Perceptual Sketch Interpretation

Markus Wuersch[1], Max J. Egenhofer[2]

[1] uLocate Communications Inc., 60 Canal St., Boston, MA 02114, USA
e-mail: markus@wuersch.net

[2] National Center for Geographic Information and Analysis, Department of Spatial Information Science and Engineering, Department of Computer Science, University of Maine, Orono, ME 04469-5711, USA
e-mail: max@spatial.maine.edu

Abstract

An automated extraction of regions from sketches can be of great value for multi-modal user interfaces and for interpreting spatial data. This paper develops the *Perceptual Sketch Interpretation* algorithm, which employs the theory of topological relations from spatial reasoning as well as good continuity from gestalt theory in order to model people's perception. The Perceptual Sketch Interpretation algorithm extracts regions iteratively, removing one region at each a time, thus making the remaining sketch simpler and easier to interpret. The evaluation of the algorithm shows that the use of gestalt theory empowers the algorithm to correctly identify regions and saves processing time over other approaches.

1 Introduction

Spatial data are being collected constantly and in large amounts. Interpreting these data poses a shear never-ending task of gaining information from raw data. This task, when dealing with spatial data, relies heavily on feature extraction. Once information can be quickly extracted from spatial data, spatial analysis can be performed based on the resulting information. The assistance of spatial analysis and its results greatly supports society in many challenging tasks and endeavors, such as emergency management, resource management, economic impact studies, and health risk assessment. Automated feature extraction is, therefore, of great importance when dealing with spatial data.

This paper defines a perceptual feature extraction algorithm to successfully identify regions in a sketch, a particularly challenging task for interacting with and using visual information. The goal is to obtain from a visual presentation (i.e., a sketch) exactly what people perceive in such a sketch. This process is called *perceptual sketch interpretation*. The success of feature extraction methods depends largely on the scope of the geometric objects that may be handled (Bennamoun and Mamic 2002). The scope of the PSI algorithm is limited to simple regions in a sketch. Sketches that describe a highly patterned texture (e.g., checkerboard) are outside the scope of this work, because reliably identifying regions in such cases is impossible without additional knowledge.

Automatic extraction of features from a sketch that was originally drawn on paper has to address analog-to-digital conversion in order to execute the feature extraction algorithm. Converting a paper sketch into a digital environment is possible through scanning, edge detection, and vectorization, for which plenty commercial tools are available; therefore, this work only addresses the task of extracting features from vectorized representations. It is also assumed that during the analog-to-digital conversion, the resulting vector representation of a sketch is topologically cleaned (i.e., removing overshoots, undershoots, and slivers).

There is a discrepancy between the elements contained in a sketch and the elements that people perceive. In a sketch, only lines are explicitly present, while regions are perceived by grouping together lines that form closed loops. 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 complex to be formalized so that a machine could carry out that task automatically and reliably. 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 (Bennamoun and Mamic 2002). This paper describes a perceptually supported algorithm for extracting regions from a sketch without prior knowledge about drawing sequences and without interactive human-computer interaction. The sketch in Figure 1 will be used throughout this paper as a running example to illustrate the steps of the region extraction algorithm.

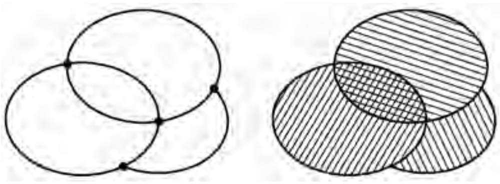


Fig. 1. A sample sketch: (a) the original sketch-nodes highlight the intersections of drawn lines and (b) the identified regions.

Solving feature extraction from sketches will be of great value for research in several domains that deal with visual data, such as computer vision and feature extraction from satellite images or aerial photographs, as well as multi-modal user interfaces, such as spatial-query-by-sketch (Egenhofer 1996), if static sketches *in lieu* of real-time sketches are used as queries.

The remainder of this paper reviews related work in Section 2. Underlying principles from gestalt theory and spatial reasoning are summarized in Section 3. Section 4 introduces the perceptual sketch interpretation algorithm, followed by the description of a prototype system (Section 5) and an evaluation with 24 sketches, and their intended meanings, collected from human subjects (Section 6). Section 7 draws conclusions and suggests future work.

2 Related work

Feature extraction involves image processing, computer vision, and machine intelligence. The most relevant work is briefly reviewed in this section.

