

# Qualitative Analysis of Sketched Route Maps: Translating a Sketch into Linguistic Descriptions

Marjorie Skubic, Sam Blisard, Craig Bailey, Julie Adams and Pascal Matsakis

**Abstract**—In this paper, we introduce our work on sketch understanding, focusing here on the analysis of a sketched route map. A route map is drawn to help someone navigate along a path for the purpose of reaching a goal. A hand-sketched route map does not generally contain complete map information and is not necessarily drawn to scale, but yet it contains the correct qualitative information for route navigation. Here we propose a methodology for extracting a qualitative model of a sketched route map, based on human navigation strategies, using spatial relationships. Linguistic descriptions are generated from the sketch, both in the form of detailed descriptions at discrete path steps and also as a high-level route description. To describe the path linguistically, one must first be able to understand the path in a qualitative sense. We assert that the translation of a sketch into linguistic descriptions illustrates that the essential qualitative path knowledge has been extracted. The methodology is demonstrated using example sketches drawn on a handheld PDA.

**Index Terms**—sketch, route map, spatial relations, force histograms, linguistic descriptions

## I. INTRODUCTION

A sketched route map is drawn to help someone navigate along a path for the purpose of reaching a goal. Although route maps do not generally contain complete map information about a region and they are not always drawn to scale, they do provide relevant information for the navigation task. People sketch route maps to include landmarks at key points along the path and use spatial relationships to help depict the route [1]. The depiction of the environment structure is not necessarily accurate and may even distort the actual configuration [2]. For example, a 60 degree turn in the physical environment may be sketched as a 90 degree turn. However, as the route follower navigates in the real environment, his motion is constrained by the environment so that the distortion is corrected and the route can be completed [2]. Indeed, in a study of 29 sketched route maps, each contained the information necessary to complete a complicated navigation task [1].

Manuscript received May 28, 2002. This work was supported by the U.S. Office of Naval Research under Grant N00014-96-0439 and the Naval Research Laboratory under Grant N00173-01-1-G023F.

M. Skubic, S. Blisard, and C. Bailey are with the Dept. of Computer Engineering and Computer Science, University of Missouri, Columbia, MO 65211 USA. (phone: 573-446-7766; e-mail: skubicm@missouri.edu).

J. Adams is with the Dept. of Electrical Engineering and Computer Science, Vanderbilt University, Nashville, TN 37235 USA. (e-mail: julie.a.adams@vanderbilt.edu)

P. Matsakis is with Dept. of Computing and Information Science, University of Guelph, Ontario, Canada. (e-mail: matsakis@cis.uoguelph.ca)

Research by Michon and Denis [3] provides insights into how landmarks are used for human navigation and what are considered to be key route points. In studying route directions, they found that landmarks were used more frequently at four types of critical nodes: (1) at the start, (2) at the end, (3) at a change in orientation, and (4) at major intersections where errors could easily occur. Thus, people use the relative position of landmarks as cues to keep on track and to determine when to turn left or right.

Tversky and Lee collected and analyzed both sketched route maps and route descriptions and found the structure to be effectively the same [4]. The route is depicted both diagrammatically and verbally as a sequence of steps which are segmented by a landmark. Each step can be designated as a triple with an orientation, an action, and a landmark. Given that sketched routes and route descriptions have the same structure, it should be possible to convert a sketch into linguistic descriptions. Furthermore, such a translation would illustrate that the essential qualitative route information has been extracted from the sketch.

In our ongoing work on spatial modeling, we have been investigating the analysis of hand-drawn route maps, in which the user sketches an approximate representation of the environment and then sketches the desired path with respect to that environment. The objective is to extract spatial information about the map and a qualitative path through the landmarks drawn on the sketch. This information is used to build a path representation and then to generate a linguistic description of the sketched route. Note that the representation and the description are based *not* on absolute position, but rather on positions *relative* to landmarks in the environment.

Previous work has been done in using sketches to represent geographic information. The strategy of using a sketch with spatial relations was proposed by Egenhofer as a means of querying a geographic database [5]. Igarashi et al. proposed a path drawing technique overlaid on a virtual scene, as a means of specifying a route through the virtual environment [6]. Ferguson et al. developed a sketch interface for military course-of-action diagrams, which supports queries using spatial relationships [7]. Cohen et al. have developed a multimodal interface (QuickSet) in which users draw gestures on top of an existing map [8]. For example, gestures may be drawn to define regions, specify a route, or indicate a heading. Freksa et al. proposed the use of a schematic map for directing robot navigation [9]. A schematic map is described as an abstraction between a sketch map and a topological map, e.g., a subway map. Finally, in Agrawala and Stolte's work, a sketch

is produced for the user to show a simplified representation of a route [10]. Unlike quantitative maps, the machine-drawn sketch is purposely not drawn to scale and includes only partial information to emphasize pertinent route landmarks.

The focus of this paper is on the analysis of sketched route maps, and we explore whether a reasonable, qualitative description of a route can be extracted, using relative spatial information. Modeling spatial relationships is based on previous work using the histogram of forces [11]-[13]. The main contributions of this paper are the application to hand-drawn route maps, the extraction of key landmark states for critical path changes, and the generation of a linguistic path description. In Section II, we briefly review the histogram of forces. In Section III, sketch interpretation is illustrated with a map sketched on a PDA. Section IV includes an analysis of PDA-generated sketches from a user study. Concluding remarks in Section V include a brief discussion on the current status and future directions.

