

To appear in the *Journal of Robotic Systems*, Special Issue on Mobile Robots.

Place Recognition in Dynamic Environments

Brian Yamauchi¹ (yamauchi@robotics.stanford.edu)

Pat Langley² (langley@robotics.stanford.edu)

Institute for the Study of Learning and Expertise

2164 Staunton Court

Palo Alto, CA 94306

Phone: (415) 494-1588

¹ Currently at the Navy Center for Applied Research in Artificial Intelligence, Naval Research Laboratory, Washington, DC, 20375-5337.

² Also affiliated with the Robotics Laboratory, Computer Science Department, Stanford University, Stanford, CA, 94305.

Place Recognition in Dynamic Environments

Abstract

We have developed a technique for place learning and place recognition in dynamic environments. Our technique associates evidence grids with places in the world and uses hill climbing to find the best alignment between current perceptions and learned evidence grids. We present results from five experiments performed using a real mobile robot in a real-world environment. These experiments measured the effects of transient and lasting changes in the environment on the robot's ability to localize. In addition, these experiments tested the robot's ability to recognize places from different viewpoints and verified the scalability of this approach to environments containing large numbers of places. Our results demonstrate that places can be recognized successfully despite significant changes in their appearance, despite the presence of moving obstacles, and despite observing these places from different viewpoints during place learning and place recognition.

1 Introduction

Place learning and place recognition are two of the central issues in mobile robotics. Unless a robot has an absolute position reference (e.g. from a global positioning satellite), it needs some method to determine its current location using its sensors. Place learning consists of associating perceptions with locations in the world. Place recognition consists of matching current perceptions with those previously learned to determine the robot's current location.

Much research has been done on these topics, but most has been confined to environments that do not change. In contrast, most environments containing people do change, and change often. People move chairs and rearrange desks. They open closed doors and close open ones. A localization algorithm that depends upon an unchanging world is likely to fail in any environment containing human beings.

Our goal is to develop methods for place learning and place recognition that are robust to the types of changes that robots may encounter in human environments. We have developed a technique that associates evidence grids with places in the world and uses hill climbing to find the best alignment between current perceptions and the learned evidence grids.

This paper presents results from five experiments performed using a real mobile robot in a real-world environment. The first experiment measured the effects of lasting changes in the world upon place recognition, whereas the second study combined lasting changes with shifts in viewpoint. The third experiment measured how well our approach scaled to a large number of places and also how well it handled transient changes in the world. The fourth experiment focused on a particularly challenging environment, a hallway containing many regions of similar appearance. The final experiment involved the same hallway, but with changed viewpoints.

2 Related Work

Many researchers have studied place learning and place recognition for mobile robots. Their proposed spatial representations include Kuipers and Byun's distinctive places¹, Kortenkamp's visual scenes², Engelson's image signatures³, and Greiner and Isukapalli's landmarks.⁴ However none of this research has addressed the issue of place recognition in dynamic environments, where the appearance of places may change over time. Some of this work may be applicable—for example, Kortenkamp's visual scenes and Engelson's image signatures may remain

distinguishable despite other changes in the environment. However, none of these techniques have been tested in dynamic environments.

Leonard and Durrant-Whyte⁵ have developed methods for localization using sonar sensors to track the positions of environment features (planes, cylinders, corners, and edges) with extended Kalman filters. While their approach has been successful in simple environments, we believe our approach is better suited for complex dynamic environments where features are subject to frequent changes.

Schiele and Crowley⁶ report a method that estimates position based on matching line segments extracted from evidence grids using Hough transforms and Kalman filtering. However, their research has only dealt with static environments, and it is unclear how robust their techniques would be in dynamic environments.

Courtney and Jain⁷ describe an approach that extracts features from evidence grids built using sonar, vision, and infrared sensors, and uses these features for place recognition. However, their research was also limited to static environments, and identifications based on these features might not be stable in dynamic environments.

Thrun⁸ has used evidence grids for Cartesian position estimation in his RHINO system, but this approach assumes that walls will only be parallel or perpendicular to each other. Although this may hold for most indoor environments, obstacles can make it difficult to determine the actual orientation of walls. Our approach differs in localizing based upon all of the detected features of the environment, rather than relying upon a priori assumptions about the structure of the world.

Schultz and Grefenstette⁹ report results on continuous localization using evidence grids. In their work, local grids constructed by the robot are continuously registered with a global grid to determine the robot's Cartesian coordinates. Our work differs in studying the ability to recognize distinct places using separate grids.

In previous work¹⁰, we have used evidence grids for place learning in a static environment. This research, conducted mostly in simulation, used an exhaustive search over all possible translations to find the best alignment between the test grid and training grids, and did not handle rotational alignment. The work we describe in this paper differs in dealing with dynamic environments in the real world, handling rotational as well as translational alignment, and using a hill-climbing algorithm to more efficiently search the space of possible transformations.

In other previous research¹¹, we have also developed a technique for using evidence grids, along with a hill-climbing algorithm for alignment, to correct errors in dead reckoning in a dynamic real-world environment. However, that localization procedure dealt only with a single location and did not address learning and recognizing multiple places.

3 Place Learning

The localization system described in this paper is the newest component of ELDEN (Exploration and Learning in Dynamic ENvironments), an integrated mobile robot system developed for exploration, learning, and navigation in dynamic, real-world environments.¹² Place learning consists of building an evidence grid for a region in space and associating it with a place in the environment. Each place is represented as a node within a topological/metric map, and each node stores the Cartesian location of the corresponding place. The topological component is included for navigation purposes, even though it is not used for place recognition.

3.1 Evidence Grids

3.1.1 Prior Probabilities and Sensor Models

Evidence grids are a spatial representation developed by Moravec and Elfes.¹³ Space is represented as a Cartesian grid where each cell has a certain probability of being occupied. Initially, each of these cell probabilities is set to the estimated prior probability of cell occupancy. For example, if one quarter of the space in a given area is occupied, one might set the prior probability to 0.25. (In practice, evidence grids tend to be insensitive to errors in the prior probability, and an estimate of 0.5 generally works well.)

Each time the robot receives a sensor input, the evidence grid is updated using the corresponding sensor model. Each sensor model describes the probability that cells are occupied given the reading received. This model depends on the characteristics of the individual sensor. One of the major advantages of the evidence grid representation is its ability to fuse sensor information. Any number of sensor readings from any number of sensors can be combined as long as models exist for each sensor type.

3.1.2 Updating Evidence Grids

Formally, evidence grids provide a means for combining information from sensor readings in the following way.¹⁴ If X represents information such as a sensor reading, then $p(o|X)$ is the probability that a cell is occupied given X , and $p(\neg o|X)$ is the probability that this cell is not occupied given X . Then, from Bayes' theorem:

$$\frac{p(o|X)}{p(\neg o|X)} = \frac{p(X|o)}{p(X|\neg o)} \times \frac{p(o)}{p(\neg o)}$$

where $p(X|o)$ is the probability of receiving information X given that this cell is occupied, $p(X|\neg o)$ is the probability of receiving information X given that this cell is not occupied, $p(o)$ is the prior probability that any given cell is occupied, and $p(\neg o)$ is the prior probability that any given cell is unoccupied, $p(\neg o) = 1 - p(o)$.