Saund (2003) uses a maximal turning path and smooth continuation between lines to identify closed or nearly closed regions. The identified figures are either accepted or rejected based on a measure for a good gestalt. This perceptually closed path finding algorithm, however, requires prior domain knowledge for a successful interpretation.

PerSketch is a perceptually supported sketch editor (Saund and Moran 1995), which offers users suggestions when editing an object in a sketch. In doing so, PerSketch tries to read the users' mind. The algorithm picks objects based on geometric properties (e.g., closure, parallelism, corners, and T-junctions). Research on building these rules can be found in the computer-vision literature (Mohan and Nevatia 1989, Sarkar and Boyer 1993).

CANC2 (Mohan and Nevatia 1992) is a computer vision system that identifies object edges from a vectorized image. The set of vectorized edges of an image is reduced to object edges by applying gestalt laws (e.g., proximity, continuity, symmetry, closure, and familiarity), thus eliminating noise. Identified edges are grouped into non-overlapping object surfaces.

The use of sketching as human-computer interaction mode is explored in *Sketching Spatial Queries* (Blaser 2000), which aims at building a spatial query from a sketch input. The query processor computes similarity values to any other sketch and returns sketches that are similar to the input sketch. Similarity between sketches is computed based on completeness, geometry, topology, metric, and directions of objects and topological relations. Sketching is also used in multi-modal interfaces. Oviatt et al. (1997) use pen input to convey location information. Likewise, Quickset (Cohen 1997) links spoken and pen input.

Based on Wuersch's (2003) use of gestalt principles to extract regions from sketches, Waranusast (2007) forms regions from sketches drawn on PDAs. Such sketches also provide the temporal information about the drawing, offering further heuristics to extract regions successfully.

3 Underlying principles

The region-extraction algorithm developed in this paper is based on the theory of *topological relations* and on *gestalt theory*. Both are briefly explained and complemented with refinements and formal definitions.

3.1 Topological Relations in Sketches

One of the main objectives that spatial reasoning can serve is to change a spatial representation into a different format (translation and interpretation) (Vieu 1997). In the case of interpreting a sketch, this means to identify spatial information of a sketch's elements and to use this information to form new objects. Topology is a most critical part in identifying significant spatial information (Egenhofer and Mark 1995, Kuipers 1979). The 9-intersection (Egenhofer and Herring 1991) provides a framework for identifying formally binary topological relations. It distinguishes eight topological relations between two 2-discs and 33 topological relations between two lines in R^2. This paper considers topological relations between two regions, between two patches (regions in a partitioned space), and between two simple lines.

Regions are homeomorphic to 2-discs and, therefore, two regions in a sketch can have any of the eight possible topological relations. When identifying regions, however, it is impossible to distinguish two regions with the topological relation *equal*.

Partitions are defined as subdivisions of space that consist of cells in the most general case, where any two distinct cells do not have a common interior (Egenhofer and Herring 1991). Patches are not true partitions, because each hole is treated as a separate patch, inside or contained by another patch. Topological relations between two patches (Figure 2) are limited to disjoint, meet, covers, coveredBy, inside, and contains, as well as the dimensional refinements of meet, covers, and coveredBy, referring to the dimension of the shared boundaries of two patches, that is, 0-meet, 1-meet, 0-covers, and 0-coveredBy (Egenhofer 1993). Any two regions with the relation 1-covers or 1-coveredBy form two patches with the relation meet. Two overlapping regions form three or more patches with the topological relation meet between each pair of adjacent patches. Alike equal relations between two regions, this relation is not detectable between two patches.

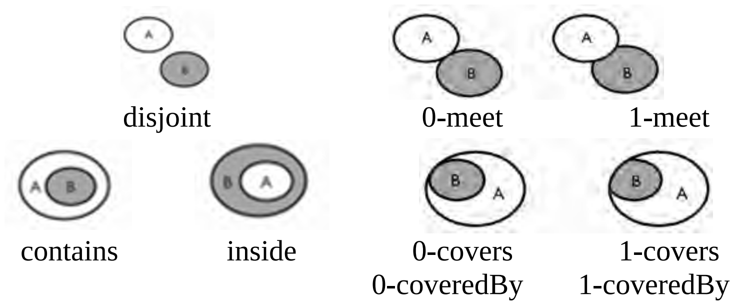
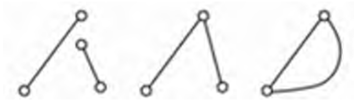


Fig. 2. Possible topological relations between two patches in a sketch.