## II. MODELING SPATIAL RELATIONSHIPS

In the context of image analysis, Matsakis and Wendling introduced the notion of the histogram of forces for modeling spatial relationships between 2D objects [11]. The histogram of forces ensures processing of both raster data and vector data. It offers solid theoretical guarantees and lends itself, with great flexibility, to the definition of fuzzy directional spatial relations such as “to the right of,” “in front of,” etc. [12]. For our purposes, the histogram of forces also allows for a low-computational handling of heading changes in orientation and makes it easy to switch between an allocentric (world) view and an egocentric (path) view. A brief overview is provided here. For details, see also [11]-[13].

### A. The Histogram of Forces

The relative position of a 2D object A with regard to another object B is represented by a function  $F^{AB}$  from  $\mathbf{R}$  into  $\mathbf{R}_+$ . For any direction  $\theta$ , the value  $F^{AB}(\theta)$  is the scalar resultant of elementary forces. These forces are exerted by the points of A on those of B, and each tends to move B in direction  $\theta$  (Fig. 1).  $F^{AB}$  is called the *histogram of forces associated with (A,B) via F*, or the *F-histogram associated with (A,B)*. The object A is the *argument*, and the object B the *referent*. Actually, the letter F denotes a numerical function. Let  $r$  be a real number. If the elementary forces are in inverse ratio to  $d^r$ , where  $d$  represents the distance between the points considered, then F is denoted by  $F_r$ . The  $F_0$ -histogram (histogram of constant forces) and  $F_2$ -histogram (histogram of gravitational forces) have very different characteristics. The former coincides with the angle histogram [14]—without its weaknesses (e.g., requirement for raster data, long processing times, anisotropy)—and provides a global view of the situation. It considers the closest parts and the farthest parts of the objects equally, whereas the  $F_2$ -histogram focuses on the closest parts.

The F-histogram associated with (A,B) is represented by a limited number of values (i.e., a set of discrete directions), and the objects A and B are assimilated to polygons and handled through vector data. The computation of  $F^{AB}$  is of complexity  $O(n \log(n))$ , where  $n$  denotes the total number of vertices [11]. Details on the handling of vector data can be found in [13].

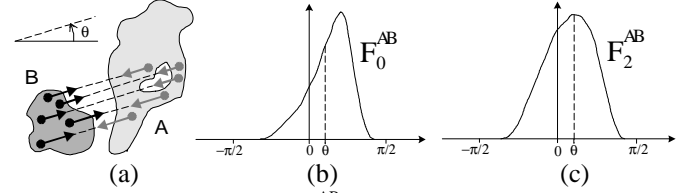


Fig. 1. Force histograms. (a)  $F^{AB}(\theta)$  is the scalar resultant of forces (black arrows). Each one tends to move B in direction  $\theta$ . (b) The histogram of constant forces associated with (A,B), i.e., the position of A relative to B. (c) The histogram of gravitational forces associated with (A,B).

### B. Linguistic Description of Relative Positions

The histogram of forces can be used to build qualitative spatial descriptions that provide a linguistic link to the user. In [12], a system that produces linguistic spatial descriptions of images is presented. The description of the relative position between any 2D objects A and B relies on the sole primitive directional relationships: “to the right of,” “above,” “to the left of” and “below” (imagine that the objects are drawn on a vertical surface). It is generated from  $F_0^{AB}$  and  $F_2^{AB}$ .

For any direction  $\theta$  in which forces are computed, different values can be extracted from the analysis of each histogram. For instance, according to  $F_r^{AB}$ , the degree of truth of the proposition “A is in direction  $\theta$  of B” is  $a_r(\theta)$ . This value is a real number greater than or equal to 0 (proposition completely false) and less than or equal to 1 (proposition completely true). Moreover, according to  $F_r^{AB}$ , the maximum degree of truth that can reasonably be attached to the proposition (say, by another source of information) is  $b_r(\theta)$  (which belongs to the interval  $[a_r(\theta), 1]$ ). The direction  $\theta$  for which  $a_r(\theta)$  is maximum is called the *main direction*.

In [12], the “opinion” given by  $F_r^{AB}$  about the position of A relative to B is represented by  $a_r$ (RIGHT),  $b_r$ (RIGHT),  $a_r$ (ABOVE),  $b_r$ (ABOVE),  $a_r$ (LEFT),  $b_r$ (LEFT),  $a_r$ (BELOW) and  $b_r$ (BELOW). Four numeric and two symbolic features result from the combination of  $F_0^{AB}$  and  $F_2^{AB}$ ’s opinions (i.e., of the sixteen corresponding values). They feed a system of fuzzy rules and meta-rules that outputs the expected linguistic description. The system handles a set of adverbs (like “mostly,” “perfectly,” etc.) that are stored in a dictionary and can be tailored to individual users.