If A represents the current state of the grid and B represents the information from a new sensor reading, then the cell occupancy probabilities can be combined using the equality:

$$\frac{p(o|A \cap B)}{p(\neg o|A \cap B)} = \frac{p(o|A)}{p(\neg o|A)} \times \frac{p(o|B)}{p(\neg o|B)}$$

This makes the approximation that A and B represent independent information, which is not true when a particular point can be sensed more than once (by the same or different sensors). In practice, this approximation means that the overall occupancy results tend to be accurate, but the numerical occupancy probabilities are not reliable. For example, if the sonar cones overlap for two sensor readings, the cells in the overlap will have their probabilities increased or decreased twice, as if the two sensor readings provided independent information about the structure within this region.

Konolige¹⁵ presents one approach to dealing with this problem. In this method, pose information is stored with each cell, indicating the incident direction of each sonar reading. Only the first sonar reading from a particular direction is considered for each cell—subsequent readings are ignored. This approach works well in static environments, but is not well-suited to dynamic environments, since the early state of the world will become "frozen" into the grid, and the grid will not be updated to reflect future changes that occur in the world. Instead we choose to accumulate multiple sensor readings over time, using the standard evidence grid formulation, and then we design our grid matching function to be tolerant to the uncertainty in cell occupancy probabilities.

3.1.3 Evidence Grid Advantages for Dynamic Environments

Accumulating multiple readings over time is an effective method of filtering out transient changes. Consider a person walking past the robot as it maps a particular region of space. This person's path will cover many grid cells, but each only for a brief moment. Each sonar reading that reflects from the person will increase the occupancy probability of the corresponding cells. However, each cell will only be occupied briefly, so all of the other sonar readings incident on this cell will reduce its occupancy probability. As a result, the cells along this path will have a low occupancy probability despite the person's passage.

In addition to providing an effective method for combining data from multiple sensor readings, evidence grids have two other advantages for use in dynamic environments. First, they can be updated quickly. Using a logarithmic transformation of the equations described above, each cell update can be computed with a single addition. Second, small changes in the environment tend to result in small changes to the corresponding grid representation. This property is important for dealing with lasting changes in the environment.

One exception to the second property is the case of specular reflections, which occur when a sonar pulse hits a flat surface and reflects away from (rather than back to) the sensor. As a result, the sensor registers a range that is substantially larger than the actual range. Because of this, a small change in the angle of a surface could potentially result in a substantial change to the evidence grid. Konolige¹⁵ also suggests a method for dealing with specular reflections by ignoring all sonar readings if they would imply that previously occupied cells are unoccupied (as would occur if a specular reflection were to overlap an obstacle). However, this would not work for dynamic environments, since a previously occupied space may actually have become unoccupied due to changes in the world. Instead, during the construction of each evidence grid, we rotate the sonar sensors through a range of angles equivalent to the width of the sonar arc. As a result, if both specular and non-specular reflections are possible from a given viewpoint, then both will be incorporated into the evidence grid.

3.1.4 Evidence Grid Specifics

We used a Nomad 200 mobile robot in our research. This robot is equipped with 16 sonar sensors, evenly spaced around the base at 22.5 degree intervals. In order to build each evidence grid, the robot remained at the center of the place region and took eleven sets of sixteen sonar readings at two degree intervals (for a total of 176 sonar readings for each grid).

Sonar sensors have a number of well-known limitations. The most severe is the problem with specular reflections described above. Other limitations include the speed of sound, which restricts the firing rate of any given sensor by the maximum time required for the pulse to travel to a distant obstacle and return, and the requirement that other sonar sensors remain inactive while the current pulse is in flight.

Despite these limitations, we decided to use sonar because these sensors are relatively low-cost and available for a wide variety of mobile robot platforms, thus providing results that would be relevant to the largest number of researchers and developers of mobile robots. We plan future work with more precise, albeit more expensive and less widely available, sensors such as laser rangefinders.

In our research, each evidence grid contained a 64 x 64 matrix of cells representing an area 30 feet by 30 feet. The grid size was selected to correspond to the area contained within the effective range of the sonar sensors (15 feet) as viewed from the center of each place region. Each cell corresponds to an area about half a foot wide, which is sufficient resolution to represent most of the significant features (e.g. walls, desks, chairs) within a typical office environment. Future experiments are planned to measure the effect of cell size on recognition accuracy. Each cell is represented by a single byte, using a logarithmic scale, so the total memory required for each grid is 4096 bytes.

4 Place Recognition

Place recognition consists of building a new evidence grid at the robot's current location (the recognition grid) and matching this grid against all of the grids that have been previously associated with places in the world (the learned grids). The recognition grid is translated and rotated to find the best match with each learned grid. A multiresolution hill-climbing algorithm searches the space of possible translations and rotations.

We designed this method to recognize a place from a number of different positions and orientations within that place. Shifts in position are handled by translating the recognition grid, whereas shifts in orientation are handled by rotating the recognition grid.

We define translations and rotations over evidence grids in the following way. The origin of the coordinate frame is located at the center of each grid, corresponding to the robot's position when the grid was constructed. Each cell in the recognition grid is translated by displacing the point corresponding to the center of

each cell and determining into which cell the new point would fall in the learned grid. Each cell in the recognition grid is rotated by computing the vector from the origin to the center of the cell, then rotating this vector around the origin, and determining into which cell the new vector would fall in the learned grid.

A match score is computed for each pair of corresponding cells (one in the recognition grid, the other in the learned grid). This match metric is given by:

$$s(i,j) = \begin{cases} 1 & \text{if } p(i) > p_0 \text{ and } p(j) > p_0 \\ 1 & \text{if } p(i) < p_0 \text{ and } p(j) < p_0 \\ 1 & \text{if } p(i) = p_0 \text{ and } p(j) = p_0 \\ 0 & \text{otherwise} \end{cases}$$

where $s(i,j)$ is the match score for corresponding cells i and j , $p(i)$ is the probability that cell i is occupied, $p(j)$ is the probability that cell j is occupied, and p_0 is the prior probability that any cell is occupied. This score is summed over all of the corresponding cells, and the total is the match score for the learned grid for the current transformation.

We developed this match metric to deal with the problem of non-independent sensor readings. Since the sonar cones overlap, their sensor readings are not independent. As a result, the occupancy probabilities in the evidence grid do not accurately reflect the precise probability that each cell will be occupied. However, what is reliable is whether each cell is more likely or less likely to be occupied than the prior probability (or whether it has not been sensed at all, in which case it will be equal to the prior probability). Thus, the match metric increases the match score whenever two corresponding cells are either both more likely to be occupied, less likely to be occupied, or unsensed in both the recognition grid and the learned grid.

The hill-climbing algorithm applies this process iteratively to find the best transformation between the recognition grid and each learned grid. The hill-climbing stepsize is halved when a local maximum is reached, in order to more precisely locate this maximum. When a local maximum is reached using the minimum step size, the search is stopped and the score for the current transformation is used as the overall match score for the learned grid. This process is repeated for each of the learned grids, and the grid with the maximum match score is selected as the winner.

Place recognition is performed using an offboard workstation that communicates with the robot through a radio ethernet. Offboard computation permits the use of more powerful computational hardware without adding to

the robot's onboard power requirements. In addition, the offboard workstation presents a graphical interface that allows users to issue commands to the robot and to visualize the spatial structure of the grids representing places learned by the robot. The radio ethernet provides a reliable, high bandwidth, low latency communication link between the workstation and the robot.

