Semantic Mapping Using Mobile Robots

Denis F. Wolf and Gaurav S. Sukhatme, *Senior Member, IEEE*

***Abstract*—Robotic mapping is the process of automatically constructing an environment representation using mobile robots. We address the problem of semantic mapping, which consists of using mobile robots to create maps that represent not only metric occupancy but also other properties of the environment. Specifically, we develop techniques to build maps that represent activity and navigability of the environment. Our approach to semantic mapping is to combine machine learning techniques with standard mapping algorithms. Supervised learning methods are used to automatically associate properties of space to the desired classification patterns. We present two methods, the first based on hidden Markov models and the second on support vector machines. Both approaches have been tested and experimentally validated in two problem domains: terrain mapping and activity-based mapping.**

***Index Terms*—Activity monitoring, robot mapping, semantic mapping, supervised learning, terrain mapping.**

I. Introduction

CREATING an internal representation (map) of the physical environment is one of the most basic and important capabilities in mobile robotics. Most tasks to be performed by mobile robots require some type of internal knowledge of the environment. Given the importance of mapmaking, roboticists and scientists have been actively working on this topic for about two decades, and several mapping techniques have been proposed in the literature over this time (see [35] for a review).

In general, the main focus of the research on robot mapping has been on representing the geometry of the environment with high accuracy. Although robot-built maps have been successfully used for tasks like path planning, navigation, and localization, they are very limited in describing the environment other than distinguishing between occupied and empty areas. Virtually any property of space can be represented in a map, but the large majority of maps built by mobile robots consist of only metric representations of the occupancy of the space. During the mapping task, most robot mapping techniques neglect a considerable amount of information that describes other aspects of the environment like its navigability, or the nature of the activity that typically occurs there. The semantic mapping problem consists of using mobile robots to create maps that represent not only metric occupancy but also other properties of the environment.

Manuscript received July 23, 2007; revised October 10, 2007. This paper was recommended for publication by Associate Editor F. Park and Editor D. Sun upon evaluation of the reviewers' comments. This work was supported in part by the National Science Foundation under Grant IIS-0133947, CNS-0509539, CNS-0331481, and CCR-0120778, and in part by the FAPESP, Brazil.

D. F. Wolf is with the Department of Computer Systems, Institute of Mathematics and Computer Science, University of Sao Paulo, 05508-900 Sao Paulo, Brazil (e-mail: denis@icmc.usp.br).

G. S. Sukhatme is with the Robotic Embedded Systems Laboratory, Department of Computer Science, University of Southern California, CA 90089-0781 USA (e-mail: gaurav@usc.edu).

Digital Object Identifier 10.1109/TRO.2008.917001

Semantics consists of assigning meaning to data. Semantic data representation has become a popular research topic that has been explored by scientists from different fields [14], [18], [21], [31], [43]. One of the main purposes of giving meaning to information is to make data interpretation easier. The use of semantics in robot-built maps is a method of facilitating the understanding of the data that is used to represent the environment, and simplifying the sharing of environmental information by robots, people, and other machines. It also lets different properties of space to be represented allowing for richer environmental models. Our approach is based on the use of supervised learning methods to automatically associate properties of space to the desired classification patterns.

In this paper, we present semantic mapping approaches for two problems: semantic terrain mapping and semantic activity mapping. The terrain mapping problem consists of creating 3-D representations and classifying the terrain according to its navigability. Applications for this approach range from local avoidance of nonnavigable areas to path planning. The activity-based mapping problem consists of creating 2-D maps that classify the environment according to the occupancy of the space by dynamic entities over time. Examples of applications for these maps are path planning and urban traffic modeling.

We propose two semantic mapping approaches, one based on hidden Markov models (HMMs) and another based on support vector machines (SVMs). Both approaches have been applied to both problem domains considered here. A map segmentation algorithm based on Markov random fields has been used to improve the semantic classification. The two semantic mapping approaches have been evaluated based on field experiments.

II. Related Work

Semantic mapping is a young research topic in mobile robotics. Few papers have been published on this subject and most of them are very recent. The approach presented in [7] suggests an idea to infer high-level information from regular robotic maps. The authors claim that the semantic information extracted from the maps can be used for planning and decision making.

In [27], the idea of semantic mapping is presented as part of a human–robot interface. Uniquely identifiable labels are assigned to specific objects of the environment and instructions making use of these labels are given to the robot. In [12], [32], and [37] specific places of indoor environments are labeled based on the presence of key objects in them. Computer vision techniques are used to extract the necessary information from images. A similar approach is presented in [20], the difference being that only range data is used.

