing a region from a sketch that results in a gap in the spatial scene: (a) the original spatial scene and (b) the scene after removing the region C∪D∪E where the patch B is lost.

4.3 Summary

The PSI algorithm of extracting regions from line drawings was formally described in detail in this chapter. This formalization can be used as the basis for the prototype implementation in chapter 5. First, the scope and underlying assumptions of the algorithm were defined Then, the algorithm was described in detail, given pseudo code and descriptions of each part and of different modules.

Chapter 5

SKETCH INTERPRETATION PROTOTYPE

The PSI algorithm in Chapter 4 is implemented in a prototype application that serves two purposes: (1) to test the model of object extraction from paper sketches during the model development as well to test the final proposed model and (2) to prove its functionality.

5.1 Digital Sketch

A digital sketch is a formal representation according to the mathematical model of a sketch as described in chapter 3. This model is here divided into its component classes (e.g., sketch, point, stroke, region, relation matrix) in order to describe a spatial scene and to perform the necessary manipulations for interpreting a sketch. Each class itself has a set of variables and methods and has been developed specifically for this project.

5.1.1 Sketch Class

The sketch class is the main class for a sketch. It contains a list of other classes (e.g., point, stroke, region) and a relation matrix along with properties such as the continuity thresholds and a set of methods. This class completely describes a spatial scene. The sketch interpretation prototype uses several of these classes: a patch sketch to describe the initial sketch with all the patches, a region sketch to describe the result of the region

extraction algorithm and an array of sketches to show the intermediate steps of the region extraction process. The latter clearly visualizes the strategy of removing one extracted region at a time.

Sketch
■ List of Points (SketchPoint)
■ List of Strokes (SketchStroke)
■ List of Regions (SketchRegion)
■ Relation Matrix
▪ Set of Properties
Set of Methods

Figure 5.1. The Sketch class.

A point or stroke in the sketch is only stored once in the Sketch class's list of points or list of strokes. Two regions that share a stroke reference the same stroke object in the sketch class and two strokes that share a point reference this object in the list of points.

5.1.2   SketchPoint Class

The SketchPoint class is a simple class storing x- and y-coordinates and a point identifier along with a set of methods. A list of points is stored in the Sketch class and points are referenced from the SketchStroke class.

SketchPoint
■ Identifier ■ X-coordinate ■ Y-coordinate
Set of Methods

Figure 5.2. The SketchPoint class.

5.1.3   SketchStroke Class

The SketchStroke class models a line or polyline. This class contains a list of references to the sketch's list of points. For further processing, a SketchStroke also contains a set of methods, a set of properties such as azimuth. Two lists contain references to other strokes that are adjacent to this stroke, one for the start, and one for the end of this stroke. These lists are important for the region extraction algorithm when tracing along adjacent strokes.

SketchStroke
■ Identifier ■ Azimuths ■ List of adjacent strokes ■ List of points (references to list of points in sketch)
Set of Methods

Figure 5.3. The SketchStroke class.

56

5.1.4 SketchRegion Class

A SketchRegion models a patch or an extracted region in a sketch. A list of strokes references objects in the sketch's list of strokes. Other properties are an identifier and gestalt values.

SketchRegion
▪ Identifier ▪ List of Strokes (references to list of strokes in sketch) ▪ Gestalt values
Set of Methods

Figure 5.4. The SketchRegion class.

5.2 Paper Sketch-To-Digital Sketch Conversion

A sketch that is drawn on paper is converted into a digital representation by scanning the sketch. Before applying the PSI algorithm, this raster image of the sketch is converted into a digital sketch. Existing methods are used for raster-to-vector conversion of a sketch. This section describes these methods. These methods, however, are not sufficient to convert complex images such as aerial photographs or satellite images. The development of such methods is an ongoing research effort and is beyond the scope of this thesis.

5.2.1 Extracting Outlines of Sketched Objects

Typically *edge detection* is used to extract object outlines from images, such as aerial photographs or satellite images. The process of edge detection yields a binary image

with pixels corresponding either to the background or to an object edge (Castelman 1996; Haralick 1985). A sketch is a binary image with white background pixels and black foreground pixels. Black pixels are all on the boundary of a region and, therefore, when working with sketches, no edge detection is necessary. Before the boundary pixels can be linked and converted into a vector representation, however, the boundary is thinned to a one-pixel-wide edge in a process called *thinning*.

Alternative approaches to thinning and linking are the use of *tracking algorithms* or *active contour models*, also called *snakes*. Tracking algorithms follow along connected pixels whereas snakes are energy-minimization splines that are pulled towards features in an image. In the case of sketches, splines are pulled towards the boundaries of sketched objects (Kass *et al.* 1988).

5.2.1.1 Binary Image Thinning

Thinning reduces a curvilinear object to a single-pixel-wide line (Castelman 1996). It erases black pixels such that an object without holes erodes to a minimally connected stroke located equidistant from its nearest outer boundaries (Pratt 2001). In the case of sketched objects, thinning reduces a thick boundary of a region to a thin edge.

5.2.1.2 Linking of Edge Points

The pixels on the boundary of an object are labeled *edge points*. A binary image showing only the locations of the edge points is called a *binary edge image*. *Edge point linking* associates nearby edge points in order to fill in gaps left by noise and to create a closed, connected boundary (Castelman 1996). Each connected boundary can then be

represented by a polyline, thus creating a vector map of the original sketched spatial scene.

5.2.1.3 Tracking Algorithms and Active Contour Models

Tracking is a semi-automatic raster-to-vector conversion of line features. It typically requires a starting point and a direction. Some algorithms give the users choices on how to proceed when reaching obstacles (e.g., intersections) or let the user specify a default action prior to starting the tracking algorithm (e.g., turn left, turn right, go straight, or stop tracing) (Environmental Systems Research Institute 2001). The PSI algorithm uses gestalt theory to automatically continue tracking when identifying regions or when building the patches in a sketch.

A different approach from conventional tracking is a technique that uses active contours, referred to as *snakes* (Kass *et al.* 1988). Snakes typically require initialization with control points, which are used in an energy minimization process to provide edge-based delineation (Doucette 2002). Rather than processing the entire image, snakes are applied selectively to parts of the image (Blake and Isard 1998).

Tracking as well as snakes are a direct raster-to-vector conversion (i.e., thinning or linking is not necessary). When applied to sketches, however, they do not outperform the more simple vectorization algorithm in today's image processing and GIS software.

5.2.1.4 Cleaning Topology

A quickly drawn sketch without concern for detail contains noise, in the form of *overshoots, undershoots,* and *slivers* (Bolstad 2002). Overshoots and undershoots

typically originate from either digitizing a line too far (overshoot) and, creating a short spike, or digitizing a line too short (undershoot), creating a gap. Slivers can occur by digitizing the same line twice or by overlaying two data layers that are not precisely registered (Figure 5.5).

In sketches that only contain regions, every form of noise must be cleaned for a correct representation of the sketched regions because of the following reasons: undershoots result in open polylines instead of closed boundaries; overshoots result in lines that are only connected at one end and thus cannot be part of any region's boundary. Slivers result in small regions, where two regions overlap and do not only introduce additional regions, but might also change the topological relation between two regions.