In experiments 1-3, an onboard compass was used to determine orientation. The value returned from the compass was not accurate in an absolute sense, but was usually repeatable (± 10 degrees) for a given location. In experiments 4 and 5, the robot was given an approximate estimate of its initial orientation. This estimate did not need to be precise. Typically, place recognition is able to compensate for errors of up to 45 degrees. In situations where an initial orientation estimate would be impossible to obtain, another option would be to have the robot perform a series of hill-climbing registrations from a set of initial angles (e.g. eight angles offset at 45 degree intervals), and then select the particular transformation that generates the maximum overall match score.

5 Experiment 1: Lasting Changes

We designed our first experiment to measure the effects of environmental modification on the robot's ability to recognize previously learned places. The robot constructed grids for five places in a real-world office environment. These places contained many different types of obstacles, including chairs, tables, bookshelves, boxes, workstations, and other robots. We changed each of these places by adding new obstacles (office swivel chairs). We placed each new obstacle approximately six feet from the robot, spaced with roughly even angular separations (90-180 degrees, as permitted by the positions of the existing obstacles). Recognition grids were constructed for each place with one, two, and three new obstacles.

Figure 1 shows the original learned grid for one of the places (adjoining both open laboratory space and a hallway) and the corresponding recognition grids with the addition of one, two, and three new obstacles. The positions of these new obstacles are circled. Cells with occupancy probabilities greater than the prior probability of occupancy are represented by small circles. Cells with occupancy probability equal to the prior probability are represented by dots. Cells with occupancy probability less than the prior probability are represented by white space.

In order to study the effectiveness of the hill-climbing matching procedure, we compared it to a simple nearest-neighbor algorithm that returns the learned grid with the highest match score without any translation or rotation. We made this comparison in order to measure the effectiveness and the robustness of the hill-climbing algorithm. If hill climbing performed no better than nearest neighbor for the trials involving new viewpoints, this would imply that it was not effective at searching the space of grid transformations. If hill climbing performed better than nearest neighbor for trials without additional obstacles, but this advantage disappeared as new obstacles were added, then this would indicate that the hill-climbing algorithm was not robust to changes in the environment.

We define a *trial* to consist of matching a recognition grid against all of learned grids. If the learned grid with the highest score corresponded to the correct place, we considered the trial successful. If a learned grid for another place had a higher score, we considered the trial unsuccessful. Each of the 15 recognition grids (5 places \times 3 grids) was matched against each of the five learned grids for a total of 75 scored grid matches for each algorithm (nearest neighbor and hill climbing). Each trial (five grid matches) required about one second for the nearest-neighbor algorithm and about 20 seconds for the hill-climbing algorithm, executing on a Decstation 3100.

In this experiment, both the nearest-neighbor algorithm and the hill-climbing algorithm were able to perform place recognition with 100% accuracy for all five places. On every trial, the recognition grid matched the correct learned grid better than any of the other learned grids. That the nearest-neighbor algorithm performed as well as hill climbing was not surprising, given that place recognition was conducted from the same viewpoint as place learning, so it was not necessary to transform the corresponding grids. However, it was useful to learn that hill climbing did not introduce "false positives" by transforming the recognition grid to match one of the wrong learned grids.

6 Experiment 2: Changed Viewpoints

We performed the second experiment to measure the effects of environmental changes on the robot's ability to recognize places from a different viewpoint than the one in which these places were originally learned. For each place, the robot learned a grid. Then we moved the robot to a new viewpoint (two feet away), and the robot constructed a recognition grid.

With the robot at the new viewpoint, we added new obstacles (swivel chairs) in the same manner as the previous experiment, and additional recognition grids were constructed for each place with one, two, and three new obstacles. Figure 2 shows the learned grid constructed at the initial viewpoint for one of the places, along with the corresponding recognition grids from the new viewpoint with no new obstacles, with one new obstacle, and with three new obstacles. The locations of the new obstacles are circled.

Each trial consisted of matching a recognition grid against each of the learned grids. Trials were successful if the learned grid with the best score corresponded to the correct place. Each of the 20 recognition grids (5 places \times 4 grids) was matched against each of the five learned grids for a total of 100 scored grid matches for each algorithm.

Figure 3 shows the recognition accuracy of the two matching algorithms as a function of the number of new obstacles added. The nearest-neighbor algorithm was able to recognize 80% of the places from the new viewpoint with no added obstacles, but its performance dropped rapidly as new obstacles were introduced, recognizing only 60% of the places with one new obstacle, and only 20% of the places with two or three new obstacles. In contrast, the hill-climbing algorithm was able to correctly recognize 100% of the places from the new viewpoint with no new obstacles and was also able to correctly recognize 80% of the places from the new viewpoint with one, two, or three new obstacles.

7 Experiment 3: Scaling and Transient Changes

Although the first two experiments provided promising results for a small number of places, an important issue is how well this technique scales to many places. In order to address this issue, we conducted a third experiment during which the robot mapped 47 different places within an office environment. In general, the number of places required for a given environment will depend on both its size and its complexity. A useful heuristic is to control the spacing between places so that a place exists for every destination to which a user may want to direct the robot and a place also exists for every branch point that may require the robot to select between paths to alternative destinations.

In this study, each place corresponded to a region five feet in diameter. We selected this place size so that the topological map could represent the traversable paths through the environment for a robot that is roughly two feet in diameter.

Initially the robot starts with an empty map. The robot's starting location becomes the first place in the new map. As the robot moves through the world, a new place is created whenever the robot moves out of the space contained in the existing place regions. A topological link is created between the new place and the place corresponding the robot's previous location. A new evidence grid is created and associated with the new place unit.

The environment for this experiment consisted of a large open area containing chairs, tables, desks, bookshelves, workstations, and bicycles, bordered by walkways and surrounded by offices. Dynamic change was present in both transient and lasting forms. Transient changes were caused by people moving through the environment, during both place learning and place recognition. Lasting changes occurred when people rearranged chairs, added and removed obstacles (i.e. bicycles), and opened and closed doors.

Figure 4 shows the topological/metric map constructed as the robot moved through the environment. This figure shows the place locations along with the topological links connecting these places. A total of 47 places were learned and an evidence grid was constructed for each place. The time required to build each evidence grid was approximately thirty seconds. The time required for place recognition was approximately five minutes (including the time required to build the recognition grid) using a Decstation 3100, with most of the time spent in the grid matching procedure. We have recently transferred this system to faster hardware, which substantially reduces the time required for place recognition, as described in the next section.

Figure 5 shows the learned grid (a) and recognition grid (b) for place 26. On the left side of this area is a wall containing open doorways leading to offices. On the right side is a large open area containing chairs, desks, and workstations. The clear area in the lower-left corner of this area is actually a (permanent) specular reflection caused by a whiteboard. This surface is sufficiently smooth that it acts as a mirror for the sonar, consistently reflecting all of the beams originating near the center of this area. In this case, these reflections can actually be useful as a distinguishing feature of this place, but only if the place regions are sufficiently small that the angle of reflection is similar during learning and recognition.

