

Spatial Relations for Tactical Robot Navigation

Marjorie Skubic, George Chronis, Pascal Matsakis, and James Keller

Computer Engineering and Computer Science Department
University of Missouri-Columbia, Columbia, MO 65211

ABSTRACT

In this paper, we provide an overview of our on-going work using spatial relations for mobile robot navigation. Using the histogram of forces, we show how linguistic expressions can be generated to describe a qualitative view of the robot with respect to its environment. The linguistic expressions provide a symbolic link between the robot and a human user, thus facilitating two-way, human-like communication. In this paper, we present two ways in which spatial relations can be used for robot navigation. First, egocentric spatial relations provide a robot-centered view of the environment (e.g., *there is an object on the left*). Navigation can be described in terms of spatial relations (e.g., *move forward while there is an object on the left, then turn right*), such that a complete navigation task is generated as a sequence of navigation states with corresponding behaviors. Second, spatial relations can be used to analyze maps and facilitate their use in communicating navigation tasks. For example, the user can draw an approximate map on a PDA and then draw the desired robot trajectory also on the PDA, relative to the map. Spatial relations can then be used to convert the relative trajectory to a corresponding navigation behavior sequence. Examples are included using a comparable scene from both a robot environment and a PDA-sketched trajectory showing the corresponding generated linguistic spatial expressions.

Keywords: spatial relations, linguistic spatial descriptions, mobile robot navigation, human-robot communication, histogram of forces

1. INTRODUCTION

Being able to interact and communicate with robots in the same way we interact with people has long been a goal of AI and robotics researchers. Much of the robotics research has emphasized the goal of achieving autonomous robots. However, in our research, we are less concerned with creating autonomous robots that can plan and reason about tasks, and instead we view them as semi-autonomous tools that can assist a human user. The user supplies the high-level and difficult reasoning and strategic planning capabilities. We assume the robot has some perception capabilities, reactive behaviors, and perhaps some limited reasoning abilities that allow it to handle an unstructured and dynamic environment.

In this scenario, the interaction and communication mechanism between the robot and the human user becomes very important. The user must be able to easily communicate what needs to be done, perhaps at different levels of task abstraction. In particular, we would like to provide an intuitive method of communicating with robots that is easy for users that are not expert robotics engineers. We want domain experts to define their own task use of robots, which may involve controlling them, guiding them, or even programming them.

In ongoing research on human-robot interaction, we have been investigating the use of spatial relations in communicating purposeful navigation tasks. Linguistic, human-like expressions that describe the spatial relations between a robot and its environment provide a symbolic link between the robot and the user, thus comprising a type of navigation language. The linguistic spatial expressions can be used to establish effective two-way communications between the robot and the user, and in this paper, we provide approaches from both perspectives.

First, from the robot perspective, we have studied how to recognize the current (qualitative) state in terms of egocentric spatial relations between the robot and objects in the environment, using sensor readings only (i.e., with no prior map or model of the environment). Linguistic spatial descriptions of the state are then generated for communication to the user.

Second, from the user perspective, we offer a novel approach for communicating a navigation task to a robot, which is based on robot-centered spatial relations. Our approach is to let the user draw a sketch of an environment map (i.e., an approximate representation) and then sketch the desired robot trajectory relative to the map. State information is extracted from the drawing on a point by point basis along the sketched robot trajectory. We generate a linguistic description for each point and show how the robot transitions from one qualitative state to another throughout the desired path. A complete navigation task is represented as a sequence of these qualitative states based on the egocentric spatial relations, each with a corresponding navigation behavior. We assume the robot has pre-programmed or pre-learned, low-level navigation behaviors that allow it to move safely around its unstructured and dynamic environment without hitting objects. In this approach, the robot does not have a known model or map of the environment, and the user may have only an approximate map. Thus, the

navigation task is built upon connected spatial states (i.e., qualitative states), which form a type of topological map. Note that we are not attempting to build an exact model of the environment, nor to generate a quantitative map. However, we do want to generate linguistic descriptions that represent the qualitative state of the robot with respect to its environment, in terms that are easily understood by human users.

The idea of using linguistic spatial expressions to communicate with a semi-autonomous mobile robot has been proposed previously. Gribble *et al* use the framework of the Spatial Semantic Hierarchy for an intelligent wheelchair [1]. Perzanowski *et al* use a combination of gestures and linguistic directives such as “go over there” [2]. Shibata *et al* use positional relations to overcome ambiguities in recognition of landmarks [3]. However, the idea of communicating with a mobile robot via a hand-drawn map appears to be novel. The strategy of using a sketch with spatial relations has been proposed by Egenhofer as a means of querying a geographic database [4]. The hand-drawn sketch is translated into a symbolic representation that can be used to access the geographic database.