A description is composed of three parts. The first part involves the primary direction (e.g., “A is mostly to the right of B”). The second part augments the description with a secondary direction (e.g., “but somewhat above”). The third part indicates to what extent the four primitive directional relationships are suited to describing the relative position (e.g., “the description is satisfactory”). That is, it indicates to what extent it is necessary

to utilize other spatial relations such as “surrounds.” When range information is available, a fourth part can also be generated to describe distance (e.g., “A is close to B”), as shown in [13].

### III. INTERPRETING A PDA-SKETCHED MAP

In this section we illustrate how qualitative route information is extracted from a map sketched on a PDA such as a PalmPilot. The stylus interface of the PDA allows the user to sketch a map much as he would on paper. The PDA captures the string of (x,y) coordinates as they are drawn on the screen, to acquire the temporal character of the sketch.

The user draws a representation of the environment by sketching the approximate boundary of each object. During the sketching process, a delimiter separates the string of coordinates for each object in the environment. After the environment has been drawn, another delimiter denotes the start of the route, and the user sketches the desired path, relative to the sketched environment. An example of a sketch is shown in Fig. 2a, and Fig. 2b shows the corresponding digital representation, where each point represents a captured screen pixel. To provide unique labels, landmark objects are assigned sequential numbers as they are sketched, as shown in Fig 2b, or they can be assigned labels by the user [15].

In the sections below, procedures are described for (1) extracting the qualitative state of each path step and generating a corresponding linguistic description, (2) extracting the movement along the path, (3) associating a state with each key turning point of the path, and (4) generating the high level path description. In analyzing the sketch and translating it into linguistic descriptions, we make two assumptions. First, we assume the sketch contains sufficient information for the sketched route, i.e., each turn change has at least one corresponding landmark. Tversky and Lee’s work on analyzing sketched route maps supports this assumption [1][4]. Second, the perspective of the generated description assumes that the starting position and orientation of the path are correct and match the description of the starting state.

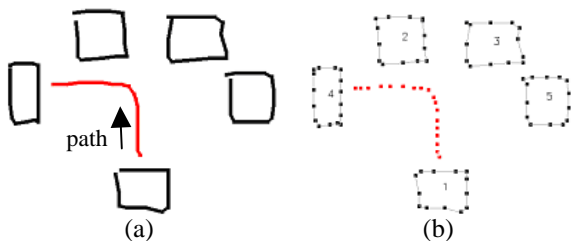


Fig. 2. Sketch 1 (a) A route map sketched on a PDA. (b) The corresponding digital representation

#### A. Extracting Spatial States

The extraction of spatial states from the sketch is summarized in Fig. 3. For each point along the sketched route, a view of the environment is built, initially using a pre-set sensory radius. For each object within the radius, a polygon is constructed using the boundary coordinates of the object as vertices. If any of the object points lies within the sensory radius, the entire object boundary is used as the object model.

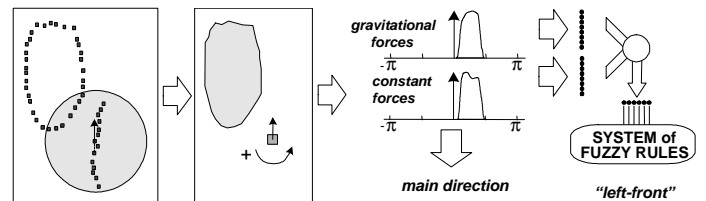


Fig. 3. Synoptic diagram showing how spatial information is extracted from the sketch.

Next, the histograms of constant forces and gravitational forces are computed as described previously. The referent is always a virtual agent positioned at a path step and modeled as a bounding box for the histogram computations. To capture route-centered spatial relationships, the path orientation must also be considered. The heading is computed using adjacent points along the sketched path to determine an instantaneous orientation. We compensate for the discrete pixels by averaging 5 adjacent points (centered on the considered path step), thereby filtering small perturbations and computing a smooth transition as the orientation changes. The filtering also smooths sharp turns. After the heading is calculated, it is used to shift the histograms along the horizontal axis to produce an egocentric view.

The force histograms associated with the path agent and each object are used to generate a detailed linguistic description of the relative position, as discussed in Sec. II. The *main direction* of each object is also extracted from the histogram of constant forces (Sec. II.B) and discretized into one of 16 possible directions (Fig. 4) and used as a symbolic representation of the landmark’s position with respect to the path. Examples of some corresponding linguistic descriptions are shown in Fig. 4.

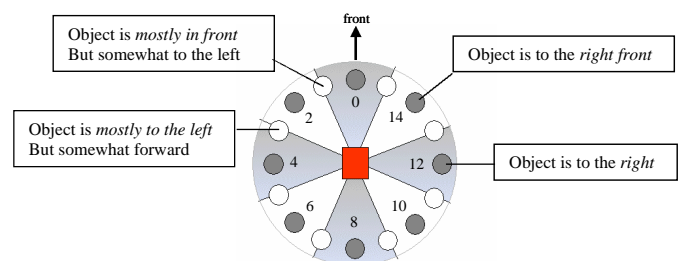


Fig. 4. Sixteen directions are situated around the path agent.

The spatial information extracted from Sketch 1 (Fig. 2) is summarized in Fig. 5. In Fig. 5c, the main direction of each object is plotted for the route steps in which the object is “in view”; labels of the corresponding directions are displayed on the graph to show the symbolic connection. For reference, we have included a sampling of detailed, egocentric linguistic descriptions generated along the route (Fig. 5e). The attached video clip (sketch1.avi) shows an animation of the sketched route and includes detailed descriptions at each step.

