A Sketch Interface for Mobile Robots*

Marjorie Skubic, Craig Bailey and George Chronis
Computer Engineering and Computer Science Department
University of Missouri-Columbia
Columbia, MO 65211
skubicm@missouri.edu

Abstract – *In human-to-human communication, a hand-drawn route map is often sketched to show a desired navigation path. In this paper, we describe a PDA sketching interface that can be used to direct a mobile robot along a specified path. Because sketched route maps are not drawn precisely or necessarily to scale, we do not attempt to analyze precise path information, but rather qualitative route information is extracted. The paper focuses on the front-end sketch understanding and includes a description of the interactive features, such as deleting, moving, and labeling landmarks. Results of a user evaluation are also presented, in which participants report that the sketching interface was as easy as using pencil and paper by a 2:1 margin.*

Keywords : human-robot interaction, route map, sketch understanding

1 Introduction

The underlying goal behind this work is the creation of robot interfaces that allow a novice user to guide, control, and/or program a robot to perform some purposeful task. To address this goal, we have been investigating the use of hand-drawn route maps, in which the user sketches an approximate representation of the environment and then sketches the desired robot trajectory with respect to that environment. The objective of the sketch interface is to extract topological information about the map and a qualitative path through the landmarks drawn on the sketch. This topological information is then used to build a task representation for the robot, which operates as a semi-autonomous vehicle. The task representation is based on sensing and relative position, *not* absolute position.

Route maps have been investigated previously in the context of human-to-human communication. Although they 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 [13]. The depiction of the environment structure is not necessarily accurate and may even distort the actual configuration [14]. 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 [14]. Indeed, in a study of 29 sketched route maps, each contained the information necessary to complete a complicated navigation task [13].

Research by Michon and Denis [8] 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. The work of the researchers noted above and others indicate the importance of environment landmarks and spatial relationships in human navigation. The work suggests that spatial relationships of landmarks with respect to the desired route may be useful not only for robot control but also as a link between a robot and its human user.

Some previous work has been done in using a computer sketch interface to direct movement. Ferguson et al. developed a sketch interface for military course-of-action diagrams, which are used for strategic planning [4]. Cohen et al. have developed a multimodal interface (QuickSet) in which users draw gestures on top of an existing map [3]. 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 [5]. A schematic map is described as an abstraction between a sketch map and a topological map, e.g., a subway map. The closest work is that of Kawamura et al. in which the user specifies a robot path by selecting intermediate points on a sketch of the environment [6]. Artificial landmarks are placed in the scene and on the sketch to act as landmarks in the navigation process.

* **0-7803-7952-7/03/$17.00 © 2003 IEEE.**

In previous work, we have presented strategies for extracting a sequence of navigation states from a sketched route map [9], translating the sketched route to a linguistic description [10], and using the sequence of states for robot navigation [2]. In this paper, the sketching interface is extended to be more interactive. Section 2 describes the new sketching interface. Sections 3 and 4 describe the experimental design and results of a usability study designed to examine the system. In Section 5, we provide concluding remarks.

2 The Sketch Interface

Our work with sketches has utilized a PDA as the interface device [9]. The user simply draws the map on the PDA screen using the stylus, and the sketch is captured as a sequence of (x,y) coordinates. The path information extracted from the sketch, as used for robot navigation, is inspired by human navigation and insights on how people draw and use sketched route maps. The extracted route is represented as a sequence of steps, each with a landmark state and an associated movement (turn right, turn left, go straight, or stop). Each landmark state consists of the qualitative spatial relationship between one or more landmarks and the robot.

Spatial relationships of the environment landmarks are computed with respect to the robot, imposing a radius around the robot position to represent the sensory range, as the robot moves along the drawn path. (See also [9, 10, 2].) The histogram of forces [7], originally developed for processing images, provides a fast and robust tool for computing the spatial relationship between each environment object and the robot. Several features are extracted from the histogram to represent the qualitative state of a landmark, including the "main direction" (the center of mass of the histogram), the histogram width, and the qualitative distance from the robot.