Today's GISs provide methods for cleaning the topology of a spatial scene. Overshoots are detected by specifying a *dangle length* and a *fuzzy tolerance* for closing gaps and merging slivers (Environmental Systems Research Institute 1994; ERDAS 1997). These methods are sufficient for cleaning the topology of a sketch and, therefore, this work does not further investigate in this task.

| (a) | (b) | (c) |
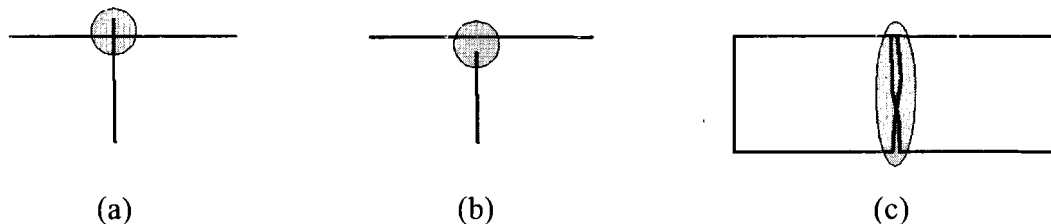
Figure 5.5. Drawing noise that can occur in a sketch: (a) overshoot, (b) undershoot, and (c) sliver.

60

5.2.2 Conversion of Data other than Sketches

The methods described in the previous section are applicable to binary images of a sketched scene. Because of the relatively simple data model of a digital sketch, these methods yield good results in extracting object edges and edge linking. One might, however, desire to apply the PSI algorithm of this thesis to other data than a binary image of a sketch.

In order to apply the PSI algorithm to more complex imagery, the chosen image is converted into a sketch-like representation. For this conversion, object boundaries are to be extracted from the image. Whereas this is a simple task performed on a sketch, it becomes a complex task when dealing with detailed gray level or color images. If a successful image-to-sketch conversion method can be found, then the proposed algorithm of this thesis can be applied to the converted image.

5.3 Guided Tour

The PSI algorithm is applied to the digital sketch obtained as outlined in section 5.2. This process is illustrated here in terms of a guided tour of the prototype application. The prototype was implemented with the user interface described in this section and the necessary project specific classes. It was developed in an object-oriented environment using the core java class libraries for the necessary methods and the java swing classes for the graphical user interface. The prototype runs on any operating system with an installed java virtual machine.

5.3.1 The Application Window

Launching the prototype application draws the application window on the screen. The window is segmented into areas for menus, toolbar, sketch properties, and a sketch pane for the visual output of a sketch.
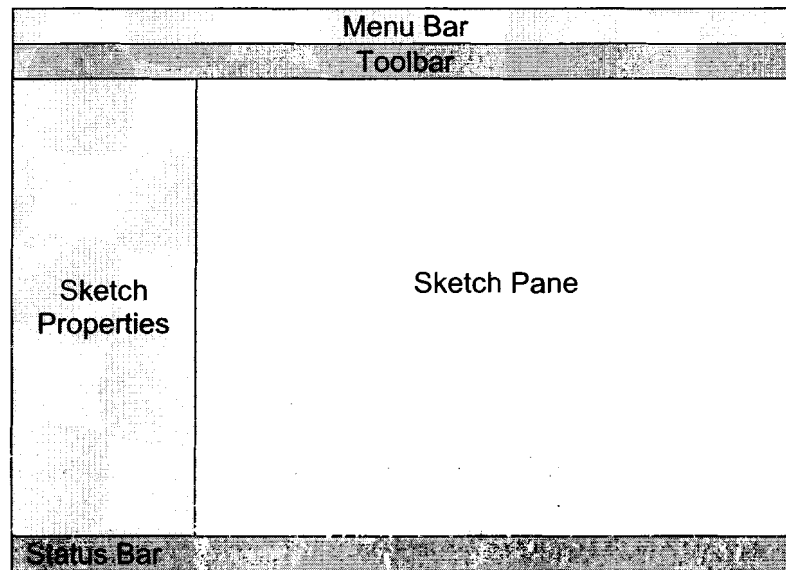
Figure 5.6. The segmentation of the application window in its components.

5.3.1.1 User Interface Metaphor

Displaying a sketch is based on the map metaphor. This visualization is closely related to today's GISs that display geographic data using a map metaphor. Tools are accessible through menus and toolbar buttons.

5.3.1.2 Interaction

Human-computer interaction (HCI) uses the intuitive mode of WIMP (windows, icons, menus, pointers) that is widely known in today's computer environments. Necessary

user input is brought to a minimum for a simple use of the prototype and is received

through a set of menu items or toolbar buttons.

Figure 5.7. The menu bar and toolbar of the prototype application. From left to right: open, reopen, save, save as, show sketch properties, show relation matrix, highlight previous region, highlight next region, zoom in, zoom out, zoom to extent, extract regions on/off, set modules, fill gaps on/off, show previous sketch in process tab, show next sketch in process tab.

5.3.2 Processing a Sketch

To process a sketch, a sketch file is chosen through a menu item or by clicking on the

open button on the toolbar. The prototype reads ESRI's interchange format (e00) or a

text file. If an interchange file is chosen, the ARC section (arc coordinates and

topology) of the interchange file format is read. The patches are built from this

information regardless of whether there is a PAL (polygon topology) section or not.

Alternatively, the prototype reads a simple text file containing a list of points with x-,

y-, and z-coordinates (z-coordinates are arbitrary and will be ignored by the prototype),

and an identifier of the line that the point is part of (Figure 5.8).

```
Pt#, X, Y, Z, Line#
1,398.0,162.0,0.0,1
2,383.0,179.0,0.0,1
3,383.0,179.0,0.0,2
4,374.0,180.0,0.0,2
5,354.0,186.0,0.0,2
```

Figure 5.8. The text file format that can be read by the prototype.

For the evaluation process of the proposed model and prototype, a simple drawing application (Sketch Draw) was developed that will save a spatial scene in the above text file format.

5.3.3   Exploring the Sketch

A sketch is read from a file and the result of the region extraction process is displayed in the sketch pane of the prototype application. The sketch pane is divided into three tabs: a patch tab, a region tab, and a process tab, each showing the same sketch at different stages of the region extraction process. Detailed information about the elements of the sketch is given in the sketch properties panel.

Region identifier

Number of strokes in region / Region is patch

Gestalt values

Stroke identifier

Point identifier

Coordinates

Azimuths of stroke

Lists of strokes connecting to this stroke
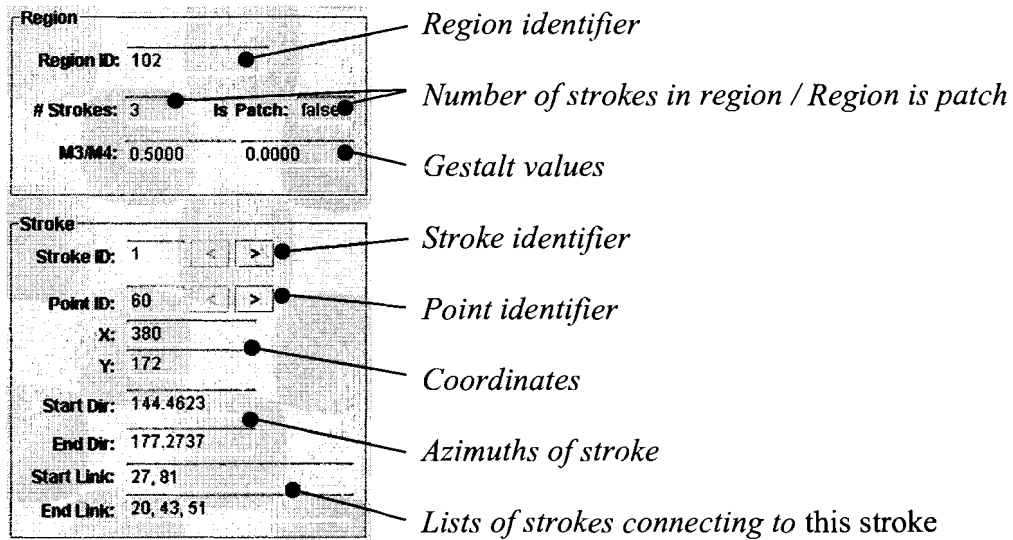
Figure 5.10. The Sketch Properties Panel.