In this paper, we show how spatial relations can be extracted both from a robot’s sensors and from a hand-drawn map sketched on a PDA. In Section 2, we discuss background material on the spatial analysis algorithms, which are an extension of work previously applied to image analysis. In Section 3, we show how the robot’s sonar readings can be used to generate inputs for the spatial analysis algorithms. In Section 4, we show a method for extracting the environment representation and the corresponding states from the PDA sketch. Experiments are shown in Section 5 using a comparable scene from both a robot environment and a PDA-sketched trajectory showing the corresponding generated linguistic spatial expressions. We conclude in Section 6 and discuss future work.

2. SPATIAL RELATIONS METHODS

Freeman [5] proposed that the relative position of two objects be described in terms of spatial relationships (such as “above”, “surrounds”, “includes”, etc.). He also proposed that fuzzy relations be used, because “all-or-nothing” standard mathematical relations are clearly not suited to models of spatial relationships. By introducing the notion of the histogram of angles, Miyajima and Ralescu [6] developed the idea that the relative position between two objects can have a representation of its own and can thus be described in terms other than spatial relationships. However, the representation proposed shows several weaknesses (*e.g.*, requirement for raster data, long processing times, anisotropy).

In [7][8], Matsakis and Wendling introduced the histogram of forces. Contrary to the angle histogram, it ensures processing of raster data as well as of vector data. Moreover, it offers solid theoretical guarantees, allows explicit and variable accounting of metric information, and lends itself, with great flexibility, to the definition of fuzzy directional spatial relations (such as “to the right of”, “in front of”, etc.). For our purposes, the histogram of forces also allows for a low-computational handling of heading changes in the robot’s orientation and makes it easy to switch between a world view and an egocentric robot view.

2.1 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 θ , the value $F^{AB}(\theta)$ is the total weight of the arguments that can be found in order to support the proposition “A is in direction θ of B”. More precisely, it 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 θ (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. 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 and very interesting characteristics. The former coincides with the angle histogram—without its weaknesses—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.

Throughout this paper, the referent B is always the robot. The F-histogram associated with (A,B) is represented by a limited number of values (*i.e.*, the set of directions θ is made discrete), and the objects A and B are assimilated to polygons (vector data). It is shown that the computation of F^{AB} is of complexity $O(n \log(n))$, where n denotes the total number of vertices.

Details can be found in [7][8].

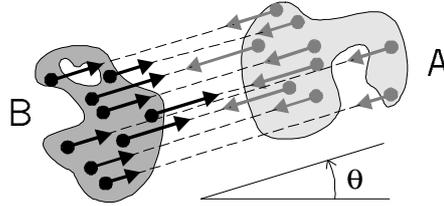


Figure 1. Computation of $F^{AB}(\theta)$. It is the scalar resultant of forces (black arrows). Each one tends to move B in direction θ .

2.2 Linguistic Description of Relative Positions

In [9][10], Matsakis *et al.* present a system that produces linguistic spatial descriptions. 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} (the histogram of constant forces associated with (A,B)) and F_2^{AB} (the histogram of gravitational forces). First, eight values are extracted from the analysis of each histogram: a_r (RIGHT), b_r (RIGHT), a_a (ABOVE), b_a (ABOVE), a_l (LEFT), b_l (LEFT), a_b (BELOW) and b_b (BELOW). They represent the “opinion” given by the considered histogram (*i.e.*, F_0^{AB} if r is 0, and F_2^{AB} if it is 2). For instance, according to F_0^{AB} the degree of truth of the proposition “A is to the right of B” is a_0 (RIGHT). 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_0^{AB} the maximum degree of truth that can reasonably be attached to the proposition (say, by another source of information) is b_0 (RIGHT) (which belongs to the interval $[a_0$ (RIGHT),1]). F_0^{AB} and F_2^{AB} ’s opinions (*i.e.*, the sixteen values) are then combined. Four numeric and two symbolic features result from this combination. 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.) which are stored in a dictionary, with other terms, and can be tailored to individual users. A description is generally composed of three parts. The first part involves the primary direction (*e.g.*, “A is mostly to the right of B”). The second part supplements the description and involves 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 of the objects (*e.g.*, “the description is satisfactory”). In other words, it indicates to what extent it is necessary to utilize other spatial relations (*e.g.*, “surrounds”).

The use of a dictionary for storing the linguistic terms provides flexibility and easy adaptability. The precise terminology and phrasing can easily be adjusted to suit the application or the user. The terminology can even be translated to create multi-lingual expressions.

3. EXTRACTING SPATIAL STATES FROM ROBOT SENSORS

In this section, we describe the application of the F_0 and F_2 histograms for extracting spatial relations from the sensor readings of a mobile robot. For this application, we use a vector data representation (*i.e.*, a boundary representation using vertices), which simplifies the computational complexity and provides a method for producing the linguistic expressions in real time. In this work, we have used a Nomad 200 robot with 16 sonar sensors evenly distributed along its circumference. The sensor readings are used to build a polygonal representation of the objects surrounding the robot. The vertices of each polygon are extracted and the F_0 and F_2 histograms are built, as described in Section 2.1. The histograms are then used to generate linguistic descriptions of relative positions between the robot and the environment objects (see Figure 2). Note that although we show a specific sensor type and layout, the methods used do not assume a particular sensor type or configuration. Any type of range sensor could be used. Also, the analysis software is designed so that the sensor layout is read during the initialization process.