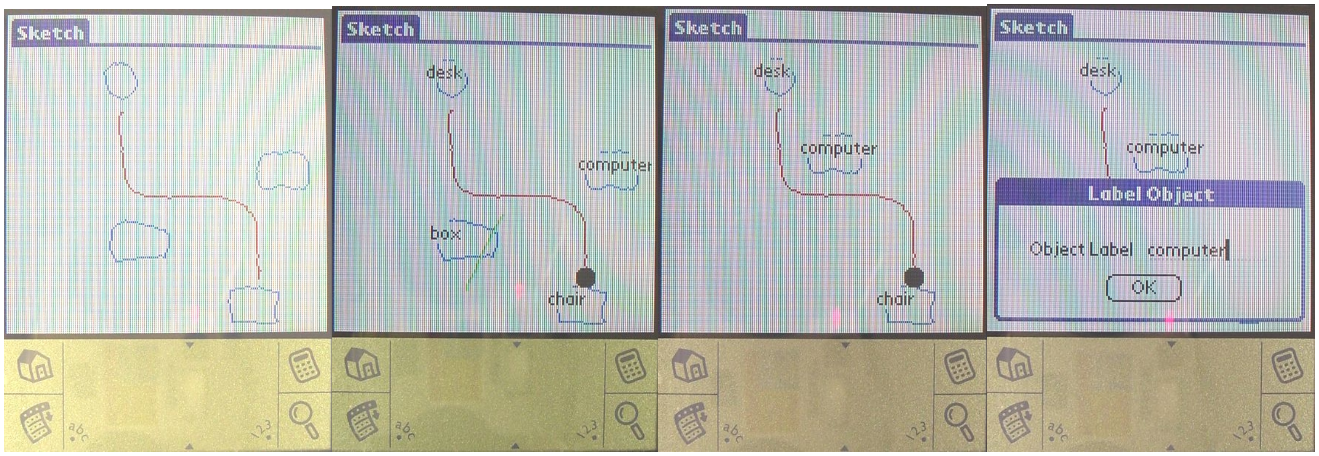
The challenge of the sketch interface is to correctly and consistently interpret the stylus markings of the user. The first sketching system was not interactive and required knowledge specific to the interface. The goal of the revised system is to create a natural sketching interface; the user should be able to draw and/or mark out objects and paths in such a way that the use of menus is minimized.

The sketch interface interprets the markings of the user and transforms the sketched image into a representation of the sketched route. An example sketch is shown in Figure 1. The drawing window is 160 x 160 pixels. The interface identifies and records each landmark drawn and distinguishes landmarks from the sketched path. This is accomplished by recognizing any closed polygon as a landmark object; thus, there are no restrictions on the shape or size of the landmarks. To determine a closed polygon, the sketch interface examines if the Euclidian distance between the starting and end points of a drawn object fall within an empirically set threshold. To improve performance on the PDA, distance calculations are confined to integer-based square root routines. A closed object signals to the user that the system interpreted the stroke as a new object. The objects are also displayed in blue as another cue that it was recognized. Audible beeps provide further feedback when the pen touches the screen, and again when the sketch interface recognizes a drawn object.

A path is interpreted as any line segment that is not recognized as an object, and with total length greater than a set threshold. A path has no restrictions on its maximum length, the number of turns, or the abruptness of the turns. A path is displayed in red to distinguish it from the drawn objects. The default direction of the path is denoted by the direction in which the segment is drawn. Alternatively, the user can add a starting point as shown in Figure lb. To add a start point, the user makes a small "squiggle" on either end of the path near the last point. The sketch interface compares the average distance of all of the drawn points of the segment to a threshold. If the average distance is less than the threshold and drawn near one of the endpoints of a path, then a round black circle is drawn to denote the starting position of the path. The sketch interface ignores any "squiggle" marks not near the path endpoints and deletes them.

Our objective was to take advantage of the interactive editing capabilities of the PDA over that of pencil and paper. An obvious extension was the ability to intuitively delete objects. We allow the user to simply cross out an object on the display with a drawn 'X', the first stroke of which is also shown in Figure 1b. When the sketch interface determines that two non-path strokes have been made, it computes whether the stroke segments cross each other, and if so, where the intersection point lies. Using slope equations, decision parameters can be computed quickly without having to fully calculate the intersection point. When an intersection is detected, then the sketch interface searches for the nearest object to that intersection point within a radius of the longest arm of the drawn 'X'. The user can also cross out a path in this manner. A path can also be deleted by simply drawing a new path; only the last path drawn is retained.