5.3.3.2   Patch Tab

The patch tab shows the sketch before the region extraction algorithm was applied. Only numbered patches are displayed (Figure 5.11a). The numbering of the patches is arbitrary.

5.3.3.3   Region Tab

The region tab shows the sketch after the region extraction process. Numbered identified regions are displayed in different colors for better visualization of the regions (Figure 5.11b). In this sketch view, the first region is highlighted. Through a menu or toolbar command the user can cycle through each region in the sketch while the region's properties are displayed in the sketch property area. If region extraction is disabled or if no regions could be identified, the region tab shows the same sketch as

the patch tab. In this case, using the command to cycle through different regions will simply cycle through the patches in the sketch.
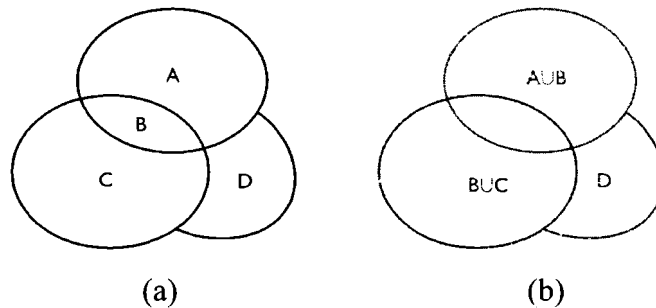
Figure 5.11. A sketch viewed in (a) the patch tab and in (b) the region tab.

5.3.3.4 Process Tab

The process tab shows the sketch at different stages of the region extraction process. The process is visualized in its steps, outlining how region extraction took place. At first, the original sketch with all its patches is displayed. Stepping forward with the corresponding menu or toolbar command will display the same sketch after removing the first identified region and so on until all regions were removed(Figure 5.12). This visualization is helpful to understand possible errors in case of misinterpretations by the algorithm.
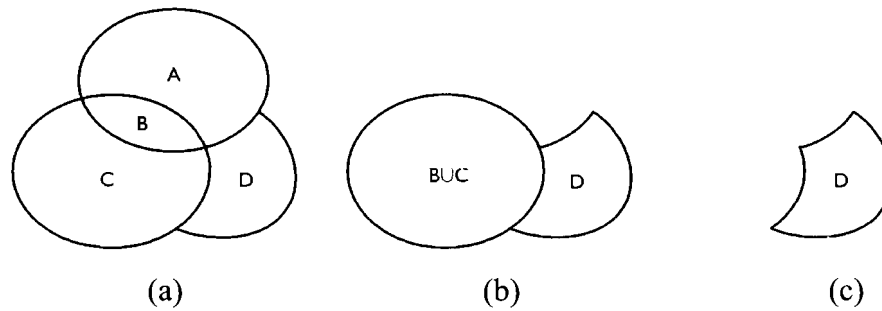
Figure 5.12. A sketch shown in the process tab: (a) the original sketch, (b) after removing an identified region, and (c) after removing the next identified region.

5.3.4 Saving a Sketch

A sketch can be saved in the sketcho (Blaser 2000a) format with the file extension *dmp*. The patches are saved if the patch tab is active, otherwise the identified regions are saved. The sketcho format allows further use of the interpreted sketch to form a spatial query executed on a set of other sketches using the prototype of Sketching Spatial Queries (Blaser 2000a).

5.4 Summary

This chapter outlined an implementation of the digital sketch and the region extraction algorithm. Project specific classes of the digital sketch were described as well as the user interface of the prototype application. Input from and output to the user were given in detail.

The prototype application served as a test bed for the proposed model and was used in the model evaluation described in Chapter 6.

Chapter 6

MODEL EVALUATION

The subset of all regions in a lattice that is created by the set module applying the notion of good continuity (continuity set module) is significantly smaller than the complete set of regions in the lattice (full set module). The hypothesis of this thesis, however, asserts that the result of the region extraction algorithm is similar using the suggested subset of regions. This chapter evaluates the hypothesis through the PSI algorithm.

6.1 Evaluation Design

The development of the PSI model and its implementation in a prototype application were two parallel processes, going back and forth between the two. During this development phase, the model was tested with sample sketches drawn by the author. To gain an unbiased view, however, a set of sketches was obtained from people that were not involved in this project's design.

Subjects were asked to draw one spatial scene and identify their interpretation of the sketch. In total, 36 sketches were collected, of which 30 could be used for the analysis. Six sketches were not used because they did not adhere to the spatial data model. The 30 sketches were processed with the PSI prototype. The resulting sketches returned by the prototype were then compared to the interpretation given by the subjects. This

comparison was done with the prototype of Sketching Spatial Queries (Blaser 2000a) which returned similarity values for each comparison.

6.2 The Survey

A Web-based survey was conducted, giving the participating subjects the opportunity to draw and submit their sketch through a Web browser interface. This human subject testing included information about the project, an informed consent form, drawing instructions, and the necessary drawing application.

6.2.1 Surveyed Subjects

The group of subjects consisted mostly of graduate students of the Department of Spatial Information Science and Engineering at the University of Maine. Other subjects from various other departments also participated in the survey. In order to ensure that sketches were collected from different subjects, the IP-address of each client machine from which a sketch was submitted was recorded. No personal information of the subjects was collected as the survey was aimed at a general group of subjects.

6.2.2 Instructions Given to Subjects

The subjects were given brief instructions on what to draw. The instructions guided the subjects to draw a sketch within the scope of this thesis and to a sketch that might challenge the model's performance (e.g., overlapping regions vs. trivial cases such as all disjoint regions):

70

- *Your sketch can be abstract (e.g., combination of circles, squares etc.) or it can represent a spatial scene such as two overlapping layers. Think of data in a GIS or CAD with two layers overlapping each other. For example:*
  - *Soil type overlapping land parcels*
  - *Habitat areas of different animals*
- *Approximately 4–6 regions*
- *Only draw regions (e.g., closed polygons; each loop should be closed)*
- *The regions can have any regular or irregular shape.*

http://www.spatial.maine.edu/~wuersch/sketchSurvey.html

Region Extraction from Line Drawings
Human Subject Testing - Collecting Sketches

Markus Wuersch's Home Page

<< 1 ... 2 ... 3 ... 4 ... submit sketch >>

Instructions:

Read carefully to the bottom of this column and use the applet to the right to draw a sketch (see instructions below the applet):

- your sketch can be abstract (e.g., combination of circles, squares etc.) or it can represent a spatial scene such as two or more overlapping layers.
  Think of data in a GIS or CAD with two layers overlapping each other. For example:

  o soil type overlapping land parcels

  o habitat areas of different animals

  -> see some examples

- approximately 4-6 regions:

- only draw regions (e.g., closed polygons; each loop should be closed)

- the regions can have any regular or irregular shape.

When you are done drawing your sketch, do not submit yet! You need to provide some more information about your sketch.