### B. Extracting Path Movement

In addition to extracting spatial information on the environment landmarks, we also extract the movement along the sketched path. The computation of the path heading provides an instantaneous orientation. However, we also want to track the change in orientation over time and compute discrete changes, i.e., move forward, turn left, or turn right. The turning rate is determined by computing the change in instantaneous heading between two adjacent route points and dividing by the distance between the points to normalize the rate. A positive rate means a turn to the left, and a negative rate means a turn to the right. The discrete path turns are generated from the turning rate using a threshold (in this case, 0.7 normalized units per time step), and then filtering out spurious changes to smooth the commands. The Fig. 5d graph shows the normalized turning rate and the discrete path turns extracted from Sketch 1. Fig 5b also shows the route turns extracted and the steps at which they occur.

### C. Extracting Landmark States at Critical Path Nodes

The discrete path turns in Fig. 5d show the general trend in the movement along the sketched route and the correlation with the relative landmark positions (Fig. 5c). At the beginning of the route, when object #1 is behind the path, the movement is straight ahead. When objects #2 and #3 are in view, the path turns to the left and stops when object #4 is in front and close.

The starting and ending states and the states at which the path turning rate changes comprise the critical path nodes. To incorporate knowledge of human navigation and the qualitative nature of the spatial information, a system of fuzzy rules is used to extract the significant landmark states at each critical node. In fact, what is needed is the change in landmark state, i.e., the landmark event. Three linguistic variables are used as inputs: (1) the event timing, which is a measure of the closeness of the landmark event to the critical path node (measured in time steps), (2) the main direction of the landmark (0-16), and (3) the object status, e.g., the object has just come in view. The output variable is the event match, a confidence measure of how significantly this landmark event matches the critical path node. The event with the highest confidence is used to identify the landmark state associated with the critical path node.

Fig. 6 shows the rules used and Fig. 7 shows the membership functions. Landmarks in the front have a higher match than those on the sides or in the rear. Landmarks that have just come in view are given a higher match, as are events that are closer in time to the critical node. The results for Sketch 1 are shown graphically in Fig. 5; the open triangles and squares show the correlation between the critical path nodes (Fig. 5d) and the matched landmark event (Fig. 5c). In the case of the critical node at step 7, three landmark events were identified as significant; the match confidence was the same for all three events.

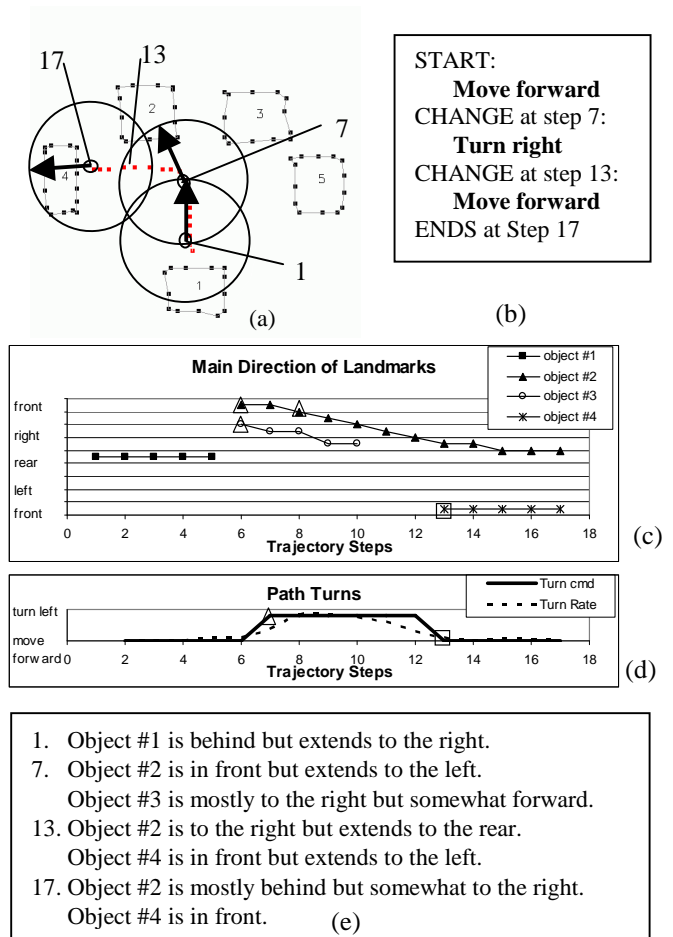


Fig. 5. Analysis of Sketch 1. (a) The digital sketch with an overlay of the sensory radius for several points along the route. (b) Extracted route commands. (c) The discrete main directions of the objects in view. (d) Path turns and turning rate (in normalized units per time step) along the sketched route. Critical path nodes and associated landmark states are identified with open triangle and square symbols (Section V.C). (e) Generated egocentric linguistic descriptions for the route points shown in part (a).