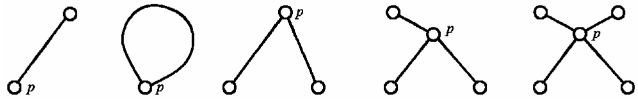
The extraction of regions from sketches requires a topologically clean sketch (i.e., no crossing lines, overshoots, undershoots, or slivers), reducing the set of possible binary topological relations between two lines from thirty-three (Egenhofer 1993) to three (Figure 3). The distinction of meet-once (Figure 3b) and meet-twice (Figure 3c) arises from the number of non-empty intersections between the boundaries of the lines.



(a) (b) (c)

Fig. 3. Possible topological relations between two lines in a sketch: (a) disjoint, (b) meet-once, and (c) meet-twice.

The node degree—that is the number of incoming and outgoing lines at each end of a line —yields a further refinement of the meet relation, which is expressed as m1, m2, m3, m4, etc. (Figure 4).



m1 m2 m2 m3 m4

Fig. 4. Intersection types of lines with metric information about the number of incoming and outgoing lines at point *p*.

3.2 Continuity and Good Gestalt

The law of good continuity and the notion of good gestalt from gestalt theory (Koffka 1935, Wertheimer 1923) are of great importance for grouping lines in a sketch, as these gestalt properties often describe people's perception. Gestalt theory, however, only provides a descriptive theory, but not specific computational processes (Zhu 1999). These theories are briefly explained and complemented with a formal definition.

The law of *good continuity* states that two lines are more likely to be grouped together if one line is perceived as the continuation of the other. In this paper, continuity is expressed by the angle γ formed by two lines, *a* and *b*, that meet (Figure 5a). This angle is then compared to a threshold resulting in either continuity or discontinuity. In cases where more than two lines meet, we rely on the symmetric property of continuity to find the best continuity. In doing so, the continuity angle γ is examined from both directions, that is, from *a* with *b* as the continuing line and from *b*, with *a* as the continuing line (Figure 5b).

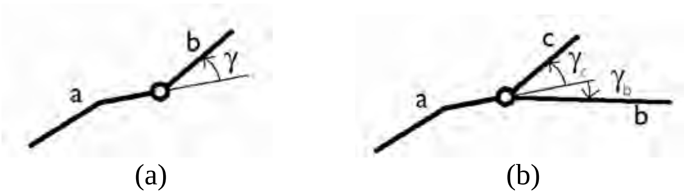


Fig. 5. Continuity for (a) two lines and (b) three lines.

In gestalt theory, the law of *pragnanz* defines 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, which refers to the simplest, most stable figure possible (Zabrodsky and Algom 1994, Zhu 1999). When describing a good gestalt people use such properties as continuity, regularity, and symmetry. Here we use the continuity property to evaluate a good gestalt of a region. 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.

4 The perceptual sketch interpretation algorithm

The Perceptual Sketch Interpretation (PSI) algorithm cycles through the following three steps: (1) identifying patches, (2) identifying regions, and (3) extracting and removing the region with the best gestalt. These steps are repeated until all the regions are identified, that is, when no patches are left in the sketch. By iteratively removing any identified region, the remaining sketch becomes less and less complex to interpret.

The PSI 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 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, the set of identified regions 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 a region of people's mental model of the same sketch.

4.1 Identifying Patches

Geographic information must be embedded in a reference system for time, space, and attribute (Chrisman 2001). Feature extraction from sketches, however, can only make use of information about space. Based on information about space, a tracking algorithm that traces along lines and continues consistently in the same direction when reaching an intersection point (e.g., always turn left, or always turn right) identifies boundaries of patches. Patches in a sketch can be used as building blocks for any region. Such a region is built as the union of two or more patches (e.g., two overlapping regions are interpreted as three patches) or a patch is itself a region.

4.2 Identifying Regions

To identify regions in a sketch means to group and union the patches into the region they form. A first step in identifying regions is to extract regions that are formed by only one patch. Such cases correspond to scenarios with a topological relation other than 1-meet (i.e., disjoint patches and patches that do not share any boundary segment with any other patch).