People walked past the robot during both place learning and place recognition, but the use of multiple sensor readings allowed the corresponding transient changes to be filtered out of these grids. The chairs on the right side of the room were rearranged between the times that these two grids were constructed, and in addition, a bicycle (not present in the learned grid) was placed in the upper-central region of the area during place recognition. Despite all of these changes, the localization system was able to correctly match the recognition grid with the learned grid.

In order to measure the effects of larger lasting changes, we removed the whiteboard that was causing the large specular reflection in the learned grid. As a result, the robot detected the wall itself rather than a specular reflection (figure 5(c)). In spite of the substantial difference between the learned grid and the new recognition grid, the place recognition system was still able to identify the robot's current location.

Overall, the robot was able to localize itself accurately throughout the environment. In about 90% of the places the robot was able to localize itself with 100% accuracy, always determining the correct place. In the remaining places, the robot was able to localize itself correctly roughly 75% of the time.

8 Experiment 4: Hallways with Lasting and Transient Changes

In the previous experiment, the presence of the central open area aided in disambiguating place locations. Hallways provide a more challenging environment, because they contain many places that appear similar. The fourth experiment tested the robot's ability to localize within a hallway, in the presence of both transient and lasting changes. The width of this hallway varied from four to six feet, and the length of this hallway was approximately 125 feet.

In this study, the robot learned 21 places within an office environment, including ten places in front of office doors, six places adjacent to posts in the hallway, three places adjacent to an open area, and two places next to hallway exits. This hallway contained large amounts of metal, electrical equipment, and wiring, resulting in magnetic fields that could change orientation 180 degrees over the space of four feet. As a result, the compass was not a reliable sensor for this environment. Instead, we provided the robot with an initial estimate of its orientation (to within roughly five degrees), and used dead reckoning to determine the relative orientation of adjacent places.

The carpeted floor introduced substantial errors into dead reckoning, with the result that the robot's orientation estimate at both ends of the hall was offset approximately 30 degrees from the robot's actual orientation.

To test the robot's ability to localize in this environment, the robot was placed at the center of each place location, a new evidence grid was constructed, and the best match was found between this grid and the learned grids using both nearest neighbor and hill climbing. This experiment was conducted in the presence of both transient changes and lasting changes. People walked past the robot during both place learning and place recognition. People also opened and closed office doors between the construction of learned and recognition grids.

Figure 6 shows the topological/metric map learned for this hallway. The orientation error is clearly visible in the curvature of the map—the actual hallway is straight. This error was also visible in the orientation of the grids. Figure 7 shows the learned grid (a) and recognition grid (b) constructed for place 9.

Although nearest neighbor performed well for places near the center of the hall, where the orientation error was small, it had some difficulty recognizing places with large amounts of orientation error. Overall, nearest neighbor could recognize 86% of the places (18 out of 21). Hill climbing was able to compensate for the substantial error in orientation throughout all regions of the hallway. Hill climbing was able to recognize 100% of the places (21 out of 21).

We used a Decstation 5100 for this experiment. Each grid was constructed in about ten seconds. The nearest-neighbor matching procedure required less than three seconds, while the hill-climbing match procedure required around 45 seconds total to match the recognition grid against all 21 learned grids.

9 Experiment 5: Hallways and Changed Viewpoints

We conducted the final experiment in the same hallway as the previous experiment. We used the same places and the same learned grids, but we changed the viewpoints used for building recognition grids by offsetting the robot's position by two feet along the hallway axis. The goal of this study was to measure the robot's ability to localize when positioned at locations other than those used to construct the learned grids. For four of the locations, this put the robot equidistant between two place locations, and for these locations, we considered either place to be a correct localization. As before, transient and lasting changes were present during this experiment.

In addition to measuring the accuracy of the places recognized, we also measured the accuracy of the Cartesian position returned from the hill-climbing match procedure. The transformation used to align the recognition grid with the learned grid was added to the stored Cartesian location of the identified place, and the result was used as the robot's estimate of its Cartesian position.

Figure 8 shows the learned grid (a) and recognition grid (b) for place 14. In this case, hill climbing was able to localize correctly, while nearest neighbor incorrectly matched this recognition grid with the learned grid for place 10 (figure 8(c)). Figure 9 shows the learned grid (a) and recognition grid (b) for place 16; in this case, both nearest neighbor and hill climbing confused the recognition grid with the learned grid for place 19 (figure 9(c)).

Overall, hill climbing performed substantially better than nearest neighbor in this experiment. Nearest neighbor was able to recognize 48% of the places (10 out of 21). Hill climbing was able to recognize 71% of the places (15 out of 21). In the 14 out of the 15 cases where hill climbing identified the correct place, it was able to estimate the robot's position to within one foot of its actual position; in the remaining case the error was 1.5 feet. The average Cartesian error was only 0.4 feet. This is sufficiently accurate to allow the robot to determine the best path toward a given destination place. When combined with a behavior-based approach for low-level obstacle avoidance, this can provide the robot with the capability to navigate robustly in a dynamic environment.

Specular reflections were the main cause of place misidentification. Translating the robot not only shifted the view, it also changed the reflections that were visible to the robot. In figure 9, for example, three of the specular reflections that were visible from the original viewpoint (a) disappeared when the robot was moved (b). We expect that using a different sensor, such as a laser rangefinder, would significantly increase recognition accuracy. We plan to test our technique using such a sensor in the near future.

10 Conclusions

We have developed a technique for place learning and place recognition in dynamic environments that involves storing and matching evidence grids. Our method uses hill climbing to find the best alignment between a grid that describes the current surroundings and a grid stored during learning, repeats this process for each stored grid, and selects the best grid and its associated alignment.

We have tested this technique in a series of experiments using a real robot in an unmodified real-world office environment. These experiments have shown that this technique is robust to transient changes, lasting changes, and changes in viewpoint, and that it scales well to environments containing many places. In addition, our studies demonstrated that the method performs well in environments such as hallways that contain many places with similar appearance, though specular reflections from sonar, combined with changes in viewpoint, cause errors in some cases. Finally, our experiments suggested that, in addition to robust place recognition, the technique provides accurate estimates of the robot's Cartesian position.

We have integrated this place recognition system with the adaptive topological path planner we developed for ELDEN.¹² The integrated system is capable of localizing robustly in dynamic environments, as described in this paper. This system is also capable of navigating through a changing world by constantly adapting its topological representation to reflect new observations and determining alternative paths when unexpected obstacles are encountered.

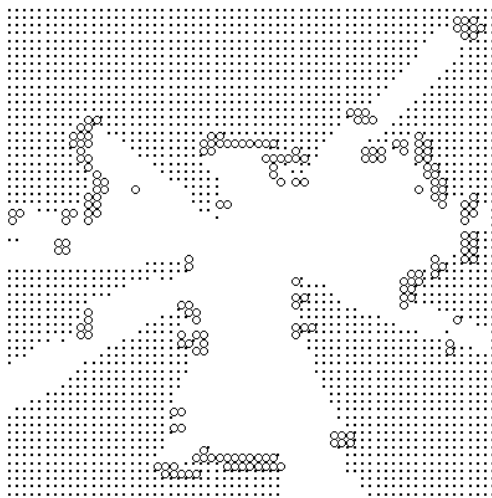
11 Acknowledgments

We thank Hans Moravec and Bill Adams for sharing their software, Alan Schultz and John Grefenstette for their useful comments and discussions, and Nils Nilsson and Barbara Hayes-Roth for providing access to a Nomad mobile robot. This research was supported in part by grant number N00014-94-1-0505 from the Office of Naval Research.