Drawing applet

1,2,3,8
2,3,4,6,7
2,4,5

OK

Start Over   Clear Labels   Label   Define Regions   Submit Sketch

having problems drawing? -> more detailed instructions
- to draw a line: click a series of points or drag mouse
- to end a line, [double click]
- to delete a line: [ctrl+click]
- crossing lines: only cross one line at a time

Figure 6.1. A screen shot of the on-line survey: instructions on the left and the drawing applet on the right. A subject drew a sketch and defined the regions by listing the patches of each region in a separate window (top right of drawing applet).

72

6.2.3 Collected Information

The survey was conducted anonymously and, therefore, the only information collected was the sketch drawn by the subject. The sketch was stored as a text file on server. This file included the topological information, labels of each patch, and a description of the composition of each region (e.g., which patches are contained by a region).

6.2.4 Spatial Data Model of the Received Sketches

The drawing applet that was provided to the subjects to draw a sketch is a vector drawing tool. It captures sketches with a simplistic data model according to the following rules:

- The only topological objects contained in a sketch are points and lines.

- A line is defined as the interior between exactly two points where the interior is the shortest path between two points (e.g., a straight line).

- The interiors of any two lines are disjoint.

This data model is chosen to guarantee a clean topology, which is further enforced by a clean operation. This operation eliminates equal lines and condenses points that are within proximity of each other to the same location. Each two end points of a line are stored in a text file along with a unique identifier of the line (Figure 6.2). Subjects were also asked to label each patch and describe each region by the patches it contains. This information is stored in a different section of the sketch file (Figure 6.3).

73

In using the described drawing tool, no raster-to-vector conversion had to be applied to any sketch. The process of extracting regions from a sketch, however, starts with a sketch drawn on paper and, therefore a raster-to-vector conversion is a necessary step. Such a conversion, however, is not part of this thesis and, therefore, the chosen approach fulfills the requirements for testing the PSI algorithm.

```
*Pt#, X, Y, Z, Line#
1,157.0,288.0,0.0,1
2,154.0,116.0,0.0,1
3,154.0,116.0,0.0,2
4,243.0,116.0,0.0,2
5,334.0,114.0,0.0,3
6,342.0,215.0,0.0,3
```

Figure 6.2. Three lines stored in a text file.

```
#labels
1,115.0,105.0,0.0,0
2,179.0,78.0.0.0,0
3,123.0,161.0,0.0,0
*

*regions:
* 1,2
* 2,3
```

Figure 6.3. Three labels and two definitions of regions stored in a text file.

6.3 Processing the Test Sketches

For each collected sketch, regions are extracted manually according to the description obtained from the online survey. The resulting spatial scene is termed *ground truth* and is used to evaluate the accuracy of the PSI algorithm.

74

Using the prototype application, each sketch is processed twice: (1) using continuity and (2) using the complete set of possible regions (the setting for continuity at m3-intersections is set to stop if no continuity is found, see Chapter 4.2.3.2). The resulting spatial scenes are then compared to the ground truth using the prototype application *Sketcho* from Sketching Spatial Queries (Blaser 2000a). This process uses the ground truth as a query input, operating on the two interpreted sketches created by the sketch interpretation prototype. As a result, Sketcho returns similarity values (between 0% and 100%) for each sketch.

If the hypothesis is true, the similarity values for each pair of sketches will be comparable. The similarity values also quantify the accuracy of the region extraction process: a similarity value of 100% shows correct interpretation of the sketch, less than 100% indicates a deviation from a correct interpretation.

6.4 Results

The results of the assessment on the collected sketches are grouped into (1) correctness, expressed in similarity to the corresponding ground truth, (2) processing time, and (3) an analysis of the shortcomings of this algorithm.

6.4.1 Correctness

In total, 36 sketches were collected through the on-line survey. From these sketches, 30 adhered to the spatial data model (i.e., they did not contain any complex regions) and were used for a first assessment (Table 6.1). Processing some of these sketches yielded

incorrect results compared to the ground truth. This misinterpretation was due to tessellations contained in three sketches. These three sketches were removed from the test set so that the test is repeated with the remaining 27 sketches.

Still, some sketches are misinterpreted and in some cases, this is due to sketches that most likely could not be interpreted correctly without additional information. We identified three such sketches and the test is repeated one more time with the remaining 24 sketches.

In this last test, 75% of the sketches were interpreted correctly using the continuity set module and 62.5% when using the full set module.

Test Set		Set Module Full	Continuity
All 30 sketches	Correctly Interpreted Sketches	50.0%	60.0%
27 sketches (no tessellations)	Correctly Interpreted Sketches	55.6%	66.7%
24 sketches (only sketches that can clearly be interpreted without additional information)	Correctly Interpreted Sketches	62.5%	75.0%
	Average Similarity	92.4%	93.9%
	Smallest Similarity	46.4%	37.6%
	Largest Similarity	99.9%	99.9%

Table 6.1. Summary of test results: correctness.

6.4.1.1 Significance Tests for Sample Sketch Set

A significance test is performed to show that the set of sample sketches is representative. One sample significance tests show that at a 5% significance level the

hypotheses of the tests that the actual means for the continuity set and the full set have similarity values of 99.99% are not rejected (Table 6.2).

| | Set Module | |
	Continuity	Full
Mean	93.86	92.34
Standard Error	3.129	3.084
Confidence Levels (95.0%)	6.472	6.379
t-stat	-1.959	-2.482
t-critical (df = 23)	2.069	2.069
Hypothesis Test Result	Not rejected	Not rejected

Table 6.2. Summary of significance test for sample sketch set.

6.4.1.2 Significance Tests for Thesis's Hypothesis

The tests in Section 6.4.1.1 showed that the means from both continuity and full set can approximate to 99.99% for large samples. A comparison between the two sets, however, is still necessary in order to support this thesis's hypothesis. The hypothesis that the continuity set yields better results than the full set at large samples is tested by using two sample one-tailed significance tests (Equation 6.1, Table 6.3). At a 11% significance level, the mean of the continuity set is greater than that of the full set. This result confirms that in 89% of the cases it can be expected that the thesis's hypothesis is true.

$$H_0 : \mu_C - \mu_F = 0 \qquad (6.1)$$
$$H_1 : \mu_C > \mu_F$$

where:
H_0 = Hypothesis of test   *μ_C = Similarity mean using continuity the set module*
H_1 = Alternative hypothesis of test   *μ_F = Similarity mean using the full set module*

Significance Level	5%	11%
t-stat	1.275	1.275
t-critical one-tail	1.714	1.261
Hypothesis Test Result	Not rejected	Rejected

Table 6.3. Summary of Significance tests for the thesis' hypothesis.

6.4.2 Processing Time

Applying the continuity set module executed on average 119 times faster than when using the full set module (Table 6.4; processing was done on a computer running Windows XP, 2.99 GHz processor, and with 1.00 GB of RAM).

	Set Module	
	Full	Continuity
Average Processing Time [s]	26.6	0.2

Table 6.4. Summary of test results: processing time.