The first step in recognizing spatial relations from sensor readings is to build object representations from the readings. Let us consider a simple case of the robot and a single obstacle, shown in Figure 3. The sonar sensor S returns a range value indicating that an obstacle has been detected. In the case of Figure 3, only one obstacle was detected, and a single object representation is plotted as a trapezoid in the center of cone S. The depth of the obstacle cannot be determined from the sonar reading; thus, we use a constant arbitrary depth when building objects. We also represent the cylindrical robot as a rectangular object, because it is easier to process using vector data, since there are only 4 vertices in a rectangle. The bounding rectangle we build around the robot is also shown in Figure 3.

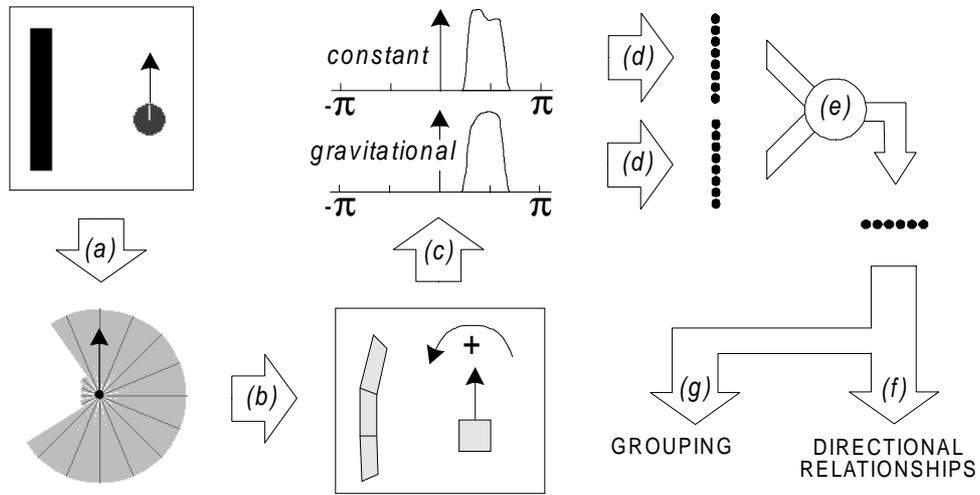


Figure 2. Synoptic diagram. (a) Sensor readings. (b) Construction of the polygonal objects. (c) Computation of the histograms of forces. (d) Extraction of numeric features. (e) Fusion of information, (f) Generated linguistic spatial descriptions for each object sensed, (g) Grouping of objects to generate a less detailed description.

In the case of multiple sonar returns, we examine the sonar readings that are adjacent to each other. There is a question on whether adjacent sonar readings are from a single obstacle or multiple obstacles. Our solution to this issue is to determine if the robot can fit between the points of two adjacent sonar returns. If the robot cannot fit between two returns, then we consider these returns to be from the same object. Even if there are actually two objects, they may be considered as one for robot navigation purposes. In the case that the distance between the two points of the sonar returns is big enough to allow the robot to travel through, we consider separate objects. To form objects from multiple sonar returns we join the centers of the corresponding sonar cones.

For example, consider the obstacle in Figure 4. Since the obstacle is relatively far from the robot, the distance between the sonar returns is rather big, and we cannot determine whether the obstacle continues between the three sonar readings, or we have three different obstacles. In this case, we plot three different objects until the robot gets closer to the obstacle and we have a better resolution of the obstacle, since more sensors would detect its presence.

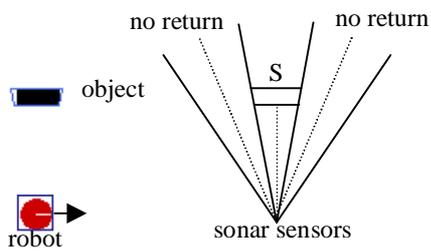


Figure 3. A single object is formed from a single sonar reading.

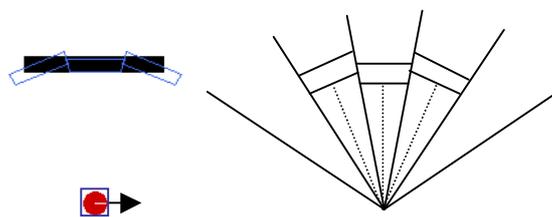


Figure 4. Three different objects are formed from 3 different sonar readings, if the readings are not “close” enough, according to the distance measure [11].