If <i>&lt;event Timing&gt;</i> is <i>&lt;very-close&gt;</i>	then <i>&lt;match&gt;</i> is <i>&lt;very-high&gt;</i>
If <i>&lt;event Timing&gt;</i> is <i>&lt;close&gt;</i>	then <i>&lt;match&gt;</i> is <i>&lt;good&gt;</i>
If <i>&lt;event Timing&gt;</i> is <i>&lt;far&gt;</i>	then <i>&lt;match&gt;</i> is <i>&lt;poor&gt;</i>
If <i>&lt;main Direction&gt;</i> is <i>&lt;front&gt;</i>	then <i>&lt;match&gt;</i> is <i>&lt;very-high&gt;</i>
If <i>&lt;main Direction&gt;</i> is <i>&lt;left&gt;</i>	then <i>&lt;match&gt;</i> is <i>&lt;high&gt;</i>
If <i>&lt;main Direction&gt;</i> is <i>&lt;right&gt;</i>	then <i>&lt;match&gt;</i> is <i>&lt;high&gt;</i>
If <i>&lt;main Direction&gt;</i> is <i>&lt;rear&gt;</i>	then <i>&lt;match&gt;</i> is <i>&lt;poor&gt;</i>
If <i>&lt;object Status&gt;</i> is <i>&lt;new&gt;</i>	then <i>&lt;match&gt;</i> is <i>&lt;very-high&gt;</i>
If <i>&lt;object Status&gt;</i> is <i>&lt;disappeared&gt;</i>	then <i>&lt;match&gt;</i> is <i>&lt;good&gt;</i>

Fig. 6. Rules for Extracting Critical Path Nodes.

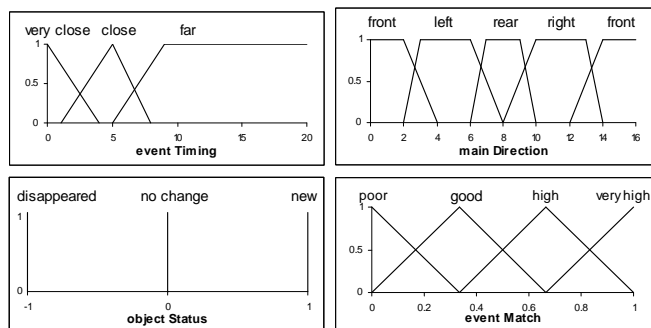


Fig. 7. Membership functions for extracting landmark states at critical path nodes.

#### D. Generating a Linguistic Description of the Path

The landmark states associated with the critical path nodes are used directly to generate a high-level linguistic description of the sketched route. Although a change in landmark state (i.e., an event) is identified from the system of rules, the linguistic descriptions are generated using the new state associated with the event, to yield a human-like description of the path. The qualitative description is generated as a sequence of path segments, each expressed in the form:

*When* <landmark state> *Then* <turn command>

The linguistic expression of the landmark state is a compilation of the objects and their main directions. For each matched object event, the main direction is converted to a shortened linguistic phrase (shown in italics in Fig. 4). For example, object #3 at main direction 14 becomes “object #3 on the right front”. If more than one event is identified for a path node, then the expressions are joined by conjunctions: “or” for the same object and “and” for a different object. The path description of Sketch 1 is then a sequential compilation of the steps, as shown below. Note that the starting and ending descriptions are generated from the states of all the objects in view at those nodes, again using conjunctions for combining multiple objects, and including coarse distance descriptions for close and very close objects.

1. When Object #1 is mostly to the rear (and close) Then Move forward
2. When Object #2 is mostly in front or in front and Object #3 is on the right front Then Turn left
3. When Object #4 is mostly in front Then Move forward
4. When Object #2 is mostly to the rear and Object #4 is in front (and close) Then Stop

#### IV. ANALYSIS TRIALS WITH PDA SKETCHES

To test the robustness of the approach, a user study was conducted, and sketches were collected of two different scenes. Each test subject in group A was shown an environment scene A (laid out on the floor) with a number of landmarks and a path through the landmarks outlined with tape on the floor. Each test subject was asked to sketch a route map using the PDA interface. Each subject in group B was shown a

different environment (scene B) and also asked to sketch a route map. Scene A was purposely constructed with close landmarks and very sharp turns. In contrast, scene B was designed with soft turns and with the landmarks farther from the designated path. Five test subjects were included in each group, all undergraduate or graduate students in computer science or engineering that had no prior knowledge of the sketch interface or the methods used to extract the path description.

Each sketch was analyzed using the methods described in Sec. III, and a linguistic path description was generated for each sketch. To further test the path descriptions, another set of test subjects was shown a generated path description of scene A or B and asked to re-sketch the scene and the designated route, using only the linguistic path description.

Results are shown in Fig. 8 and 9 for 2 example sketches analyzed as part of the study. Two short movie files (sketch2.avi and sketch3.avi) are also included with the electronic version of the paper to animate the sketches. In each animation, a detailed spatial description is displayed at each step along the path (as in Fig. 5e), and the final frame shows the generated, high-level path description.

The sketches collected illustrated a variety of styles. Test subjects sketched the environment using different orientations. In some sketches, the route started from the top, whereas others showed the route starting from the bottom or a side. About 75% of the sketches attempted to use an accurate shape to depict each landmark; however, the rest did not, and the relative size of each landmark was not necessarily accurate. Some users sketched very slowly, which resulted in a very high point density; others sketched rather quickly.