This difference is due to often very large sets of regions identified by the full set module. Figure 6.4 shows the number of regions that are returned by the set modules during the first iteration of the region extraction process compared to the number of patches in the sketch. The average number of regions returned is 11 for the continuity set module, and 2056 for the full set module, which is a multiple of 186. This number depends on the number of patches in a sketch and the topological relations between the patches. Generally, the more patches in a sketch, the more regions are identified by the set modules.
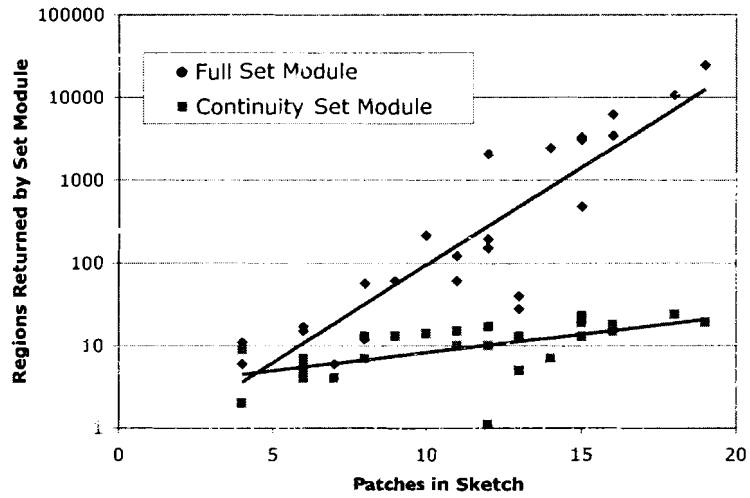
Figure 6.4. The number of regions identified by the set modules in relation to the number of patches in the sketch.

6.4.3 Shortcomings of the Algorithm

The evaluation shows a shortcoming of correctly interpreting six sketches of the set of

24 test sketches (Figure 6.5).
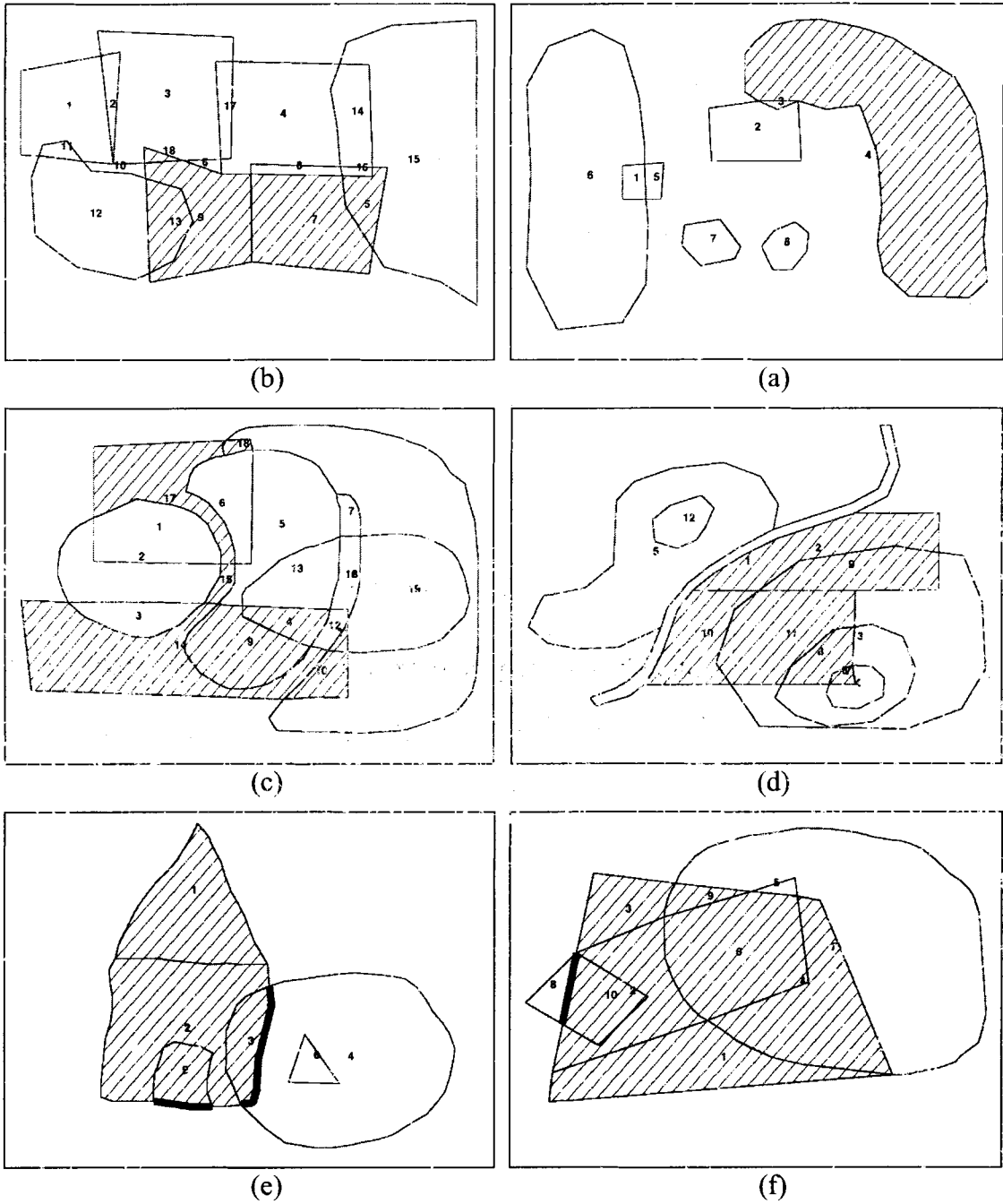
Figure 6.5. Six sketches that were misinterpreted by the PSI algorithm: (a, b, c, d, and e) using continuity identified incorrect regions; (e and f) after identifying a region, boundary segments (thick lines) were removed incorrectly from the sketch.

80

The region extraction from these six sketches was analyzed in more detail and two reasons for an incorrect interpretation were found: either an incorrect region was identified as having the best gestalt, or a region was removed incorrectly.

6.4.3.1 Insufficient Gestalt Measure

The analysis showed that in five cases, continuity is not the major factor used by people's perception to order their visual input (Figure 6.5 a, b, c, d, and e). Regions were identified by the continuity set module, that should not have been extracted. These regions, however, had a very good gestalt value and, therefore, extracted from the sketch. On the other hand, regions with a regular shape (e.g., squares, rectangles) were not identified by the continuity set module or the gestalt module did not classify these regions as having a good gestalt. In any case, however, the continuity set module returned at least one correct region.

6.4.3.2 Incorrect Removing of a Region's Boundary

In two cases the rationale of removing an identified region's boundary (Chapter 4.2.4) from the sketch returned incorrect results (Figure 6.5 e and f). In order to gain better results refinement of this rationale is necessary.

6.5 Summary

This chapter evaluated the performance of the PSI model in regarding correctness and processing time. For this purpose, pairs of results of the PSI process gained by applying the two set modules were compared to a manually extracted spatial scene.

Sketches collected through an on-line survey were used in order to gain a non-biased test procedure. The results from 30 sketches showed strong support for the hypothesis that using the continuity set module yields results comparable to using the full set module. The number of correctly interpreted sketches was 12.5% higher, the average similarity is 1.5% higher, and the processing time is on average 119 times less. The PSI algorithm misinterpreted sketches where continuity was not sufficient to identify the correct set of regions, and/or a region was not correctly removed from the sketch.

Chapter 7

CONCLUSIONS AND FUTURE WORK

This chapter summarizes the findings of this thesis and discusses possible future research topics as well as extensions and refinements to the current model.

7.1 Summary

This thesis is concerned with feature extraction from line drawings. It focused on the development of an approach of automated, perceptual sketch interpretation algorithm -- the extraction of regional objects from paper sketches. It followed Blake and Isard (1998) who stated that part of a feature extraction task is to group recovered features according to the object to which they belong. In this case, grouping refers to combining lines into region boundaries.