After building the objects around the robot based on the sonar sensor readings, we represent the relative position between each object and the robot by the histograms of constant and gravitational forces associated with the robot/object pair, as described in Section 2. We then generate an egocentric linguistic description, *i.e.*, from the robot’s point of view. Thus, the descriptions also depend on the robot’s orientation or heading. A change in robot heading is easily accomplished by shifting the histogram along its horizontal axis. Figure 5 shows an example of the linguistic expressions generated for the 5 objects detected. More details and examples can be found in [11].

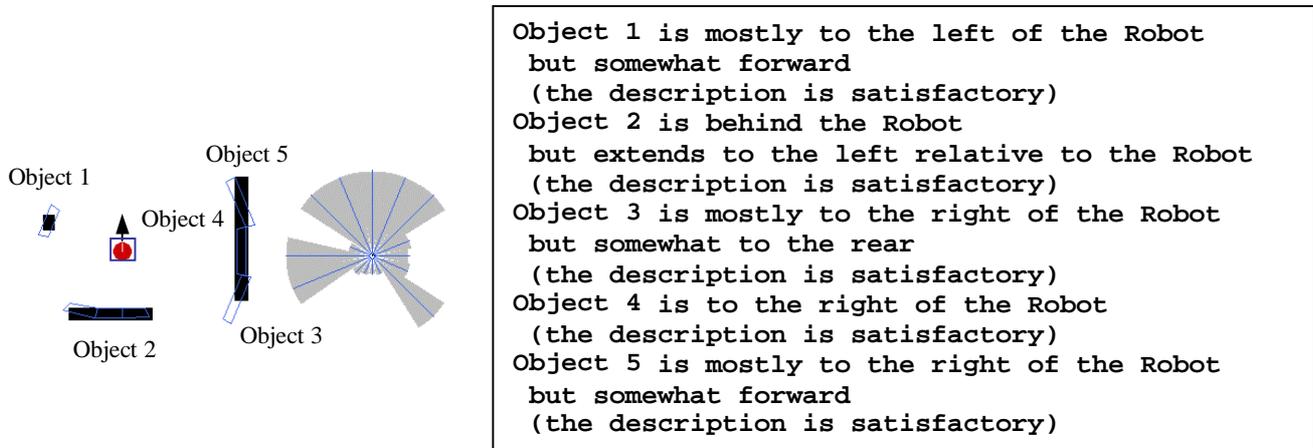


Figure 5. The robot detects 5 obstacles. Object representations are shown as plotted rectangles. The generated linguistic spatial descriptions are shown on the right.

4. INTERPRETING A SKETCHED MAP

The interface used for drawing the robot trajectory map is a PDA (e.g., a PalmPilot). The stylus allows the user to sketch a map much as she would on paper for a human colleague. The PDA captures the string of (x,y) coordinates sketched on the screen and sends the string to a computer for processing.

The user first draws a representation of the environment by sketching the approximate boundary of each object. During the sketching process, a delimiter is included to separate the string of coordinates for each object in the environment. After all of the environment objects have been drawn, another delimiter is included to indicate the start of the robot trajectory, and the user sketches the desired path of the robot, relative to the sketched environment. An example of a sketch is shown in Figure 6(a), where each point represents a captured (x,y) screen pixel.

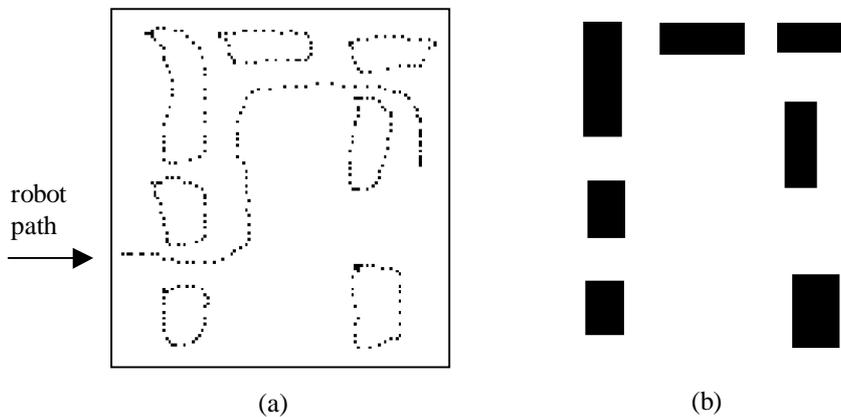


Figure 6. (a) The sketched map on the PDA used for experiments in Section 5. Environment objects are drawn as a boundary representation. The robot path starts from the left. (b) The corresponding environment defined using the robot simulator.

For each point along the trajectory, a view of the environment is built, corresponding to the radius of the sensor range. The left part of Figure 7 shows a sensor radius superimposed over a piece of the sketch. The sketched points that fall within the scope of the sensor radius represent the portion of the environment that the robot can sense at that point in the path.