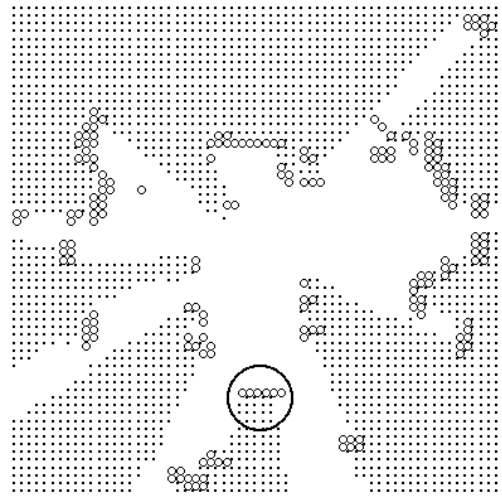
References

1. B. Kuipers and Y. Byun, "A robot exploration and mapping strategy based on a semantic hierarchy of spatial representations," *Journal of Robotics and Autonomous Systems*, **8**, 47-63, 1993.
2. D. Kortenkamp, *Cognitive Maps for Mobile Robots: A Representation for Mapping and Navigation*, Ph.D. Thesis, Electrical Engineering and Computer Science Department, University of Michigan, 1993.
3. S. Engelson, *Passive Map Learning and Visual Place Recognition*, Ph.D. Thesis, Department of Computer Science, Yale University, 1994.
4. R. Greiner and R. Isukapalli, "Learning to select useful landmarks," *IEEE Transactions on Systems, Man, and Cybernetics-Part B, Special Issue on Robot Learning*, **26**, 437-449, 1996.

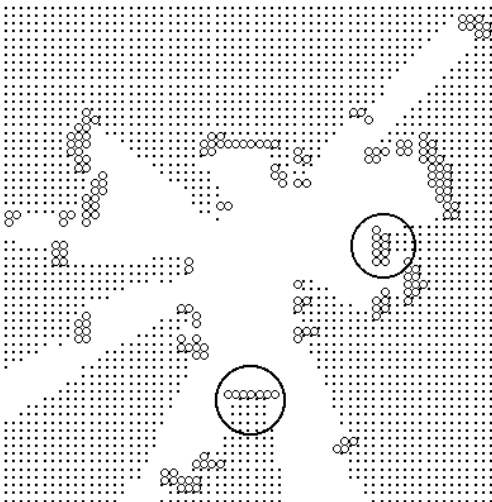
5. J. Leonard and H. Durrant-Whyte, *Directed Sonar Sensing for Mobile Robot Navigation*, Kluwer Academic Publishers, Norwell, MA, 1992.
6. B. Schiele and J. Crowley, "A comparison of position estimation techniques using occupancy grids," *Robotics and Autonomous Systems*, **12**, 163-171, 1994.
7. J. Courtney and A. Jain, "Mobile robot localization via classification of multisensor maps," *Proceedings of the IEEE International Conference on Robotics and Automation*, San Diego, CA, 1994, pp. 1672-1678.
8. S. Thrun, "The Mobile Robot RHINO," *AI Magazine*, **15**, 31-38, 1994.
9. A. Schultz and J. Grefenstette, "Continuous localization using evidence grids," NCARAI Technical Report AIC-95-024, Naval Research Laboratory, Washington, DC, 1995.
10. P. Langley and K. Pflieger, "Case-based acquisition of place knowledge," *Proceedings of the Twelfth International Conference on Machine Learning*, Tahoe City, CA, 1995, pp. 244-352.
11. B. Yamauchi, "Mobile robot localization in dynamic environments using dead reckoning and evidence grids," *Proceedings of the IEEE International Conference on Robotics and Automation*, Minneapolis, MN, 1996, pp. 1401-1406.
12. B. Yamauchi and R. Beer, "Spatial learning for navigation in dynamic environments," *IEEE Transactions on Systems, Man, and Cybernetics-Part B, Special Issue on Learning Autonomous Robots*, **26**, 496-505, 1996.
13. H. Moravec and A. Elfes, "High resolution maps from wide angle sonar," *Proceedings of the IEEE International Conference on Robotics and Automation*, St. Louis, MO, 1985, pp. 116-121.
14. H. Moravec, "Sensor fusion in certainty grids for mobile robots," *AI Magazine*, **9**, 61-74, 1988.
15. K. Konolige, "A refined method for occupancy grid interpretation," *Proceedings of the International Workshop on Uncertainty in Robotics*, Amsterdam, Netherlands, 1995.



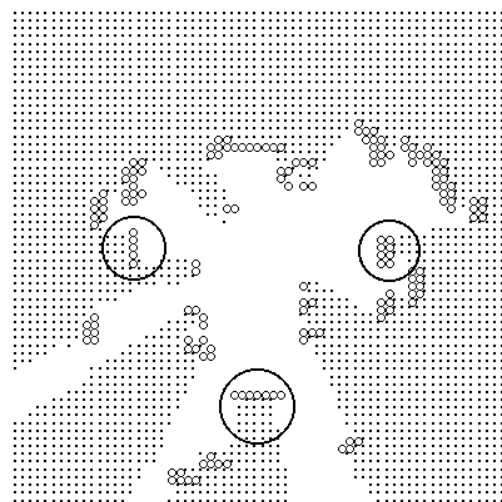
(a)



(b)

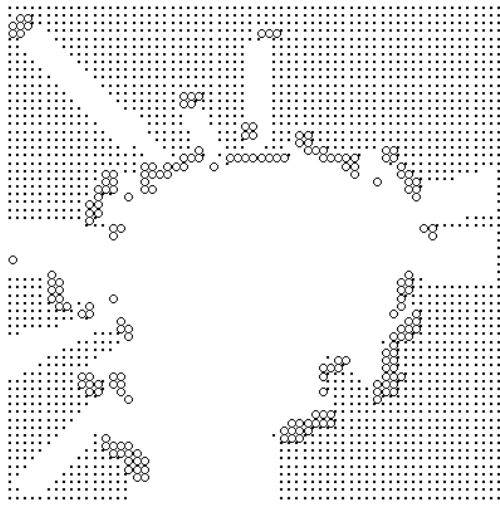


(c)

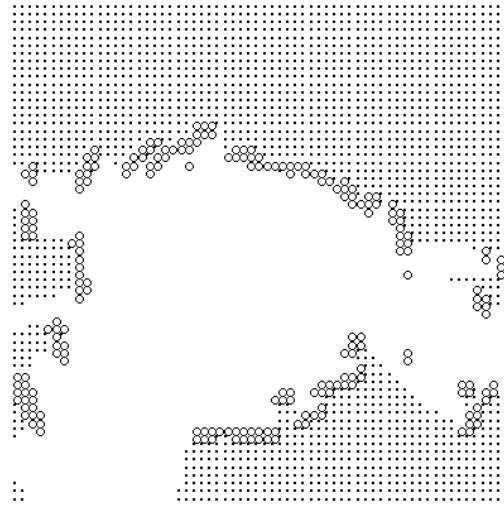


(d)

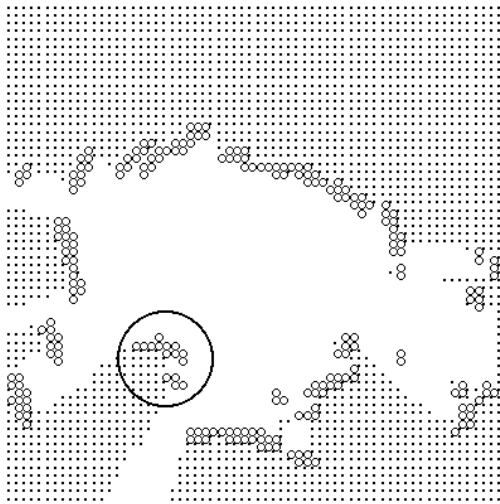
Figure 1: Learned grid for place 0 (a) and recognition grids for place 0 with one (b), two (c), and three (d) new obstacles. The locations of the new obstacles are circled.