The law of good continuity allows the algorithm to identify regions in a sketch based on Assumption 1: 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. 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 a good continuation are combined to build a new region. This process is repeated until a closed boundary is found or no further continuous boundary lines can be found. At this point, the patches used so far are combined to form a region. By repeating this task for any line in the sketch, a set of regions is created that are candidates to be extracted. This set is generally much smaller than the set of all possible regions in a sketch and does not necessarily contain all the regions to be extracted.

An identified region has to satisfy a set of conditions to be a valid region. First, the interior of the region has to be connected, which leads to the conclusion that only patches that share a boundary segment (i.e., the patches have a 1-meet topological relation) can be identified to form new regions (Figure 6a and 6b). Second, the region that is formed as the union of two patches has to contain the two lines that formed the good continuity. In order to satisfy this constraint, both lines cannot be the shared boundary between the two patches, as the shared boundary is not contained in the union of the two patches (Figure 6c and 6d). Third, the resulting new region must be simple, that is, it has no holes, separations, or spikes.

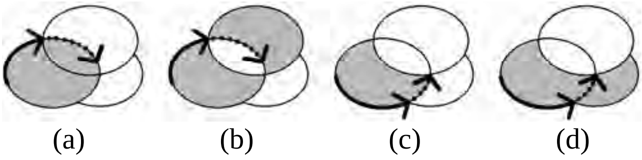


Fig. 6. An example sketch with a starting line (solid arrow) and a continuing line (dotted arrow); (a, c) conditions are satisfied, (b, d) not satisfied.

At the intersection of three lines (i.e., an m3-intersection), there are two possibilities on how the PSI algorithm should proceed if no continuing lines are found. One could argue that the next line in the current patch should serve as the continuing line, because such intersections occur likely where two patches meet (Figure 7). Alternatively, the PSI algorithm simply stops and proceeds to the next patch. The latter approach follows the idea of finding a continuous boundary and, therefore, this approach is chosen for the evaluation of the PSI algorithm.

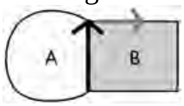
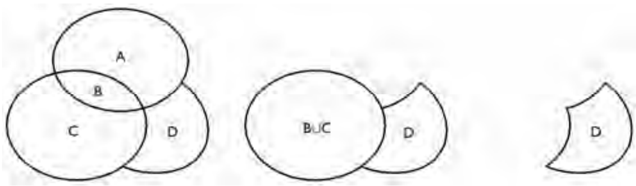


Fig. 7. Continuing at m3-intersection when no continuous boundary is found as an alternative to stopping.

4.3 Extracting Regions with Best Gestalt

Removing at each iteration of the algorithm only the region with the best gestalt value 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. Iteratively removing an identified region from the sketch is crucial for a successful interpretation of all the regions in a sketch (Figure 8).



(a) (b) (c)

Fig. 8. A sample sketch: (a) the original sketch, (b) after one iteration, and (c) after two iterations.

A region is removed from a sketch by first removing its boundary. In some cases, however, only some parts of the regions boundary can be removed, because its remaining parts are still used for building other patches. In order to outline a rationale on deciding what boundary segments can be safely 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 1).

Table 1. Classification of line types

Classification	Intersection Type		Patches
	End 1	End 2	
A	2	2	1
B	3	3	1
C	3	3	2
D	3	4+	1
E	3	4+	2
F	4+	4+	1
G	4+	4+	2

For any line type, except for types A and C, two representations can be found, one where the specific line can be removed and another one where the line cannot be removed, making it necessary to define a rationale whether or not to remove the line. Further analysis shows that lines of type A should not be present in a sketch at this point of the PSI algorithm. Since these lines form a closed loop that was identified as a region, the lines were already removed from a sketch. Lines of type C are the common boundary segments of two patches. Removing such a line when removing the boundary of one of the patches always results in a semi-open set in the sketch and, therefore, lines of type C are kept in the sketch in any case.

For the remaining line types B, D, E, F, and G it is uncertain whether or not to remove the specific line. The difference between cases where the line can be removed and cases where the line cannot be removed is in the number of regions that the line is part of. When 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 is not available before the extraction algorithm has completed 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 (Figure 9). In that case, one or more lines that close the semi-open set have to be brought back into the sketch.