The points within the radius are used as boundary vertices of the environment object that has been detected. To accommodate convex-shaped objects, an additional point on the sensor radius is included. Together, they define a polygonal region (Figure 7, step (a)) whose relative position with respect to the robot (assimilated to a square) is represented by the two histograms (Figure 7, step (b)): the histogram of constant forces and the histogram of gravitational forces, as described in Sec. 2. The heading is computed along the trajectory using a filtering algorithm that compensates for the discrete pixels.

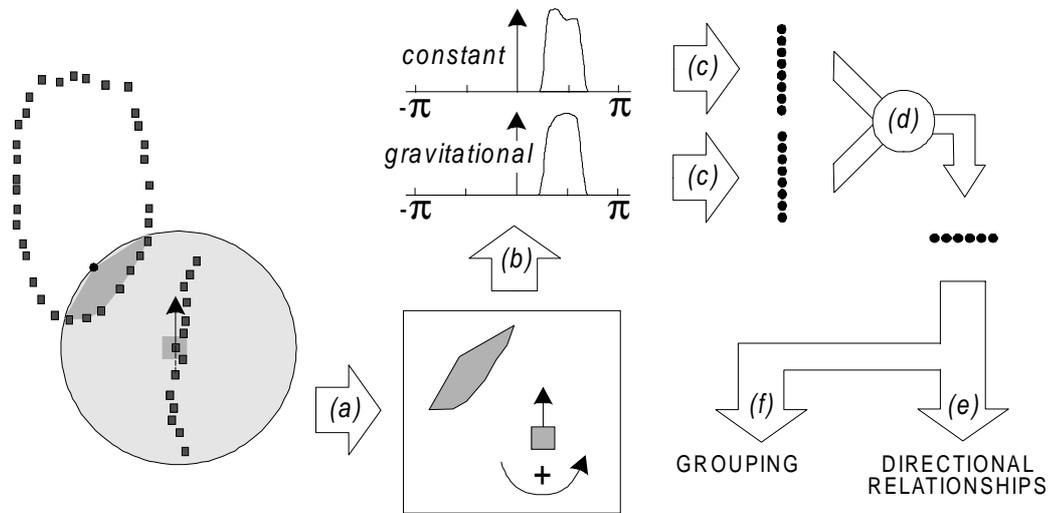


Figure 7. Synoptic diagram. (a) Construction of the polygonal objects. (b) Computation of the histograms of forces. (c) Extraction of numeric features. (d) Fusion of information, (e) Generated linguistic spatial descriptions for each object sensed, (f) Grouping of objects to generate a less detailed description.

The histograms of constant and gravitational forces associated with the robot and the polygonal region are then used to generate a linguistic description of the relative position between the two objects. The method followed is the same as that used for the sensor readings (Sec. 3). Figure 8 shows the linguistic description generated for a point on the robot path. As before, a three-part linguistic spatial description is generated for that point. See also [12] for details and more examples.

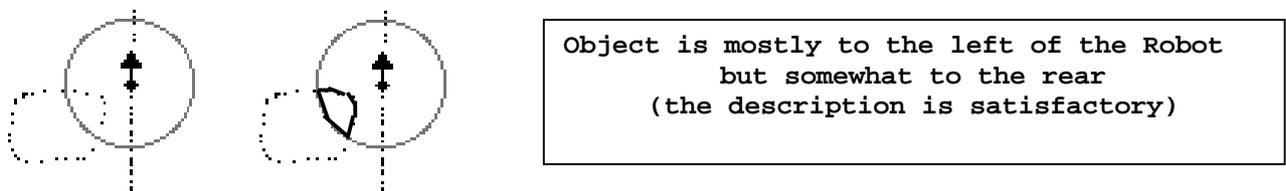


Figure 8. Building the environment representation for one point along the trajectory, shown with the generated, three-part linguistic expression.

5. EXPERIMENTS

To test the compatibility of the two methods for producing comparable linguistic expressions, we created an environment in the simulator and sketched an approximate representation on the PDA. The two representations are shown side by side in Figure 6. For the PDA sketch, environment objects are drawn using a boundary representation; the seven bounded figures represent the environment obstacles. The desired robot trajectory is sketched relative to the environment and shown in the Figure, starting from the left.

Using the methods described in Sections 3 and 4, the linguistic spatial descriptions are generated for corresponding robot trajectory points in both environments. Figures 9 through 14 show representative points along the trajectory. The PDA sketch is analyzed using a top-down view, but constrained by the effective radius of the robot's sensors, as shown in the figures. Only the portion of the object that falls within the sensor radius is used to generate the linguistic descriptions. The corresponding robot environment is analyzed using the simulated robot sonar sensors which provide an egocentric (relative) view from the robot's perspective. The object representations built from the sonar sensors are shown on the figures as overlaid trapezoids.