These variations were all handled automatically with the sketch interface. As outlined in Sec. III, the sketch orientation does not matter, as the change in the path turning rate is a relative computation, and landmark representations are captured relative to an egocentric heading. The method used to model spatial relationships adequately captured the relative position of each landmark with respect to the path, in spite of variations in the size and shape of the landmarks drawn. The variable point density is also accommodated by parsing for a fixed increment between points, resulting in consistent digital representations from which the sketch interpretation algorithms are run.

The two scenes were chosen to test two potential problem areas. First, we wanted to compare routes sketched with sharp turns vs. routes sketched with soft turns. Fig. 8a shows a typical sketch from scene A, which illustrates very sharp turns. The filtering algorithm correctly processes the turns; however, they are effectively smoothed and softened into more gradual turns. The softer turns of scene B, as illustrated in Fig 9a, are also correctly processed, but again there is no indication of the sharpness of the turn. In future work, we will consider a qualitative measure of the turning arc.

Second, we wanted to explore the variability in scale. Initial results showed that the analysis method was invariant to shape but only somewhat invariant to scale due to the fixed sensory radius that limits when objects can be “seen” and, as a result,

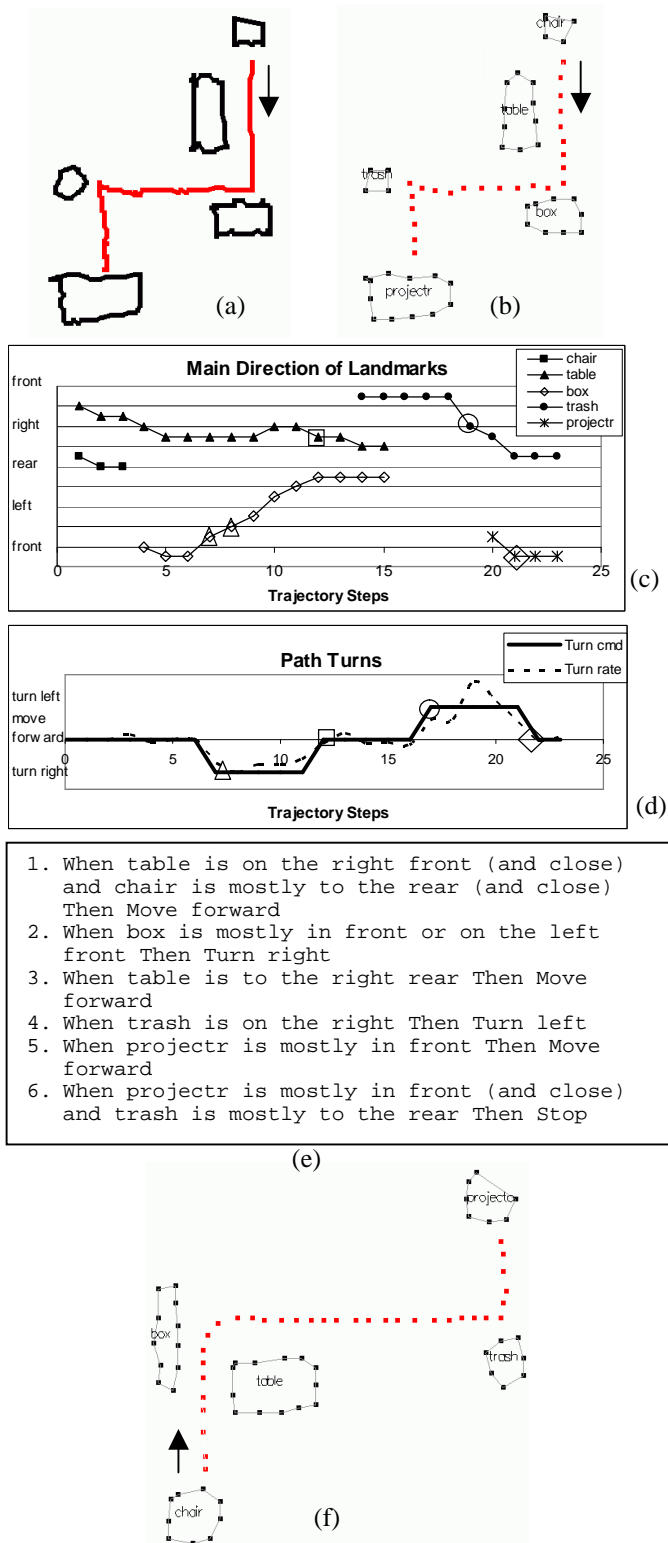


Fig. 8. Sketch 2 (scene A). (a) The sketch (b) The digital representation showing the object labels (c) The discrete main directions of the objects in view (d) Path turns and turning rate (in normalized units per time step) along the sketched route. Critical path nodes and associated landmark states are identified with open triangle, square, circle, and diamond symbols. (Section III.C). (e) The generated linguistic path description. (f) The sketch drawn from the generated path description.