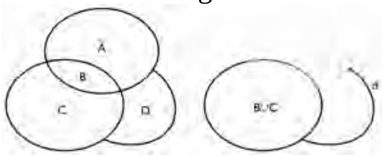
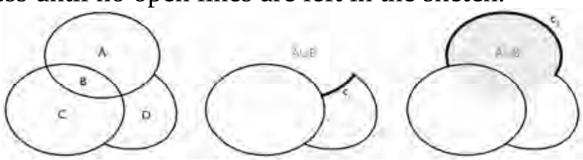


Fig. 9. After removing the region A∪B an open line d is left in the sketch.

During this process, more than one line, called the *closing line,* can be brought back into a sketch. Any such line must be part of the removed region's boundary, it must meet one or more open lines at their open end: it must not be an open line itself, it must be contained in at least one patch, and bringing back the closing line should not introduce any complex object to the sketch. If there is more than one line that meets these constraints, the line that reintroduces the least number of elements of the removed region is chosen. For example, if several lines fulfill 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 introduces back a part of the removed region's boundary and a part of its interior. In this case, the first closing line is chosen, because it does not bring back any part of the region's interior. Finding the closing lines of a sketch after removing a region is an iterative process until no open lines are left in the sketch.



(a) (b) (c)

Fig. 10. The closing line: (a) a sketch with patches A-D; (b) and (c) the same sketch with the region A∪B removed. In case of (b) the closing line c_1 only brings back a part of the boundary of (A∪B) whereas in (c) the closing line c_2 also brings back a part of the interior of A∪B into the sketch.

The PSI algorithm uses a minimum and a maximum continuity threshold for identifying continuous lines. At first, the minimum threshold is set low (e.g., 10 degrees) in an attempt to identify regions with a high gestalt value. Only if no regions are identified the continuity threshold is increased until the maximum threshold is reached or until there are no patches left in the sketch.

In cases where the PSI algorithm finishes with patches left in the sketch, these remaining patches are added to the set of extracted regions in order to complete the spatial scene.

It is possible that a patch is lost after a region is removed from the sketch. Whereas the extracted regions will not cover the same space as the original sketch, the extracted regions might match with people's mental model. In this case, an option is given whether or not to fill gaps at the end of the region extraction process. In order to illustrate such a case, two partitions have been added to the sample sketch (Figure 11).

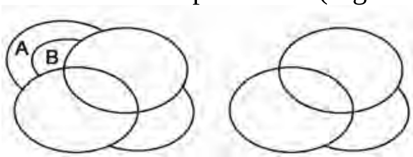


Fig. 11. Patch B is lost after removing region A.

5 Prototype

The PSI algorithm (Figure 12) was implemented in a prototype application that serves as a test bed for the model evaluation. It extracts features from a digital sketch. Any pre-processing of such a line drawing (i.e., scanning, raster-to-vector conversion, and cleaning topology) was completed with commercial hardware and software.

The prototype uses a map metaphor for displaying the sketch and allows a user to interact through a WIMP (windows, icons, menus, pointers) interface. A preference pane lets users adjust any setting used in the feature extraction process, such as continuity thresholds. Supported file formats are a text file containing a list of points grouped by line numbers and ESRI's interchange format (e00). Upon opening a sketch, three different views of the sketch are displayed: the original sketch with patches, the processed sketch containing regions (Figure 13), and a process view displaying feature extraction process at different stages. These visualizations help with analyzing possible errors in case the algorithm commits any misinterpretations. The interpreted sketch can be saved in the Spatial-Query-by-Sketch format (Blaser 2000), enabling a subsequent spatial query that can be executed on a set of other sketches.

6. Evaluation

The PSI algorithm was evaluated for correctness and compared to an alternative approach where, instead of using continuity, every possible region in a sketch is analyzed to identify regions. The result of the evaluation shows a significant advantage in efficiency using continuity to identify regions over that alternative approach. In addition, the PSI algorithm correctly interpreted 75% of the analyzed sketches. The comparison to this approach shows the advantage in processing load when the notion of good continuation is used.

```
    function PSI (sketch): sketch
    newSketch := empty sketch;
    newRegions := empty list of regions;
    continuityThreshold := minThreshold;
    loop
      remove from sketch patches that are not 1-meet to any other patch and
      add them to newSketch;
      find set of all possible regions in sketch;
      or find set of possible regions using continuity:
          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 newRegions is empty: increment continuityThreshold;
      else
          remove region from newRegions with best gestalt
              and add removed region to newSketch;
          build patches with remaining lines in sketch;
      end if;
    loop until patches is empty
          or continuityThreshold > maxThreshold;
    add unused patches to newSketch;
return newSketch;
```

Fig. 12. Pseudo code of the PSI algorithm.