Another advantage of the interactive environment is the ability to move an object. This is currently an option for the user that does require using either the menu, or writing the graffiti character 'm' on the graffiti pad. The sketch interface prompts the user to tap on an object to move, and if an object was selected, prompts the user for the new location. The user taps the screen for either event. The object is translated using the distance between the two tapped points. An example is shown in Figure 1c.



(a) (b) (c) (d)

Figure 1. (a) An unlabeled sketch. (b) A sketch showing the start point and one deletion stroke on the "box" object. (c) The sketch after the box was deleted and the computer was moved. (d) A dialog for labeling the objects.

We also give the user the ability to undo any operation that has been done, invoked either from the menu or with the graffiti character 'u'. Only one level of undo was incorporated. Deleted objects and paths can be restored, as well as removing the most recently added object or path. A moved object will be moved back to its original position with an undo. The start point can be undone as well. A final menu item, also available as the graffiti character 'c', is a clear screen option to start a new sketch.

The ability to label the landmarks in the sketched map allows for further interaction between the user and the robot in the context of its environment. Default landmark numbers are assigned as the landmarks are drawn. The user can then edit the labels. The labeling system uses the PDA graffiti system (and includes the PDA stylus "keyboard"). The user simply taps the existing label and a dialog box appears to add/edit the object label as shown in Figure 1d.

After a landmark is labeled, the user (or the robot) can now refer to that landmark by name. This provides a means to interact with any object recognition systems that the robot may possess. Indeed, we are incorporating such capabilities into our robot system [11] with morphological shared weight neural networks (MSNN) [15] for landmark recognition. Note, however, that the robot navigation of the sketched route does not *depend* on object recognition.

Our existing robot control system [11] communicates via TCP-IP client-server systems. The sketch interface has been designed to interact with this system through wireless Ethernet. Thus, real-time sketches of the environment can be input to the robot from a distance. A sketch server captures the sketches sent from the sketch client on the PDA. The server stores the sketched information as a listing of (x,y) coordinates with text delimiters denoting objects, paths, and labels.

Finally, to properly process the sketched route, we have found it necessary to prune out duplicate points that might have been captured. The object boundary points are used to compute the histogram of forces, as in [12], and the algorithm is more efficient if the point density is less than what is typically drawn. We have found that users have vastly different drawing styles, resulting in widely varying point densities in their sketches. To alleviate this problem consistently, the object and path point listings are parsed using a minimum distance between points to determine inclusion into the analyzed sketch file. Duplicate points, and points too close together are then discarded as shown in Figure 2.

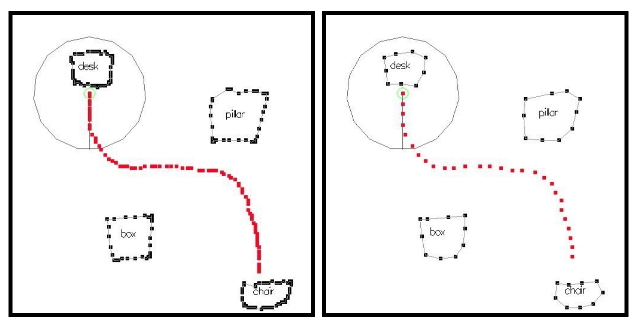


Figure 2. On the left is a digital representation of a sketch showing all of the captured points. The sketch on the right is the same sketch after parsing the points, resulting in a more equal and sparse spacing of points.

3 Experimental Design

To test the sketching system, we conducted a user evaluation to examine three questions:

1. To test the basic sketching features of the interface. (How easy is the interface to use?)

2. To test the route representation extracted and its use for robot navigation. (Does it extract the route intended and in a form usable by the robot?)

3. To test the translation of the sketched route into linguistic descriptions. (Are the generated route descriptions understandable by another person?)

This paper addresses primarily the first question, ease of use, although we also provide some examples of the extracted route for robot navigation. Details of the route extraction methodology can be found in [2]. Complete results of the study will be reported in [1].

Participants were first asked to fill out a pre-experimental survey. They were then given a short five-minute how-to on the interface and the available features. They were then asked to demonstrate drawing objects and paths and to demonstrate using the delete, move, and undo features. It was suggested to users that they use a moderate drawing speed when sketching for best results. The users were then given several minutes to practice sketching and to ask any questions they might have.