(a)



(b)



(c)



(d)

Figure 2: Learned grid for place 4 (a) and recognition grid for place 4 with no new obstacles (b), one new obstacle (c), and three new obstacles (d). The locations of the new obstacles are circled.

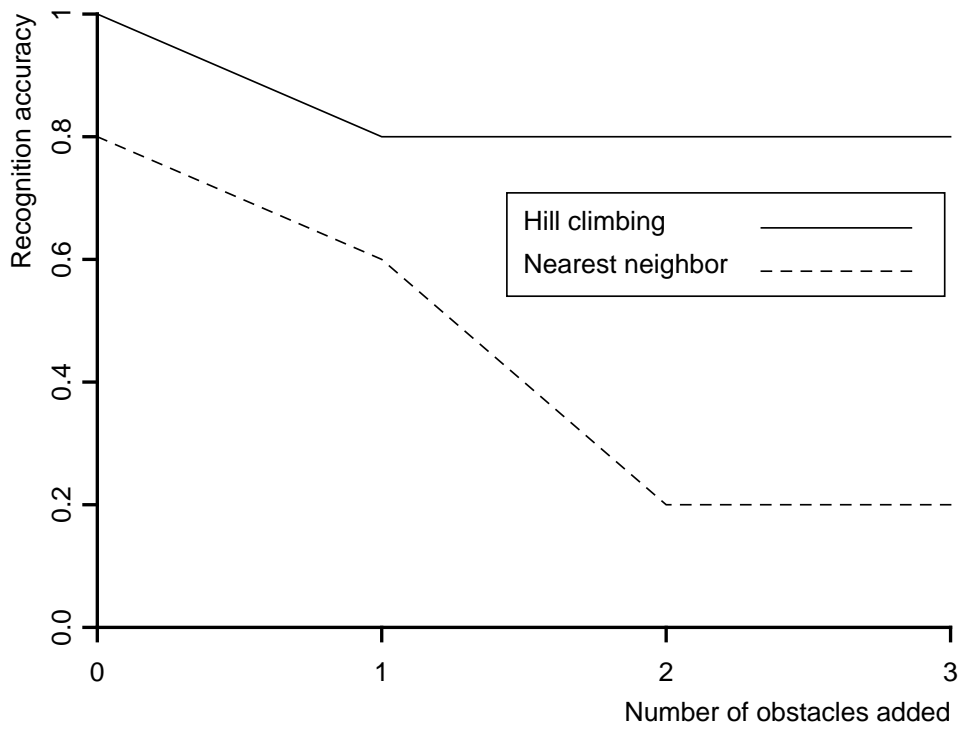


Figure 3: Recognition accuracy of matching algorithms that use hill climbing and nearest neighbor as a function of the number of new obstacles added to the environment.

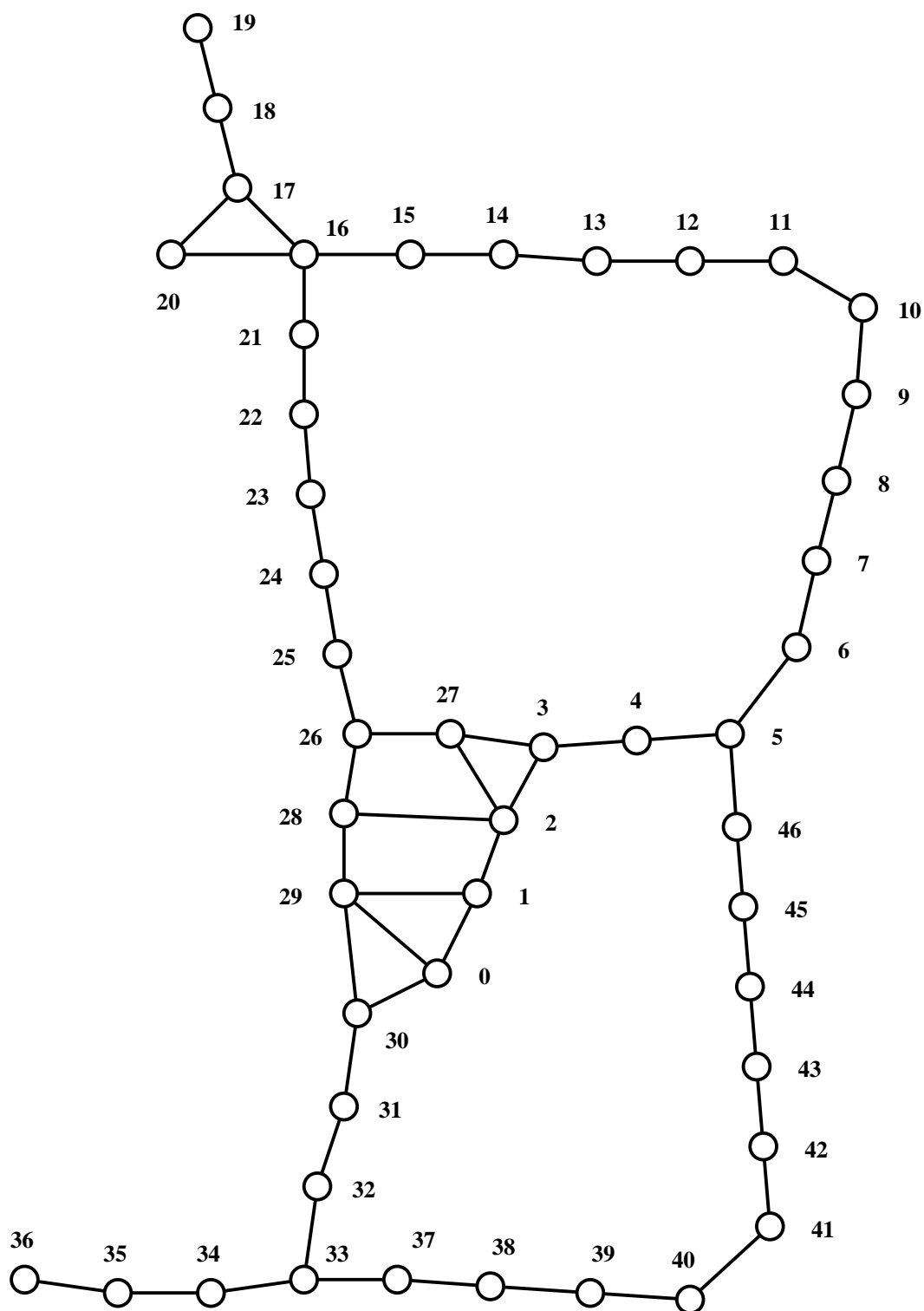


Figure 4: Topological/metric map learned for experiment 3.