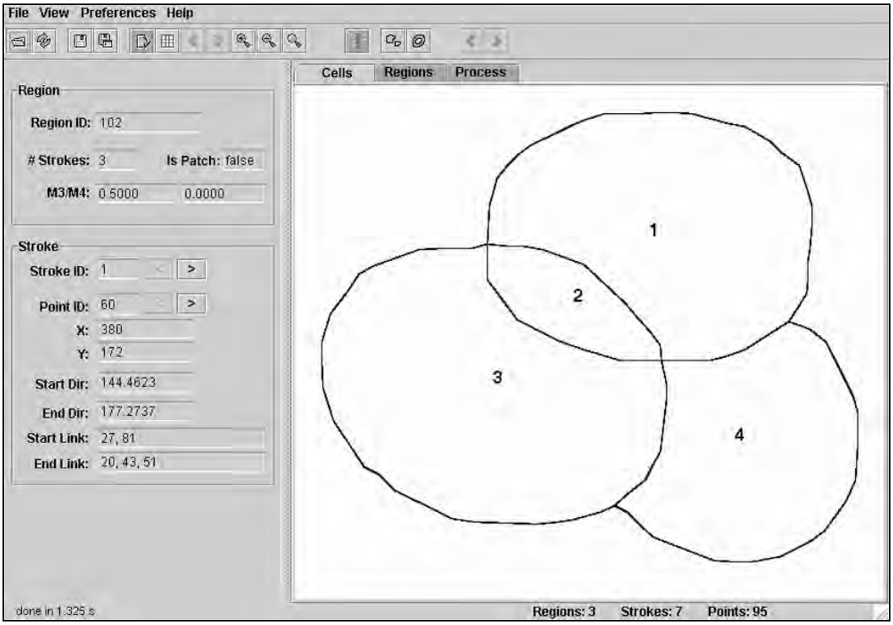


Fig. 13. The application window showing the processed sketch in the region tab. A region is highlighted and the corresponding attributes of that region are displayed in the sketch properties panel on the left.

6.1 Evaluation Design

In order to objectively evaluate the PSI algorithm, a set of sketches were obtained from people who were not involved in the design of the PSI algorithm. For this purpose, a web-based survey was conducted, giving the participating subjects the opportunity to draw and submit their sketch through a Web browser interface. Participants were also asked to submit their interpretation of the sketch. The collected information, therefore, contained 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). In total, 36 sketches were collected of which 24 did not contain any complex regions or tessellations and were selected for the evaluation.

6.2 Correctness

For each collected sketch, regions were extracted manually according to the description obtained from the online survey. The resulting spatial scene was termed *ground truth* and used to evaluate the correctness of the PSI algorithm's results. This ground truth was compared to a spatial scene identified by the PSI algorithm. The prototype for Spatial-Query-by-Sketch (Blaser 2000) was used to determine similarity values between the two sketches. The ground truth was used as a query input, operating on the interpreted sketch created by the PSI prototype. Spatial-Query-by-Sketch returns similarity values (0% to 100%), which 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 (100% is a theoretical value and because of rounding errors the actual received similarity value for identical sketches were 99.9%). From the 24 analyzed sketches, 18 (i.e., 75%) were interpreted correctly (Table 2).

6.3 Advantage of the Continuity Approach

The PSI algorithm uses the notion of good continuation to identify what patches should be combined to form regions with the best possible continuous boundary. An alternative to this approach would analyze all possible regions in a sketch.

Test sketches were processed again using all possible regions in a sketch. First, the results of this process are compared to the ground truth and the results from the extraction process of the continuity-based approach. This approach using all regions produced 15 correct interpretations (i.e., 62.5%), a less accurate result than the continuity-based approach—the PSI algorithm performed on these samples 12.5% better in absolute numbers, and 20% better with respect to the success rate of all-regions approach (Table 2).

Table 2. Correctness of the PSI algorithm

	Using continuity	Using all regions
Correctly interpreted sketches	75%	62.5%
Average similarity	93.9%	92.4%
Smallest similarity	37.6%	46.4%
Largest similarity	99.9%	99.9%

The analysis of the processing times of both approaches clearly shows the advantage of the continuity-based approach as it executed on average 119 times faster than when analyzing all possible regions in a sketch (Table 3). This difference in processing time is due to the often very large numbers of possible regions that can be extracted from a sketch.