Participants were divided into 2 equal groups. In Part I, each participant 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. Participants were told which end of the path to consider the start point. Landmarks included storage cabinets, boxes, a table, a chair, a desk, a wooden crate, and a trashcan. Each participant was asked to draw a sketch of the route map using the designated PDA interface. Finally participants were asked to fill out a post-experimental survey. Each participant in group B was shown an environment scene B and also asked to draw a sketch of the route map using the PDA interface and fill out the two surveys.

4 Results

Twenty-six participants consisted of computer science students taking first and second year undergraduate courses. Pre-survey results revealed the following details. There were twenty-two males and 4 females ranging in age from 18-49. Nineteen were of age 20-29. One participant was colorblind. Only 5 of the participants owned a PDA themselves, although several others have used them at one time or another. None of the participants rated their expertise in using a PDA as expert. In fact, on a scale of 1 to 5 (5 being expert) only one participant claimed a 4. Only three of the participants knew Palm OS Graffiti, and only two of those participants claimed to have previously used it. When asked how good they were at following correct directions provided by other people, on a scale of 1 to 5 (5 being very good) 21 participants claimed at least a score of 4. The other five participants claimed the rest of the scores, including a score of 1. When asked about their ability to accurately provide directions, 20 participants claimed a score of at least 4, while the remaining six participants claimed a score of 3.

Some example sketches, viewed after capture by the server, are shown in Figure 3. Overall, the participants rated the ease of using the sketch interface with a mean of 4.4 on a scale of 1 to 5 (5 being very easy) with a standard deviation of 0.64. When asked if they felt that the PDA was as easy to use as pencil and paper for drawing a sketch, they surprisingly responded 2:1 that it was. In fact, no aspect of the system was rated less than an average of 4.1 using the same scale. When asked how well they felt their final sketch of the scene represents the environment, on a scale of 0 to 100, the average response was 84 with a standard deviation of 7.4. When asked how well they felt their sketch represented the scale of the environment, the average score was 68, with a standard deviation of 15. (Scores ranged from 30 to 90.) This lower score might not be unexpected as the scenes that were sketched were in a rectangular room, while the drawing window on the PDA was square. In fact, several comments gathered from the surveys were that the screen was too small.

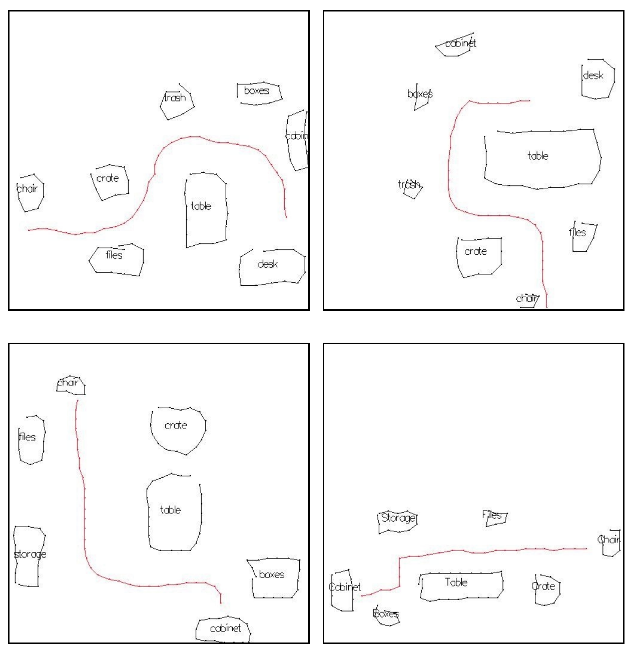


Figure 3. Example sketches from participants. Top row: scene A. Bottom row: scene B. Note the differences in the use of curved and sharp turns in the paths.

Participants were asked to score the usefulness of various sketching functionalities, such as drawing paths, drawing and labeling objects, moving objects, undo actions, and clearing the screen. The lowest average score was 4.2 (for clearing an entire sketch). They rated the ability to change a sketch an average 4.8 with a standard deviation of 0.49. This gives support that the interactivity of the PDA is a useful and novel medium to capture sketched route paths when compared to traditional pencil and paper. The ability to place labels on the objects also rated highly with an average of 4.8 and a standard deviation of 0.59. The abilities to remove objects, remove paths, and undo a stroke were also all rated an average of 4.7. Interestingly, the question regarding the use of color cues had the most variability in answers with a standard deviation of 1.1 and an average of 4.3. The scores for the use of color ranged from 2 to 5. This contrasts to a common comment that color was one of the liked features of the interface. Figure 4 shows example sketches that were drawn by two of the participants.