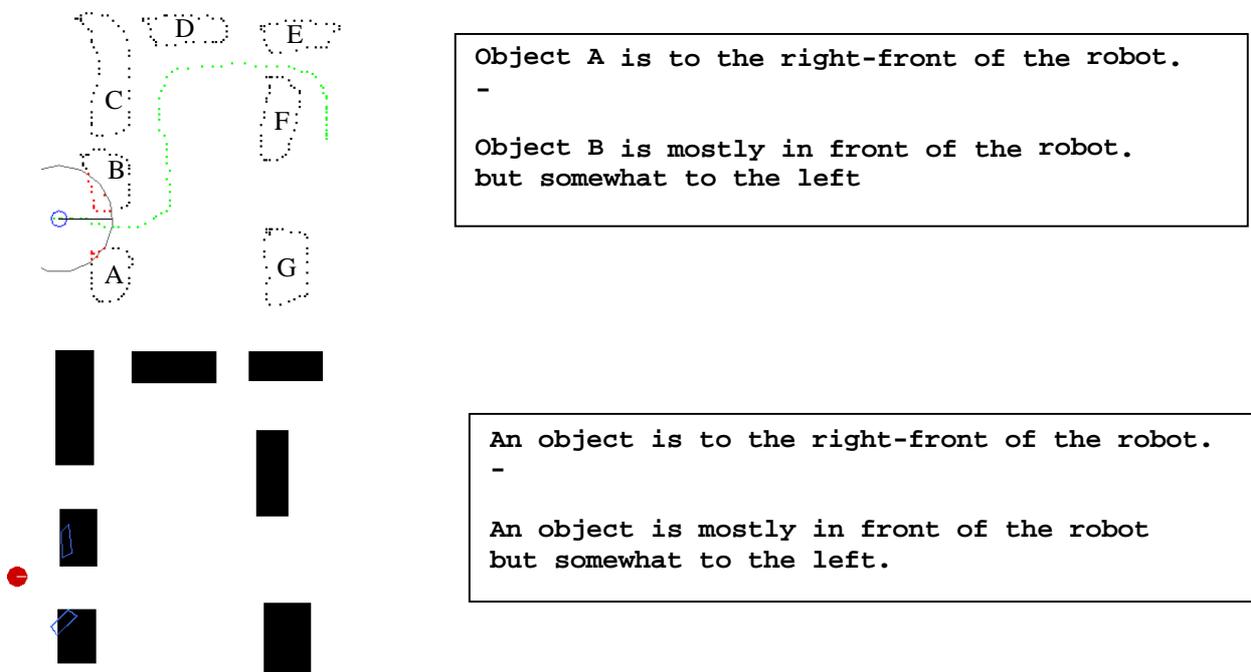


Figure 9. Position 1. The first point along the robot trajectory. The PDA sketch is shown on the top left with the effective sensor radius used for the experiments. The corresponding robot simulator view is shown on the bottom left with the object representations built from the sonar sensors overlaid as trapezoids. The generated linguistic spatial descriptions are shown on the right for each environment. Note the robot heading.

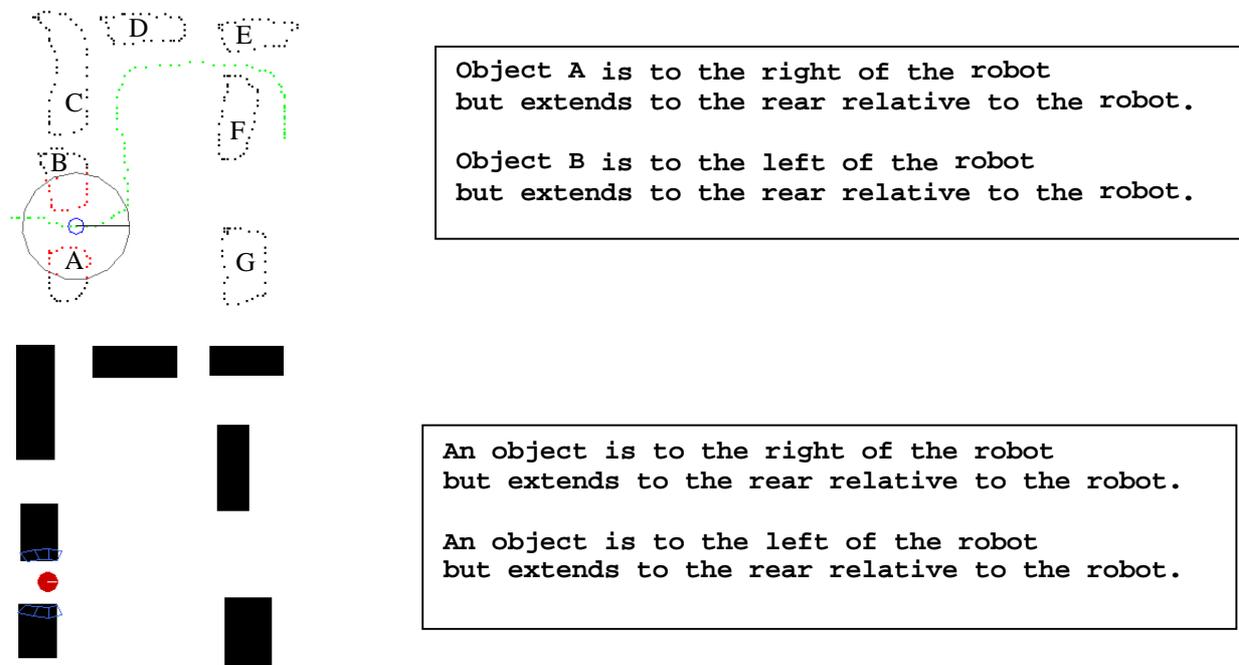
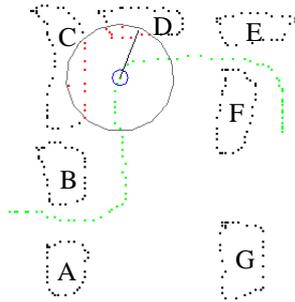
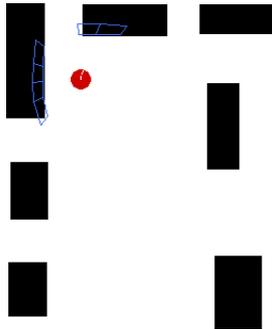