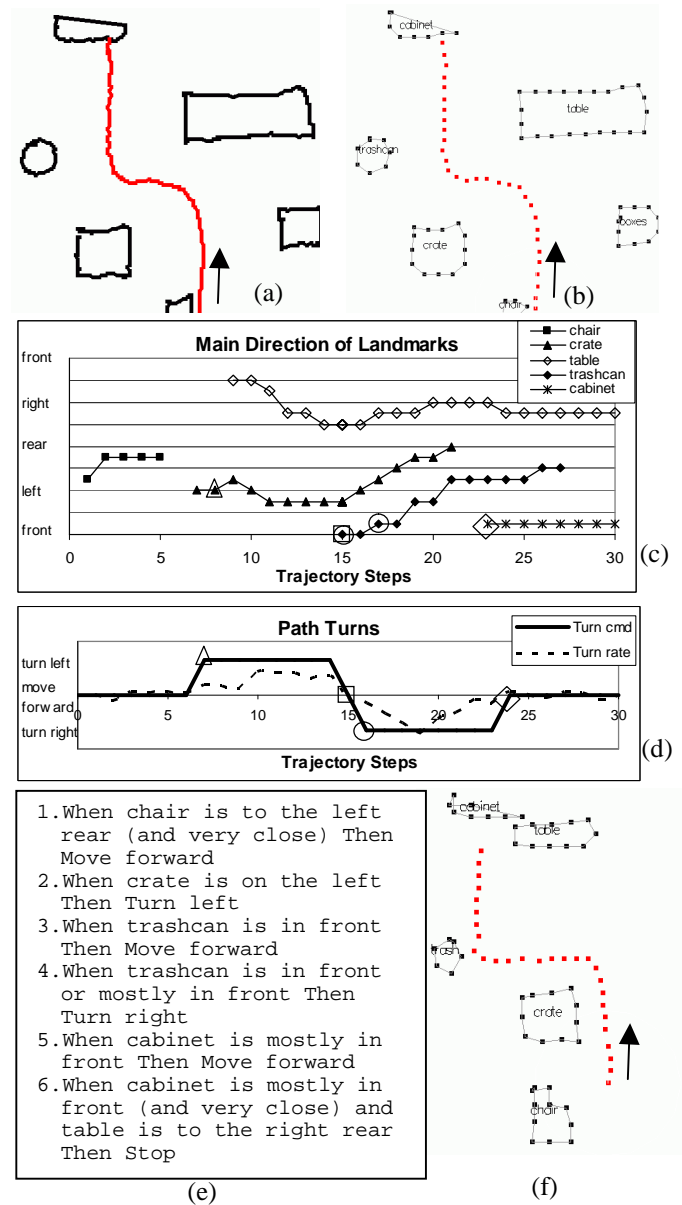


Fig. 9. Sketch 3 (scene B). (a) The sketch (b) The digital representation showing the object labels (c) The discrete main directions of the objects in view (d) Path turns and turning rate (in normalized units per time step) along the sketched route. Critical path nodes and associated landmark states are identified with open triangle, square, circle, and diamond symbols. (Section III.C). (e) The generated linguistic path description. (f) The sketch drawn from the generated path description.

affects the landmark events associated with the turning points. This was not a problem in scene A, in which landmarks were located close to the path. However, there were some problems observed with scene B, in which the landmarks are located farther from the path. To correct this limitation, we incorporated an adaptive sensory radius algorithm. As a first step in identifying the critical landmark states (Sec. IIIC), the algorithm checks for candidate landmark events at each turn change. If no such events are present, the sensory radius is

increased until events are observed. The effect of the adjustable radius can be seen in the animation of Sketch 3.

Overall, the path descriptions generated were reasonable, qualitatively correct, and similar within a scene. The complete results will be reported in [16]. The sketches drawn using only the linguistic path description (Fig. 8f and 9f) are also similar to the original sketches. The Fig. 9f sketch did not include the “boxes” landmark, as this was not included in the path description. Also, the positioning of the table varied because it was included only in the final step.

The path knowledge extracted from the sketches provides turning points based on relative landmark positions; however, there is additional spatial information extracted that includes knowledge along the path between turning points. This between-nodes information could be useful in enabling more complete descriptions. A special case is the path node that occurs at a major intersection. As noted in Sec. I, major intersections are often included if there is some potential for confusion. This is particularly applicable for outdoor urban environments with roads and indoor environments with hallways. The sketches analyzed here do not include these types of major intersections; however, such intersection nodes will be considered in future work.

## V. CONCLUDING REMARKS

The work presented here is an initial approach in analyzing sketched route maps and extracting qualitative information from the sketch. We described the four steps in the analysis process: (1) extracting the qualitative state of each path step and generating the corresponding detailed linguistic description, (2) extracting the qualitative path movement, (3) associating a qualitative state with each key turning point of the path, and finally, (4) generating the high level path description. By extracting relative spatial positions of landmarks with respect to the path, we were able to model the sketched route using human-like navigation and generate a route description in both symbolic terms and linguistic terms. In each of the sketches analyzed, a reasonable description of the path was generated, in terms easily understood by people. Furthermore, we assert that the translation of a sketched route map into linguistic descriptions illustrates that we have extracted the essential route information from the sketch.