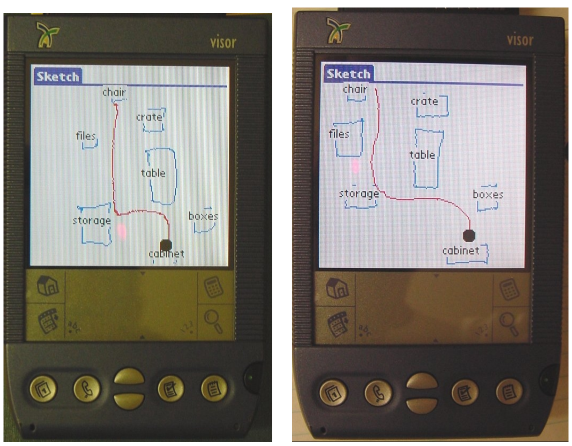


Figure 4. Two sketches of the same scene drawn by different participants, as viewed on the PDA.

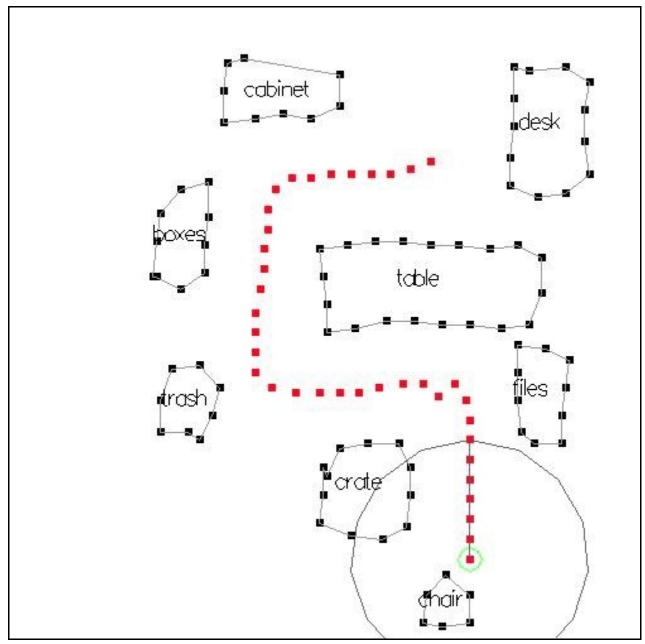
To address the second question of the user evaluation, the sketches were processed off-line to extract the route information. Sample results of each scene are shown in Figures 5 and 6. The digitized representation of each sketch is shown with the corresponding output of the spatial relations algorithms. At each time that a path heading change is detected, the key landmark states are identified, and a robot command is generated [2]. The first scene (Figure 5) had one left turn and two right turns as well as a start and stop state. Between each turn and before and after the stop and start, respectively, there is a forward movement. The algorithm extracted these state features from the sketched path. The path in scene B (Figure 6) had one left turn and one right turn in addition to the start and stop states.

5 Concluding remarks

We have shown how a PDA can be used to easily capture sketched route maps from users. These maps can then be processed and used to guide a real mobile robot. Feedback from participants in a user evaluation felt that the sketching interface was as easy as using pencil and paper by a 2:1 margin. This shows it is possible to create an intuitive sketch-based navigation system for mobile robots. Future extensions of this work will focus on the robustness of the algorithms to create consistent and reliable robot directives based on different sketches of the same scenes.