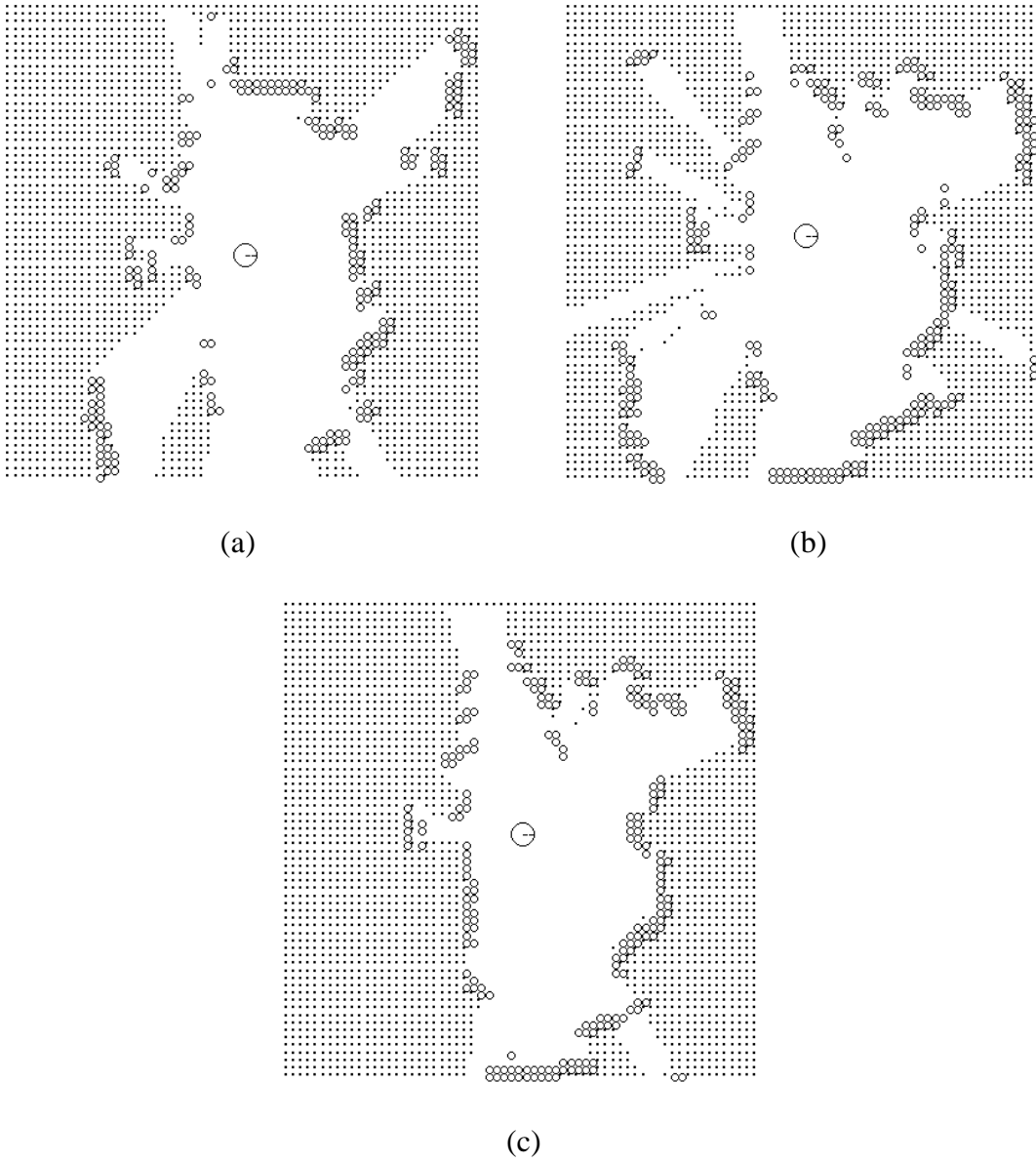


Figure 5: Learned grid for place 26 (a) and recognition grid for place 26 before (b) and after (c) specular reflector (whiteboard) was removed.

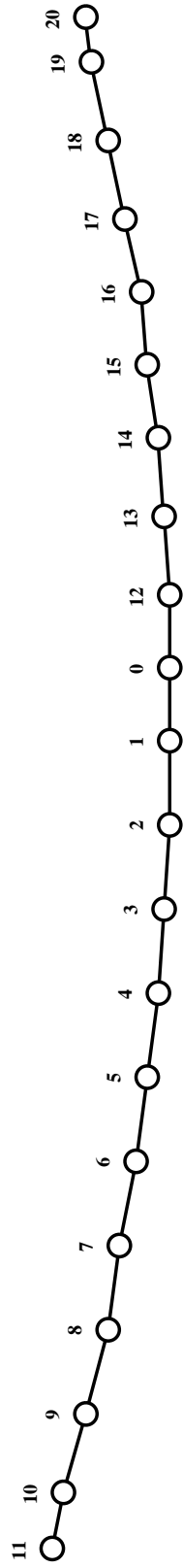


Figure 6: Topological/metric map learned for experiment 4.



(a)



(b)

Figure 7: Learned grid (a) and recognition grid (b) for place 9.