The analysis methodology described in the paper is adequate as an offline approach; however, we believe the real potential lies in a more interactive approach with the sketch interface. For example, the qualitative states and turns can be displayed as the route is being sketched so that the user can change them if necessary. Also, editing gestures can be added, as in [17], allowing the user to delete landmarks, add labels to landmarks, and specify qualitative distances. We have begun to explore such extensions and will report the results in a forthcoming paper [15]. In addition, we have been exploring the use of a sketched route map in communicating route instructions to a robot. Preliminary work can be found in [18].

## ACKNOWLEDGMENT

Several colleagues and students have contributed to this overall effort, including Professor Jim Keller and students George Chronis and Andy Carle.

## REFERENCES

- [1] B. Tversky and P. Lee, “How Space Structures Language,” in *Spatial Cognition: An Interdisciplinary Approach to Representing and Processing Spatial Knowledge*, C. Freksa, C. Habel, K. Wender (ed.), Berlin: Springer, 1998, pp. 157-176.
- [2] B. Tversky, “What do Sketches Say about Thinking?” 2002 AAAI Spring Symposium, Sketch Understanding Workshop, Stanford University, AAAI Technical Report SS-02-08, March, 2002, pp. 148-151.
- [3] P.-E. Michon and M. Denis, “When and Why Are Visual Landmarks Used in Giving Directions?” in *Spatial Information Theory: Foundations of Geographic Information Science*, D. Montello (ed.), Berlin: Springer, 2001, pp. 292-305.
- [4] B. Tversky and P. Lee, “Pictorial and Verbal Tools for Conveying Routes,” in *Spatial Information Theory. Cognitive and Computational Foundations of Geographic Information Science*, C. Freksa and D.M. Mark (ed.), Lecture Notes in Computer Science 1661, Springer, Berlin, 1999, pp. 51-64.
- [5] M.J. Egenhofer, “Query Processing in Spatial-Query-by-Sketch”, *Journal of Visual Languages and Computing*, vol. 8, no. 4, pp. 403-424, 1997.
- [6] T. Igarashi, R. Kadobayashi, K. Mase, H. Tanaka, “Path Drawing for 3D Walkthrough”, in *Proc. of the 11th Annual Symposium on User Interface Software and Technology*, ACM UIST 1998, San Francisco, CA, Nov., 1998, pp. 173-174.
- [7] R. Ferguson, R. Rasch, W. Turmel, and K. Forbus, “Qualitative Spatial Interpretation of Course-of-Action Diagrams”, in *Proc. of the 14th Intl. Workshop on Qualitative Reasoning*, Morelia, Mexico, 2000.
- [8] P.R. Cohen, M. Johnston, D. McGee, S. Oviatt, J. Pittman, I. Smith, L. Chen, and J. Clow, “QuickSet: Multimodal Interaction for Distributed Applications”, in *Proc. of the Fifth Intl. Multimedia Conference*, ACM Press, 1997, pp 31-40.
- [9] C. Freksa, R. Moratz, and T. Barkowsky, “Schematic Maps for Robot Navigation”, in *Spatial Cognition II: Integrating Abstract Theories, Empirical Studies, Formal Methods, and Practical Applications*, C. Freksa, W. Brauer, C. Habel, K. Wender (ed.), Berlin: Springer, 2000, pp. 100-114.
- [10] M. Agrawala and C. Stolte, “Rendering Effective Route Maps: Improving Usability Through Generalization”, *ACM SIGGRAPH 2001*, Los Angeles, CA, August, 2001, pp. 241-250.
- [11] P. Matsakis and L. Wendling, “A New Way to Represent the Relative Position between Areal Objects”, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 21, no. 7, pp. 634-643, 1999.
- [12] P. Matsakis, J. Keller, L. Wendling, J. Marjamaa, and O. Sjahputera, “Linguistic Description of Relative Positions in Images”, *IEEE Trans. on Systems, Man and Cybernetics*, part B, vol. 31, no. 4, pp. 573-588, 2001.
- [13] M. Skubic, P. Matsakis, G. Chronis and J. Keller, “Generating Multi-Level Linguistic Spatial Descriptions from Range Sensor Readings Using the Histogram of Forces,” *Autonomous Robots*, Jan., 2003, vol. 14, no. 1, pp. 51-69.
- [14] K. Miyajima and A. Ralescu, “Spatial Organization in 2D Segmented Images: Representation and Recognition of Primitive Spatial Relations,” *Fuzzy Sets and Systems*, vol. 65, no. 2/3, pp. 225-236, 1994.
- [15] M. Skubic, C. Bailey, and G. Chronis, “A Sketch Interface for Mobile Robots,” accepted to the IEEE 2003 International Conference on SMC, Washington, DC, Oct., 2003.
- [16] C. Bailey, “A Sketch Interface for Understanding Hand-Drawn Route Maps”, M.S. Thesis, CECS Department, Univ. of Missouri-Columbia, in preparation.
- [17] J. Landay and B. Myers, “Sketching Interfaces: Toward More Human Interface Design,” *IEEE Computer*, vol. 34, no. 3, pp. 56-64, 2001.
- [18] G. Chronis and M. Skubic, “Sketch-Based Navigation for Mobile Robots”, in *Proc. of the IEEE 2003 Intl. Conf. on Fuzzy Systems*, St. Louis, MO, May, 2003, pp. 284-289.