Table 3. The processing load of the PSI prototype (processing was done on a computer running Windows XP, 2.99 GHz processor, and with 1 GB of RAM)

	Using continuity	Using all regions
Average processing time [sec]	0.2	26.2
Average number of regions analyzed	11	2,056

6.4 Shortcomings

The six sketches that were incorrectly interpreted using the continuity approach were 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. In two cases, the rationale for removing an identified region's boundary from the sketch returned incorrect results, while in five cases continuity was not the major factor used by people's perception to order visual input so that regions were identified that should not have been extracted. Regions with a regular shape (e.g., squares and rectangles) were not identified or were not classified as having a good gestalt. The set of identified regions, however, contained at least one region to be extracted, thus Assumption 2 (Section 4) holds.

7 Conclusions and future work

We developed an algorithm to extract features from sketches. This algorithm makes use of the law of good continuity and the notion of a good gestalt to identify regions and to rank these regions by their gestalt value. A prototype implementation was used to evaluate the PSI model from which we can draw conclusions and suggest future work.

7.1 Results

The results of the PSI algorithm's evaluation lead to three major conclusions:

- By using continuity to identify a set of regions from patches, the resulting set contains at least one region that corresponds to a region in people's mental model. This conclusion confirms Assumption 2 (Section 4), supported by the results of the model assessment. Because this assumption clearly holds for the approach of analyzing all possible regions and because the results of the assessment have shown evidence that using continuity produces equal or better results, it can be concluded that the assumption also holds for the continuity approach. This conclusion is also supported by the analysis of the incorrect interpreted sketches, where other reasons were identified as the cause of incorrect results.

- Good continuity is, amongst other gestalt laws, one of the major factors used by people's perception to order visual input into meaningful objects. This conclusion is drawn because the PSI algorithm has correctly interpreted 60% to 75% of the test sketches. Sketches that were misinterpreted, however, ask for additional reasoning other than the notion of good continuity. This conclusion refers to Assumption 1, but also applies to Assumption 3 (Section 4), which 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 continuity approach to identify regions is preferred over analyzing all possible regions. The analysis of the number of correctly interpreted sketches and the analysis of the processing times strongly support this conclusion.

7.2 Future Work

The model assessment indicates possible future research topics as well as refinements and extension to the current model. Further analysis using the PSI prototype could be performed using a variety of algorithm preferences. Such analysis could show correlations between scene characteristics and distinct preferences of the algorithm. Where such correlations exist, the PSI algorithm could be tailored towards different types of sketches. Additional analysis of the algorithm's settings could also reveal which settings are the most relevant and would allow us to aim future work at the most important parts of the algorithm.

The PSI algorithm was tested on a set of sample sketches that were hand drawn. Further valuable results could be gained by evaluating the PSI algorithm on data other than sketches. Such data could be vectorized aerial photographs, satellite imagery, or any other data type that can be transformed into a vector representation.

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 work. If a scene is to be analyzed on a more detailed level, however, metric aspects as well as direction information become relevant. Incorporating such refinements for the topological relations would possibly result in more accurate sketch interpretations. It would also allow for a purely automated process, as it would rely less on generating a clean topology of the scanned sketch.

Drawing errors, such as overshoots, undershoots and slivers, are corrected for by a cleaning function before the actual region extraction process commences. In doing so, some information that could reveal more details about the possible regions in a sketch might be compromised. For example, slivers indicate that a line was drawn twice. In cases where drawing errors occur, they could give better insights on regions in a sketch thus improving the result of the region extraction. In the example of a line drawn twice, the PSI algorithm could use this information to make sure that the line is used in two different regions.

The analysis of the shortcomings of the PSI algorithm has shown that the rationale of removing an identified region's boundary, outlined in section 0, could not be relied on at all times. The result of the PSI algorithm could be improved by refining this rationale.

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 instead or in addition to continuity. 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). 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 PSI algorithm. For example, because regular shapes were not identified or were not assigned a good gestalt, regularity could be of great value for this algorithm.

While these recommendations for future work show room for improvement of the PSI algorithm, the algorithm showed convincing results supporting the perceptual approach of interpreting sketches.

Acknowledgments

This work was partially supported by the National Geospatial-Intelligence Agency under grant numbers NMA201-01-1-2003.