Figure 10. Position 2.



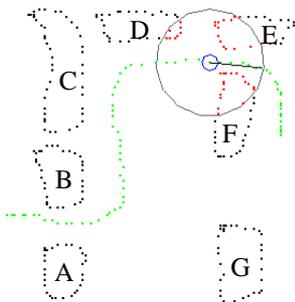
Object C is to the left of the robot
 -
 Object D is in front of the robot
 but extends to the left relative to the robot.



An object is to the left of the robot
 but extends to the rear relative to the robot.

 An object is in front of the robot.

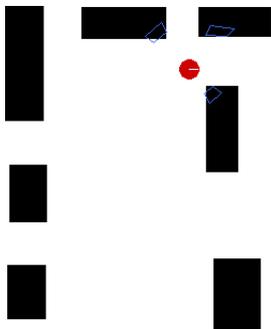
Figure 13. Position 5.



Object D is mostly behind the robot
 but somewhat to the left.

 Object E is mostly to the left of the robot
 but somewhat forward.

 Object F is to the right-front of the robot.
 -



An object is behind-left of the robot.
 -

 An object is mostly to the left of the robot
 but somewhat forward.

 An object is to the right-front of the robot.
 -

Figure 14. Position 6

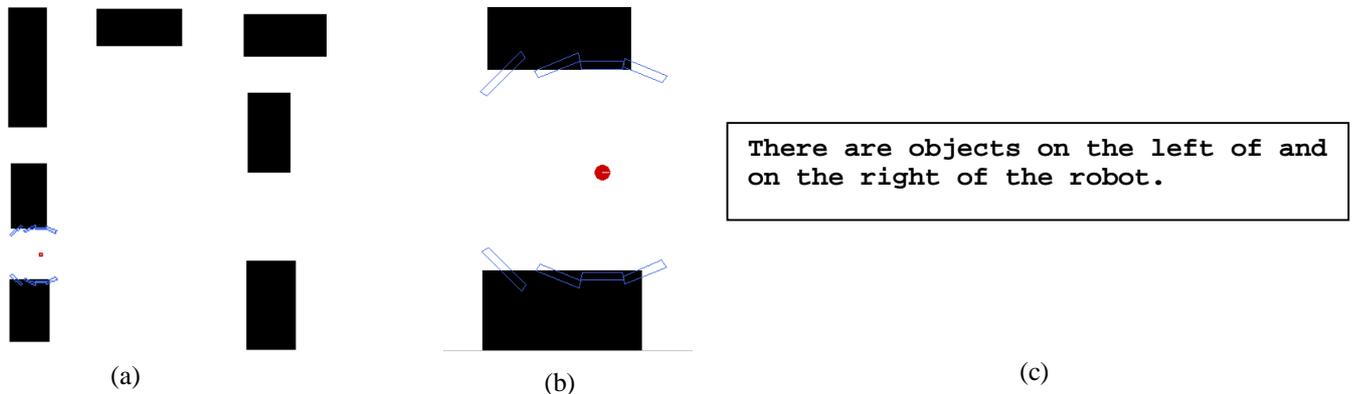


Figure 15. An example showing groups of objects sensed. (a) A top view of the environment (b) The robot senses 4 objects of the left and 4 objects on the right. (c) The high-level linguistic description generated, using object grouping.