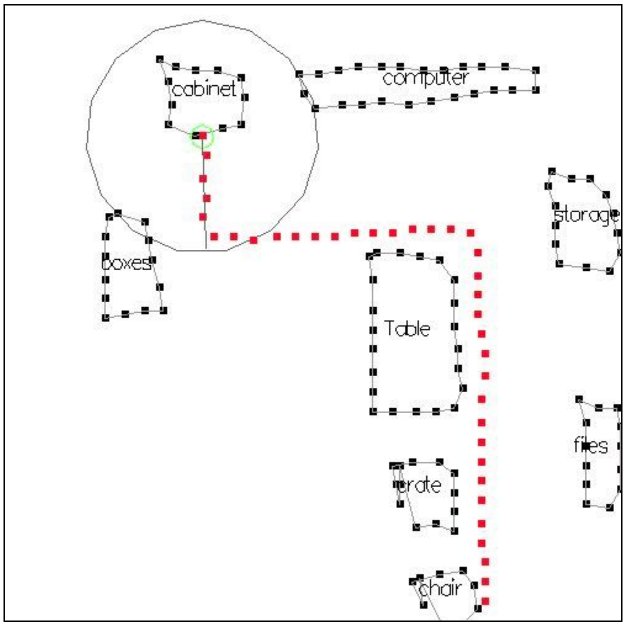
Acknowledgements

Several MU colleagues and students have contributed to this overall effort, including Professors Jim Keller and Pascal Matsakis and students Sam Blisard and Andy Carle. We also thank Dr. Julie Adams for assistance with the user evaluation. This research has been supported by the Naval Research Laboratory.



Route Step	Key Landmarks Identified	Robot Commands
Start	Chair behind Crate left	Move forward
Step 1	Table front	Turn left
Step 2	Crate mostly left Table right-rear Trash front	Move forward
Step 3	Trash front	Turn right
Step 4	Table mostly right	Move forward
Step 5	Cabinet mostly front	Turn right
Step 6	Desk front	Move forward
Step 7	Table right Desk front Cabinet mostly rear	Stop

Figure 5. Top : An example sketch of scene A. Bottom : The robot commands extracted for each route step. Other landmark information extracted includes the histogram width and the qualitative distance to the sketched route.



Route Step	Key Landmarks Identified	Robot Commands
Start	Computer left rear Boxes right front Cabinet rear	Move forward
Step 1	Boxes mostly right	Turn left
Step 2	Computer mostly left Table right-front	Move forward
Step 3	Storage mostly front	Turn right
Step 4	Table right Files left-front Storage mostly left	Move forward
Step 5	Crate mostly rear Chair right	Stop

Figure 6. Top : An example sketch of scene B. Bottom : The robot commands extracted for each route step. Other landmark information extracted includes the histogram width and the qualitative distance to the sketched route.

References

[1] C. Bailey, "A Sketch Interface for Understanding Hand-Drawn Route Maps", M.S. Thesis, CECS Department, Univ. of Missouri-Columbia, in preparation .

[2] G. Chronis and M. Skubic, "Sketch-Based Navigation for Mobile Robots", in *Proc. of the 2003 IEEE Intl. Conf. on Fuzzy Systems*, St. Louis, MO, May, 2003.

[3] 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 Proceedings of the Fifth International Multimedia Conference, ACM Press, 1997, pp 31-40.

[4] R. Ferguson, R. Rasch, W. Turmel, and K. Forbus, "Qualitative Spatial Interpretation of Course-of-Action Diagrams", in Proceedings of the 14th International Workshop on Qualitative Reasoning, Morelia, Mexico, 2000.

[5] 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.

[6] K. Kawamura, A.B. Koku, D.M. Wilkes, R.A. Peters II and A. Sekmen, "Toward Perception-Based Navigation Using Egospheres," *International Journal of Robotics and Automation*, to appear.

[7] 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.

[8] 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.

[9] M. Skubic, P. Matsakis, B. Forrester and G. Chronis, "Extracting Navigation States from a Hand-Drawn Map", in *Proc. of the IEEE 2001 Intl. Conf. on Robotics and Automation*, May, 2001, Seoul, Korea, pp. 259-264.

[10] M. Skubic, S. Blisard, and P. Matsakis, "Qualitative Analysis of Sketched Route Maps: Translating a Sketch into Linguistic Descriptions," *IEEE Transactions on SMC*, Part B, to appear.

[11] M. Skubic, C. Bailey, S. Blisard, G. Chronis, C. Huggard, R. Luke, G. Scott, and M. Williams, "Guinness Phase I," Technical Report, CECS Dept., Univ. of Missouri-Columbia, June, 2003.

[12] 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.

[13] 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.

[14] 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.

[15] Y. Won, P. Gader and P. Coffield, "Morphological Shared-Weight Networks with Application to Automatic Target Recognition," *IEEE Transactions on Neural Networks*, vol. 8, pp. 1195-1203, 1997.