References

Bennamoun M, Mamic G (2002). Object recognition. Springer-Verlag, London

Blake A, Isard M (1998). Active contours. Springer-Verlag, London

Blaser A (2000). Sketching spatial queries. Ph.D. Thesis, University of Maine, Orono, ME

Blaser A, Egenhofer M (2000). A visual tool for querying geographic databases. In Di Gesù V, Levialdi S, Tarantini L (eds), AVI 2000—Advanced visual databases, Salerno, Italy, pp 211-216

Chrisman N (2001) Exploring geographic information systems. John Wiley, New York

Cohen P, Johnston M, McGee D, Oviatt S, Pittman J, Smith I, Chen L, Clow J (1997). Quickset: multimodal interaction for distributed applications. Proceedings of the fifth ACM international multimedia conference, pp 31-40

Egenhofer M (1993). A model for detailed binary topological relationships, *Geomatica*, 47(3&4), 261-273

Egenhofer M (1993). Definitions of line-line relations for geographic databases. *IEEE data engineering bulletin* 16(3), 40-45

Egenhofer M (1996). Spatial-Query-by-Sketch. In Burnett M and Citrin W (eds) VL '96: IEEE symposium on visual languages, Boulder, CO, 60-67

Egenhofer M (1997). Query processing in Spatial-Query-by-Sketch. Journal of visual languages and computing 8(4): 403-424

Egenhofer M, Herring J (1991). Categorizing binary topological relationships between regions, lines, and points in geographic databases. Technical Report, Department of Surveying Engineering, University of Maine, Orono, ME, (http://www.spatial.maine.edu/~max/9intreport.pdf)

Egenhofer M, Mark D (1995). Naive geography, In: Frank A, Kuhn W (eds), COSIT '95, Spatial information theory. Lecture Notes in Computer Science vol 988, pp 1-16

Egenhofer M, Shariff AR (1998). Metric details for natural-language spatial relations, ACM transactions on information systems 16(4): 295-321

Koffka, K (1935). Principles of gestalt psychology, Harcourt, Brace and Company, New York

Kuipers B (1979) Modeling spatial knowledge. Cognitive science 2(2): 129-153

Lowe D (1990) Visual recognition as probabilistic inference from spatial relations, In Blake A, Troscianko T (eds), AI and eye, John Wiley, New York

Mohan R, Nevatia R (1989). Using perceptual organization to extract 3-d structures, IEEE transactions on pattern analysis and machine intelligence 11(11): 1121-1139

Mohan R, Nevatia R (1992) Perceptual organization for scene segmentation and description, IEEE transactions on pattern analysis and machine intelligence 14(6): 616-635

Oviatt S, DeAngeli A, Kuhn K (1997) Integration and synchronization of input modes during multimodal human-computer interaction. Proceedings of the conference on human factors in computing systems (CHI '97), pp 415-422

Park S-H, Gero J (1999) Qualitative representation and reasoning about shapes, In Gero J, Tversky B (eds.), Visual and spatial reasoning in design, Key Centre of Design Computing and Cognition, University of Sydney, Sydney, Australia, pp 55-68

Sarkar S, Boyer K (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

Saund E (2003) Finding perceptually closed paths in sketches and drawings. IEEE transactions on pattern analysis and machine intelligence 25(4): 475-491

Saund E, Moran T (1995) Perceptual organization in an interactive sketch editing application. International conference on computer vision (ICCV '95), IEEE Computer Society Press, pp 597-604

Shariff AR, Egenhofer M, Mark D (1998) Natural-language spatial relations between linear and areal objects: the topology and metric of English-language terms, International journal of geographical information science 12(3): 215-246

Vieu L (1997) Spatial representation and reasoning in artificial intelligence, in Stock, O. (ed.) Spatial and Temporal Reasoning, Kluwer, Dordrecht, pp 5-41

Waranusast R (2007) Perceptual-based region extraction from hand drawn sketches. Proceedings of the third IASTED international conference, advances in computer science and technology, Phuket, Thailand, pp 222-227

Wertheimer M (1923) Laws of organization in perceptual forms, In Ellis W (ed.), A source book of gestalt psychology, Routledge & Kegan Paul, London, pp 71-88

Wuersch M (2003) Perceptual sketch interpretation. M.S. thesis, University of Maine

Zabrodsky H, Algom D (1994) Continuous symmetry: a model for human figural perception. Spatial vision, 8(4): 455-467

Zhu S-C (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