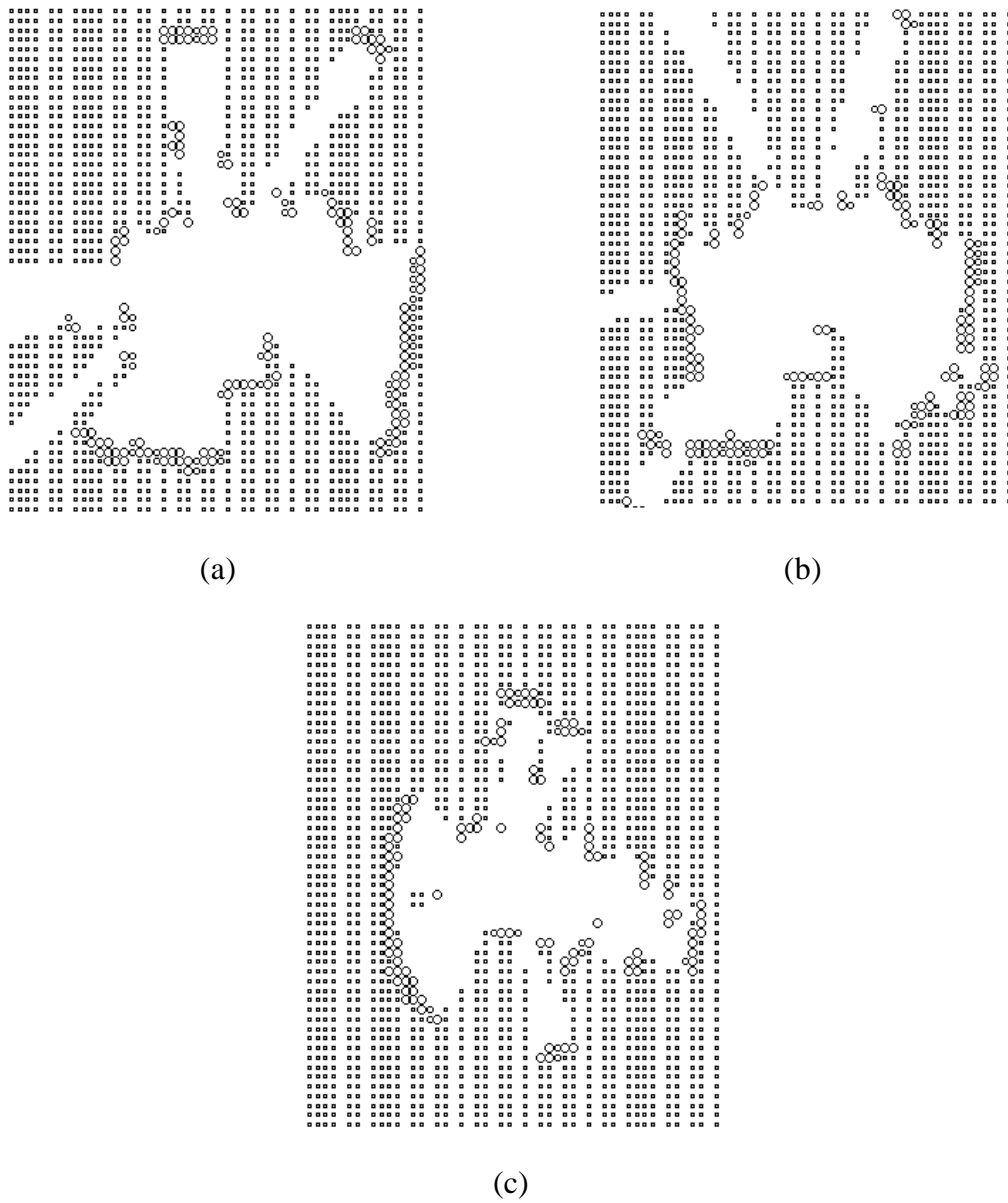


Figure 8: Learned grid for place 14 (a) which was matched correctly with the recognition grid for place 14 (b) by hill climbing, and learned grid for place 10 (c) which was incorrectly returned as the match for the recognition grid for place 14 by nearest neighbor.

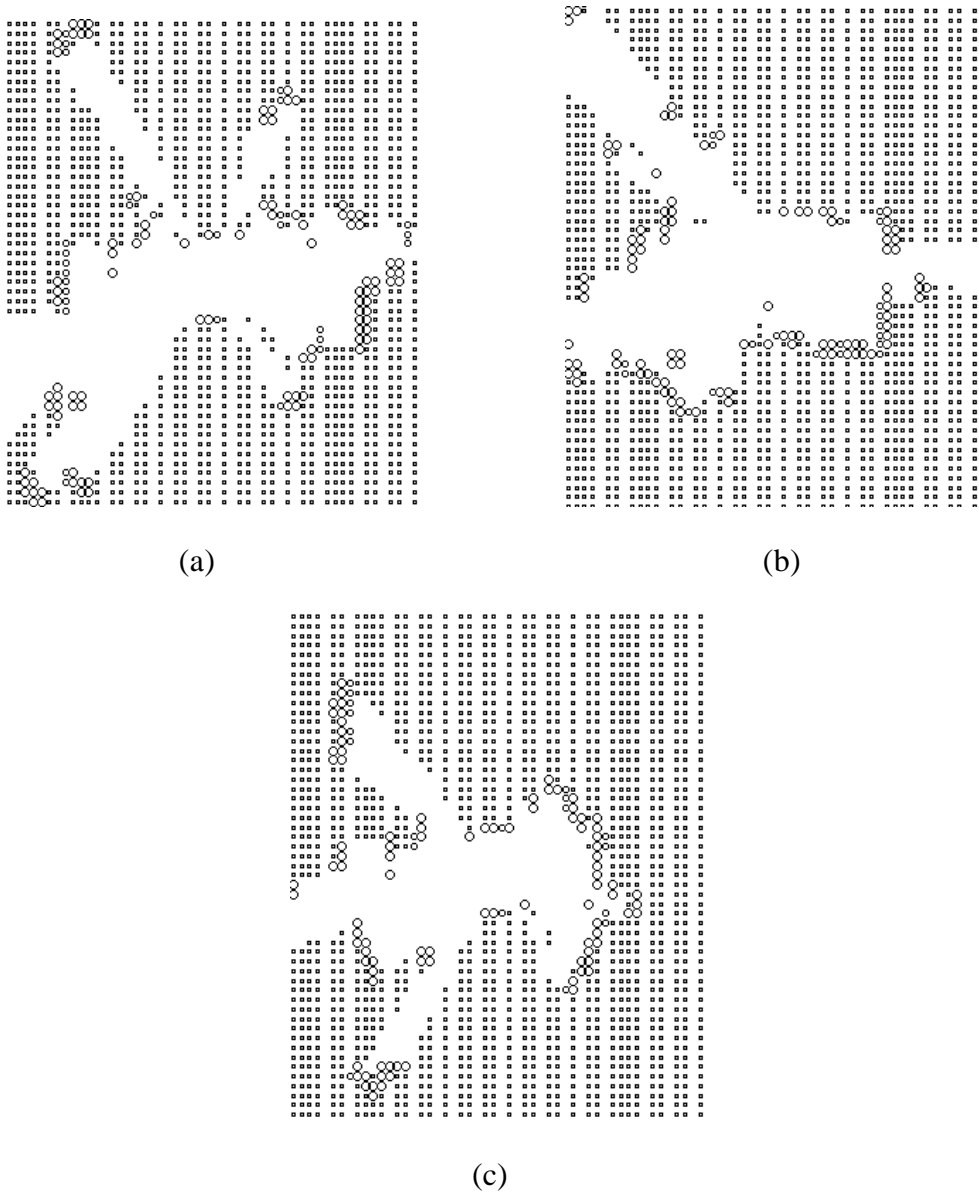


Figure 9: Learned grid for place 16 (a) and recognition grid for place 16 (b) which was confused with learned grid for place 19 (c) by both nearest neighbor and hill climbing.

Figure Captions

Figure 1: Learned grid for place 0 (a) and recognition grids for place 0 with one (b), two (c), and three (d) new obstacles. The locations of the new obstacles are circled.

Figure 2: Learned grid for place 4 (a) and recognition grid for place 4 with no new obstacles (b), one new obstacle (c), and three new obstacles (d). The locations of the new obstacles are circled.

Figure 3: Recognition accuracy of matching algorithms that use hill climbing and nearest neighbor as a function of the number of new obstacles added to the environment.

Figure 4: Topological/metric map learned for experiment 3.

Figure 5: Learned grid for place 26 (a) and recognition grid for place 26 before (b) and after (c) specular reflector (whiteboard) was removed.

Figure 6: Topological/metric map learned for experiment 4.

Figure 7: Learned grid (a) and recognition grid (b) for place 9.

Figure 8: Learned grid for place 14 (a) which was matched correctly with the recognition grid for place 14 (b) by hill climbing, and learned grid for place 10 (c) which was incorrectly returned as the match for the recognition grid for place 14 by nearest neighbor.

Figure 9: Learned grid for place 16 (a) and recognition grid for place 16 (b) which was confused with learned grid for place 19 (c) by both nearest neighbor and hill climbing.