The line drawings used in this thesis were converted into a digital sketch which is a vector representation, free of such drawing errors as overshoots, undershoots, and slivers. To obtain a clean vector representation, methods in commercially available GISs were used.

Because the PSI algorithm attempts to model people's perception, laws of organization and the law of pragnanz from gestalt theory were essential to the success of the algorithm. Gestalt theory was, therefore, used in combination with spatial

reasoning. These theories are viewed under the scope of this thesis and were refined where necessary.

We introduced an algorithm that uses the law of good continuation to work with a small set of possible regions in comparison to considering all the regions of a full lattice. These two parts are defined as set modules within the algorithm and the comparison of these two modules led to the definition of the hypothesis in chapter one: *Using the continuity set module in the Perceptual Sketch Interpretation algorithm yields comparable results of interpreted sketches to using the full set module.*

The results of the model evaluation show strong support for the hypothesis. Conclusions of these results are described in the following paragraph followed by possible refinements and extensions to the current algorithm and future work.

7.2 Major Results

The research conducted within the scope of developing the PSI algorithm led to three result statements that are described here.

- *The continuity set module returns a set of regions that contains at least one region that corresponds to a region in people's mental model.*

This conclusion is stated as an assumption in Chapter 4.2, supported by the results of the model assessment. Because this assumption clearly holds for the full set module and because the results of the assessment have shown evidence that the

84

hypothesis is true, it can be concluded that the assumption also holds for the continuity set module.

This conclusion is supported by the analysis of the incorrect interpreted sketches (Chapter 6.4.3), where other reasons were identified as the cause of the incorrect results.

- *Good continuity is, amongst other gestalt laws, one of the major factor used by people's perception to order visual input into meaningful objects.*

The model assessment shows that, depending on the set of test sketches used, 60% to 75% of the sketches were interpreted correctly using the continuity set module. This result gives evidence that good continuity is a major factor used in people's perception. Some of the test sketches, however, were extreme cases that were misinterpreted by applying the notion of good continuity. Whereas the PSI algorithm can be applied to any regular drawing, these cases ask for additional reasoning than the notion of good continuity. Other gestalt laws may need to be incorporated into the PSI algorithm.

This conclusion refers to assumption 1 in Chapter 4.2 but also applies to assumption 3. The latter assumption states that the region with the best gestalt has a corresponding region in people's mental model. Clearly, this assumption depends on the definition of a good gestalt, which in turn depends on the gestalt laws used.

The analysis of shortcomings of the PSI algorithm (Chapter 6.4.3) also shows support for this conclusion. It follows that, in order to improve the results of the

region extraction process, the set module and the gestalt module should be refined and/or extended with additional laws of organization.

- *The continuity set module is the preferred approach over the set module, which returns all the regions of the full lattice.*

The comparison of the two approaches using different set modules overwhelmingly shows support for the continuity set module. The number of regions processed is remarkably lower (by a factor of 186) for the continuity set module, but it also resulted in a slightly higher correctness of the interpreted sketches.

When processing small sketched spatial scenes, the processing time gained is between a split second and 7:36 minutes. If this algorithm is to be applied to a large spatial scene, the gained processing time is important.

7.3 Future Work

The model assessment has shown possible future research topics as well as refinements and extensions to the current model. Such topics are described here and serve as a guidance to future work and inspiration for new research efforts. First, further analysis of the algorithm's different settings is proposed as a mean to guide future research. Then, the possible application of the algorithm on other data than sketches drawn on paper is discussed before refinements and extensions to the algorithm are proposed.

7.3.1 Detailed Analysis of Different Settings of the Algorithm

The model assessment used default settings for the PSI algorithm. These settings were chosen based on the definition of the model and by experience through the development of the model and its implementation (see Chapter 4.2.3.2, 4.2.6, and 4.2.8). Experimenting with different settings could improve the results of the region extraction process and answer the following research questions:

Question 1:   *Is there a correlation between scene characteristics and distinct settings of the algorithm?*

Answering this question requires the definition of scene classes according to their characteristics. A correlation analysis between results of the region extraction process on a particular class and the settings used could reveal a correspondence.

Question 2:   *Which settings influence the result the most?*

To find the answer to this question would allow us to aim future work at the most important parts of the algorithm.

In order to answer this question, a regression analysis could be performed on a set of sketches that are the result of processing a sketch with different combinations of settings. A brief regression analysis is described here and shows that settings can be distinguished as having little or much influence on the result (Figure 7.1). This analysis, however, was run on the results of only eight processed sketches and should be repeated with a larger set of sketches to gain reliable results.
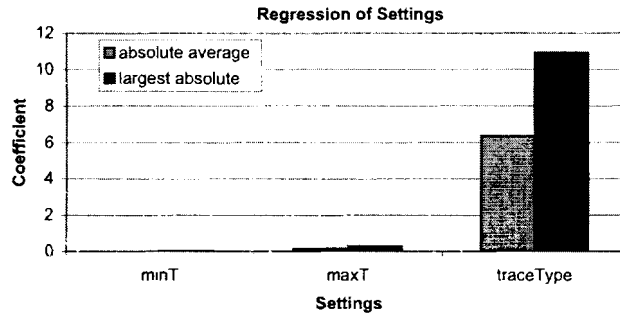
Figure 7.1. The result of a regression analysis over different settings. Minimum and maximum continuity thresholds have little influence on the result (for thresholds see Chapter 4.2.6, and for traceType see Chapter 4.2.3.2).

7.3.2 Use of the Algorithm on other Data than Sketches

In theory, the algorithm of this thesis can be applied to any data that can be transformed into a sketch-like representation. This is either a raster or a vector representation of a line drawing. Accordingly, researching region extraction from data other than sketches involves a conversion to a sketch-like representation, the analysis of the performance of the PSI algorithm, and proposing further refinements and extension.

7.3.3 Refinements of the Algorithm

Refinements to the current algorithm that would improve the results of the regions extraction process are described here.

7.3.3.1 Metric Refinement of Meet Relation Between Two Regions

The current data model requires a completely clean topology of the sketched lines in order to apply qualitative reasoning as it is described in this thesis (e.g., binary

topological relations). If a scene is to be analyzed on a more detailed level, however, metric aspects as studied by Egenhofer (1997) and Shariff (1996) become relevant (Egenhofer and Mark 1995).

A hand drawn sketch contains drawing errors that can influence the outcome of the region extraction process by changing the topological relation between two regions. Specifically if a 0-meet relation is mistakenly drawn as a 1-meet relation, then the two regions can be combined to a new region. In the case of a 0-meet relation, however, the regions cannot be combined as this would result in a complex region.

To incorporate such metric refinements for the topological relations would possibly result in more accurate sketch interpretations. It also would allow it to be more automated as it would rely less on a clean topology of the sketch.

7.3.3.2 Continuity

Currently, the continuity angle is measured between two adjacent boundary segments. In a case where any of the two boundary segments is very short, the continuity angle might not be representative for the perceived continuity (Chapter 3.4.1). Instead, a somewhat generalized boundary segment or a global curvature measure could be used to calculate a continuity angle.