The generated linguistic expressions in Figures 9 through 15 mostly agree between the two representations. However, in some cases, e.g., when there are a large number of objects in the environment, the description may be more detailed than necessary or even too detailed to be useful. We have also been developing a grouping algorithm that is used to generate a less detailed description. An example is shown in Figure 15, which is a view of the previous environment but scaled so that the robot is much farther from the obstacles. In this case, there are several objects sensed on both the left and right sides, 8 individual objects in total. The generated description is shown in Figure 15(c), which provides a higher level interpretation of the robot's situation. Details on the grouping algorithm will be discussed in a forthcoming paper [13].

6. CONCLUSIONS

In this paper, we have shown how the histogram of forces can be used to generate linguistic spatial descriptions representing the qualitative state of a mobile robot. We have described two ways in which spatial relations can be used for robot navigation. The robot can be a physical robot moving in an unknown environment with range sensors to interpret its environment, as well as a virtual robot whose environment and trajectory are sketched on a PDA. A boundary approximation of the obstacles is made, and their vertices are used as input to the histogram of forces. The approach is computationally efficient, and the spatial descriptions can be generated in real time.

We have presented an experiment in which a robot is placed in a physical environment, and a corresponding approximation is sketched on a PDA. The results show that the linguistic descriptions generated from the two different representations are comparable. This provides justification for this novel approach to human-robot interaction, namely showing a robot a navigation task by sketching an approximate map on a PDA. The approach represents an important step in studying the use of spatial relations as a symbolic language between a human user and a robot for navigation tasks.

As an extension, we are developing algorithms to incorporate other spatial relations, such as surrounds, and distance, such as *close* or *far*. The surrounds relation is determined directly from the histogram of forces. The distance descriptions are generated after processing the range information returned from the robot's sensors or the distances calculated from the PDA sketch. In some cases, a less detailed description is more useful, and we are also working on generating multi-level linguistic descriptions.

Future work may utilize linguistic spatial descriptions to facilitate natural communication between a human and a robot (or a group of robots). Image spatial analysis can be used to provide a direction relative to something in the image. For example, the user can issue instructions such as go to the *right* of the building, or go *behind* the building.

ACKNOWLEDGEMENTS

The authors wish to acknowledge support from ONR, grant N00014-96-0439 and the IEEE Neural Network Council for a graduate student summer fellowship for Mr. Chronis.

REFERENCES

1. W. Gribble, R. Browning, M. Hewett, E. Remolina and B. Kuipers, "Integrating vision and spatial reasoning for assistive navigation", In *Assistive Technology and Artificial Intelligence*, V. Mittal, H. Yanco, J. Aronis and R. Simpson, ed., Springer Verlag, , 1998, pp. 179-193, Berlin, Germany.
2. D. Perzanowski, A. Schultz, W. Adams and E. Marsh, "Goal Tracking in a Natural Language Interface: Towards Achieving Adjustable Autonomy", In *Proceedings of the 1999 IEEE International Symposium on Computational Intelligence in Robotics and Automation*, Monterey, CA, Nov., 1999, pp. 208-213.
3. F. Shibata, M. Ashida, K. Kakusho, N. Babaguchi, and T. Kitahashi, "Mobile Robot Navigation by User-Friendly Goal Specification", In *Proceedings of the 5th IEEE International Workshop on Robot and Human Communication*, Tsukuba, Japan, Nov., 1996, pp. 439-444.
4. M.J. Egenhofer, "Query Processing in Spatial-Query-by-Sketch", *Journal of Visual Languages and Computing*, vol. 8, no. 4, pp. 403-424, 1997.
5. J. Freeman, "The Modelling of Spatial Relations", *Computer Graphics and Image Processing (4)*, pp. 156-171, 1975.
6. K. Miyajima, 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.
7. P. Matsakis, *Relations spatiales structurelles et interprétation d'images*, Ph. D. Thesis, Institut de Recherche en Informatique de Toulouse, France, 1998.
8. 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.
9. P. Matsakis, J. M. Keller, L. Wendling, J. Marjamaa and O. Sjahputera, "Linguistic Description of Relative Positions in Images", *IEEE Trans. on Systems, Man and Cybernetics*, to appear.
10. J. M. Keller and P. Matsakis, "Aspects of High Level Computer Vision Using Fuzzy Sets", *Proceedings, 8th IEEE Int. Conf. on Fuzzy Systems*, Seoul, Korea, pp. 847-852, 1999.
11. M. Skubic, G. Chronis, P. Matsakis and J. Keller, "Generating Linguistic Spatial Descriptions from Sonar Readings Using the Histogram of Forces", to appear in the *Proceedings of the 2001 IEEE Intl. Conf. on Robotics and Automation*.
12. M. Skubic, P. Matsakis, B. Forrester and G. Chronis, "Extracting Navigation States from a Hand-Drawn Map", to appear in the *Proceedings of the 2001 IEEE Intl. Conf. on Robotics and Automation*.
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", in preparation for submission to *Autonomous Robots*.