7.3.3.3 Filling Gaps

The option to fill gaps can be vital for a correct interpretation of a sketch. In some cases, however, gaps should be extracted the same way as regions in order to fill the gaps with regions that correspond to people's mental model (Figure 7.2).
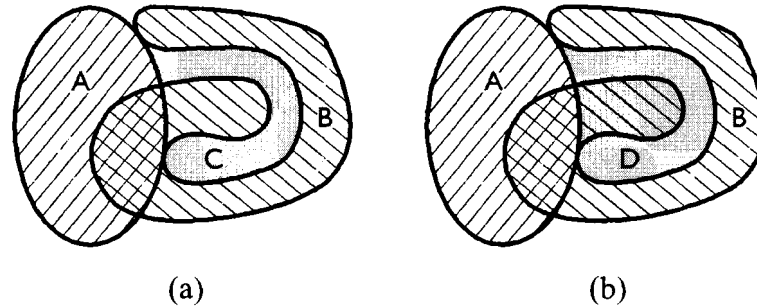
Figure 7.2. Refining the method to fill gaps: (a) using patch C as it is done by the current algorithm and (b) using continuity to find region D with a better gestalt.

7.3.3.4 Drawing Errors

Drawing errors, such as overshoots, undershoots and slivers, are corrected by a clean function before the actual region extraction process. In doing so, some information that could reveal more details about the possible regions in a sketch might be compromised. For example, a sliver indicates that the same line is drawn twice and, therefore, these lines are most likely where two regions meet. In cases where drawing errors occur, they could give better insights on regions in a sketch thus improving the result of the region extraction.

7.3.3.5 Removing a Region's Boundary

The analysis of the shortcomings of the PSI algorithm in Chapter 6.4.3 has shown that the rationale of removing an identified region's boundary, outlined in Chapter 4.2.4, could not be relied on at all times. The result of the PSI algorithm could be improved by refining this rationale.

7.3.4 Extensions to the Algorithm

The PSI algorithm of extracting regions from line drawings has shown satisfying results for sketches within the project scope. To improve the correctness and to extend the use of the algorithm, however, further extensions are proposed here.

7.3.4.1 Open Lines

Currently, any open line in the sketch is ignored and deleted from the sketch. This is done, because no open line can be part of a simple region. By deleting a line, valuable information might be lost. For example, if a line crosses through a region (starting from the region's exterior, crossing through the interior and ending again in the region's exterior), then it is clearly perceived as such (a line crossing through *one* region). Deleting the two open segments of the line, however, will result in two patches that in turn will result in more than one region.

7.3.4.2 Refinement for Good Gestalt Measure and Identifying Regions with other Gestalt Laws

The PSI algorithm uses the notion of good continuity to identify a set of possible regions from a sketch and to describe a region's gestalt. The laws of organization also define other principles (e.g., regularity, symmetry, proximity, co-linearity, co-circularity, parallelism, closure, similarity, and simplicity) that can possibly be used for both, the set module and the gestalt module instead or in addition to continuity (Figure 7.3). The analysis of the shortcomings of the PSI algorithm showed that such an extension of the algorithm could lead to a better performance of the set module and of

91

the gestalt module. For example, because regular shapes were not identified or were not assigned a good gestalt, regularity could be of great value for this algorithm.

In order to successfully describe any of these principles, the geometry of a shape has to be described in a formal way. Three basic shape attributes represent the characteristic physicality of a shape: vertex *angle* (e.g., absolute value or right/acute/obtuse; always on inside or always on outside of region), relative *length* of edges (e.g., relative to stroke length; relative to sketch extend), *curvature* of a boundary segment (Park and Gero 1999).

Probably the most challenging task of using multiple set modules and gestalt modules is to decide which module to use for a particular spatial scene. An ontology-based approach could choose appropriate modules according to information given by the user about the nature of the sketched scene (e.g., if the sketch contains buildings then rectangles are considered to have a better gestalt than a circle, hence, regularity would be chosen. If the scene contains land parcels, then no overlapping regions are allowed).

Research on using laws of organization in computer vision can be found in Lowe (1990), Mohan and Nevatia (1992), Park and Gero (1999), Saund (2003), Saund and Moran (1995), Zabrodsky and Algom (1994), and Zhu (1999).
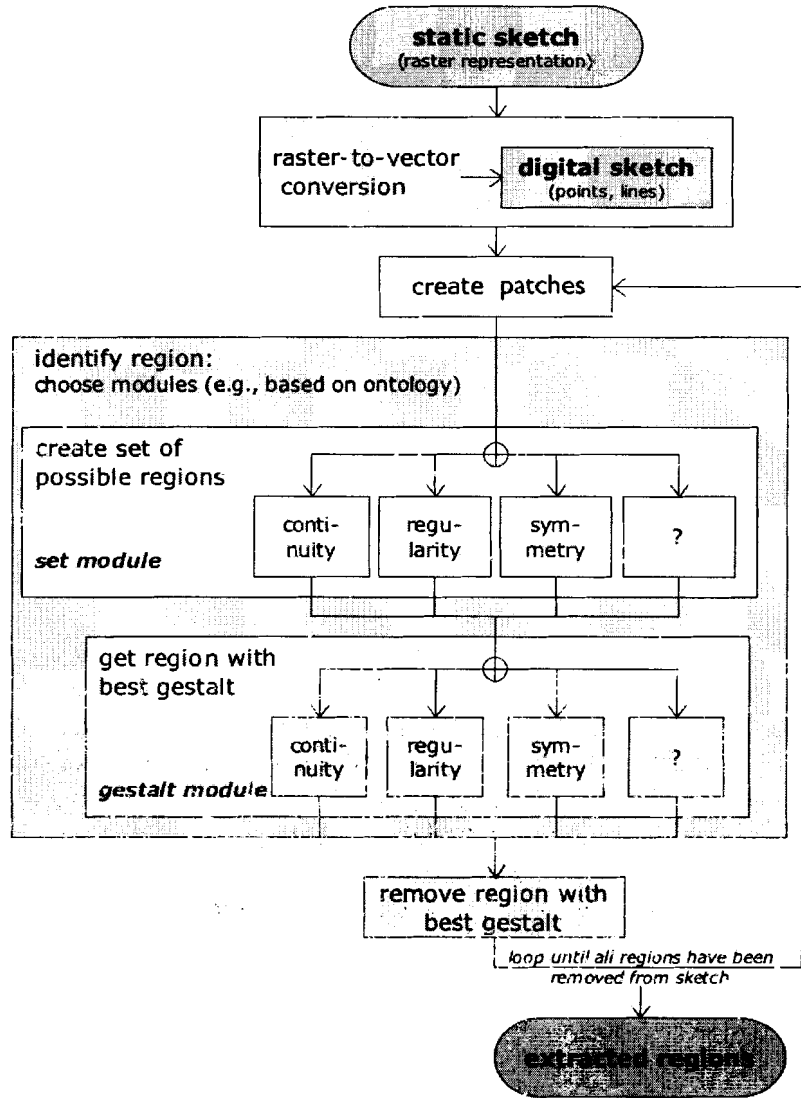
Figure 7.3. The algorithm of extracting regions from a sketch with additional set modules and gestalt modules.

93

BIBLIOGRAPHY

H. R. Arm (1997) *Wahrnehmung*. Rapperswil, Switzerland, Lecture Notes in Psychology and Sociology.

D. Ballard and C. Brown (1982) *Computer Vision*, Prentice Hall, Englewood Cliffs, NJ.

M. Bennamoun and G. J. Mamic (2002) *Object Recognition*, Springer-Verlag, London.

G. Birkhoff (1948) *Lattice Theory*. vol. 25. Revised Edition, American Mathematical Society, New York, NY.

A. Blake and M. Isard (1998) *Active Contours*, Springer-Verlag, London.

A. Blaser (2000a) *Sketching Spatial Queries*. PhD Thesis, University of Maine, Orono, ME.

A. Blaser (2000b) A Study of People's Sketching Habits in GIS. *Spatial Cognition and Computation* 2(4): 393-419.

P. Blosser (1973) Principles of Gestalt Psychology and their Application to Teaching Junior High School Science. *Science Education* 57: 43-53.

P. Bolstad (2002) *GIS Fundamentals*, Eider Press, St. White Bear Lake, MN.

K. Castelman (1996) *Digital Image Processing*, Prentice Hall, Englewood Cliffs, NJ.

N. Chrisman (2001) *Exploring Geographic Information Systems*. 2nd, John Wiley & Sons, Inc., New York, NY.

D. Clark (1999) Gestalt Theory - Instructional Technology Foundations and Theories of Learning. *Accessed:* 04.22.2003,

http://chd.gse.gmu.edu/immersion/knowledgebase/strategies/cognitivism/gestalt/gestalt2.htm.

P. Doucette (2002) *Automated Road Extraction From Aerial Imagery By Self-Organization*. PhD Thesis, University of Maine, Orono, ME.

M. Egenhofer and R. Franzosa (1991) Point-Set Topological Spatial Relations. *International Journal of Geographical Information Science* 5(2): 161-174.

M. Egenhofer and J. Herring (1991a) High-Level Spatial Data Structures for GIS. in: *Geographical Information Systems*, vol. 1, 227-237, M. Maguire, M. Goodchild, and D. Rhind, (Eds.), Longman, London.

M. Egenhofer and J. Herring (1991b) *Categorizing Binary Topological Relationships Between Regions, Lines, and Points in Geographic Databases*. Department of Surveying Engineering, University of Maine, Orono, ME.

M. Egenhofer and J. Sharma (1992) Topological Consistency. in: *Fifth International Symposium on Spatial Data Handling*, Charleston, SC, 335-343.

M. Egenhofer (1993) A Model for Detailed Binary Topological Relationships. *Geomatica* 47(3&4): 261-273.

M. Egenhofer and J. Sharma (1993) Assessing the Consistency of Complete and Incomplete Topological Information. *Geographical Systems* 1(1): 47-68.

M. Egenhofer, D. Mark, and J. Herring (1994) *The 9-Intersection: Formalism and its use for Natural-Language Spatial Predicates*. National Center for Geographic Information and Analysis 94-1.

95

M. Egenhofer and D. Mark (1995) Naive Geography. in: *Conference on Spatial Information Technology (COSIT '95)*, Semmering, Austria, Lecture Notes in Computer Science, 988, Springer-Verlag, 1-15, September 1995.

M. Egenhofer (1997) Query Processing in Spatial-Query-by-Sketch. *Journal of Visual Languages and Computing* 8(4): 403-424.

Environmental Systems Research Institute (1994) *Arc Commands*. Redlands, CA, User Manual.

Environmental Systems Research Institute (2001) *ArcScan and Image Integration*. Redlands, CA, User Manual.

ERDAS (1997) *ERDAS Imagine Tour Guides*. Atlanta, GA, User Manual.

A. Frank (1992) Spatial Concepts, Geometric Data Models, and Geometric Data Structures. *Computers and Geosciences* 18(4): 411-436.

G. Goldschmidt (1991) The Dialectics of Sketching. *Creativity Research Journal* 4(2): 123-143.

R. Haralick (1985) Survey: Image Segmentation. *Computer Vision, Graphics, Image Processing* 29: 100-132.

M. Kass, A. Witkin, and D. Terzopoulos (1988) Snakes: Active Contour Models. *International Journal of Computer Vision* 1(4): 321-331.

K. Koffka (1935) *Principles of Gestalt Psychology*, Harcourt, Brace and Company, New York.

D. G. Lowe (1990) Visual Recognition as Probabilistic Inference from Spatial Relations. in: *AI and Eye*, A. Blake and T. Troscianko, (Eds.), John Wiley & Sons Ltd., New York, NY.

K. Lynch (1960) *The Image of a City*, MIT Press, Cambridge, MA.

A. Mackworth (1977) Consistency in Networks of Relations. *Artificial Intelligence* 8: 99-118.

R. Mohan and R. Nevatia (1989) Using Perceptual Organization to Extract 3-D Structures. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 11(11): 1121-1139.

R. Mohan and R. Nevatia (1992) Perceptual Organization for Scene Segmentation and Description. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 14(6): 616-635.

S.-H. Park and J. S. Gero (1999) Qualitative Representation and Reasoning about Shapes. in: *Visual and Spatial Reasoning in Design*, 55-68, J. S. Gero and B. Tversky, (Eds.), Key Centre of Design Computing and Cognition, University of Sydney, Sydney, Australia.

B. Petermann (1932) *The Gestalt Theory and the Problem of Configuration*, Harcourt, Brace and Company, New York, NY.

W. Pratt (2001) *Digital Image Processing*. 3rd, John Wiley & Sons, Inc., New York, NY.

F. P. Preparate and R. T. Yeh (1973) *Introduction to Discrete Structures for Computer Science and Engineering*, Addison-Wesley, Reading, MA.

S. Sarkar and K. Boyer (1993) Integration, Inference, and Management of Spatial Information Using Bayesian Networks: Perceptual Organization. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 15(3). 256-274.

E. Saund and T. Moran (1995) Perceptual Organization in an Interactive Sketch Editing Application. in: *International Conference on Computer Vision (ICCV '95)*, Cambridge, Massachusetts, IEEE Computer Society Press, 597-604, June 1995.

E. Saund (2003) Finding Perceptually Closed Paths in Sketches and Drawings. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 25(4): 475-491.

R. Shariff (1996) *Natural Language Spatial Relations: Metric Refinements of Topological Properties*. PhD Thesis, University of Maine, Orono, ME.

M. Wertheimer (1923) Laws of Organization in Perceptual Forms. in: *A Source Book of Gestalt Psychology*, 71-88, W. Ellis, (Ed.) Routledge & Kegan Paul, London.

H. Zabrodsky and D. Algom (1994) Continuous Symmetry: A Model for Human Figural Perception. *Spatial Vision* 8(4): 455-467.

S.-C. Zhu (1999) Embedding Gestalt Laws in Markov Random Fields - a theory for shape modeling and perceptual organization. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 21(11): 1170-1187.

BIOGRAPHY OF THE AUTHOR

Markus Wuersch was born in Birmenstorf, Switzerland on June 24, 1973. He was raised in Birmenstorf, Switzerland. From 1989 to 1994 he worked as a land surveyor with H. Heri, Surveying and GIS, Baden, Switzerland and earned a degree in surveying from the Vocational College, Zurich in 1993. From the summer of 1994 to the summer of 1995, Markus Wuersch traveled with Up With People's Worldsmart-Program. In 1996 he began studies for a BS degree in urban and regional planning at the University of Applied Sciences, Rapperswil, Switzerland. During one year within his studies, Markus Wuersch worked as an urban planner intern with Remund + Kuster, Pfäffikon and as a transportation engineer intern with Roland Müller, Küsnacht, Switzerland. In 2000 he graduated from the University of Applied Sciences, Rapperswil and was awarded a stipend from the Leica Fond for graduate studies abroad.

In 2001 he worked as a Survey Crew Chief and Survey Technician with the Sewall Company in Old Town, Maine before beginning studies leading to a MS at the University of Maine, Orono, USA. From fall 2001 he worked as a research assistant in the department of Spatial Information Science and Engineering at the University of Maine. Markus is a candidate for the Master of Science degree in Spatial Information Science and Engineering from The University of Maine in December, 2003.