The paper described in [23] presents a technique to identify certain objects in the environment based on range

information analysis. In [1], a similar work is presented based on the expectation-maximization (EM) algorithm. In [29], an approach to extract semantic information from indoor 3-D laser maps is presented. This technique is capable of differentiating walls, floor, ceiling, doors, and other key parts of the environment in the map. A similar technique by [33] uses line segments to build object-oriented environment representations. The mapping approach, based on Adaboost, presented in [25], is able to semantically classify different places in indoor environments like rooms, corridors, and doorways based on rage data.

III. Theoretical Background

A. Hidden Markov Model (HMM)

An (HMM) consists of a discrete time and discrete space Markovian process that contains some hidden (unknown) parameters and emits observable outputs. The challenge is to estimate the hidden parameters based on observable information. This statistical tool is largely used for pattern recognition and is particularly popular in speech recognition. For a complete tutorial, see [30].

An HMM can be defined by five elements: a set of possible states (parameters), a set of possible observation symbols, the state transition probability table, the observation symbol probability distribution, and a initial state distribution. For convenience, λ will be used to characterize an HMM model, which consists of the last three elements of the HMM.

Let us assume that O is an observation sequence and each observation symbol in the sequence is denoted by o_n where n means the position of that observation symbol in the sequence ($O = o_1, o_2, \ldots, o_t$ for a sequence of size t). Let us also assume that Q is a state sequence and each state is denoted by q_n, therefore $Q = q_1, q_2, \ldots, q_t$. There are three general problems that can be solved with HMMs. In the first one, given the observation sequence O, and the model λ, how to compute the optimal corresponding state sequence Q (i.e., the state sequence that best explains the observations). In a nutshell, we are trying to maximize the expression $P(Q \mid O, \lambda)$. In the second problem, given the observation sequence O, and the model λ, how to compute $P(O \mid \lambda)$, the probability of the observations given the model. The third problem consists of how to estimate the model λ to maximize $P(O \mid \lambda)$, the probability of the observations given the model. In this paper, we use the first HMM problem framework to formulate our semantic mapping approach. The solution for this problem can be obtained using the Viterbi algorithm [10], which is based on dynamic programming techniques.

B. Support Vector Machines (SVMs)

SVMs is a general class of supervised learning techniques based on statistical learning theory and used for classification and regression problems introduced by [36]. Basically, SVM performs classification by estimating hyperplanes in multidimensional spaces, separating data from different classes. A complete survey is presented in [5]. In order to handle nonlinear classification problems, the kernel trick proposed by [3] can be used. This idea consists of using a nonlinear function to map nonlinearly separable data to a different Euclidean space where it can be linearly separable.

Four standard kernel types have been used during our experiments: linear, polynomial, radial basis function (RBF), and sigmoid. The package SVM-Light [15] has been used for the SVM learning and classification.

IV. Semantic Terrain Mapping

Autonomous navigation is an important capability for a mobile robot. When traversing rough terrain, the robot must have the ability to avoid not only obstacles but also parts of the terrain that are considered not safe for navigation [2]. This is an important problem when one is exploring unknown terrain. Applications for terrain mapping range from path planning and local obstacle avoidance to detection of changes in the terrain and object recognition [42]. The paper described in [38] presents a terrain mapping technique used for outdoor navigation. Their approach combines vision and range sensor information to create accurate representations of the terrain. Vision-based navigation and mapping are also addressed in [19] and [16].

Planetary exploration is an interesting example of a practical application for this research topic [26]. In this context, we use a semantic mapping approach based only on information provided by range sensors and an odometer. Our approach builds 3-D maps of the terrain and classifies mapped regions into two categories: navigable and nonnavigable.

Experiments have been performed using ground robots equipped with 2-D laser range finders (mounted pitched down). As the robot moves, the range information generates a 3-D point cloud map [41], which models the terrain. In this paper, we consider flat parts of the terrain such as walkways as navigable areas. Grass and gravel are considered nonnavigable (or less desirable) areas. Depending on the application, different types of terrain may be considered navigable and nonnavigable. For example, in a planetary exploration context, areas with large rocks that may damage the robot may be considered nonnavigable. Classified maps can be very useful for path planning and safe navigation. Even though sometimes the difference in the roughness of concrete walkways and grass is very small, our approach is capable of classifying them successfully.

It is important to mention that the approach proposed here can easily be used to identify different types of navigable and nonnavigable terrains. As long as the variability of the altitude is detectable by the laser range finder, our algorithm is able to differentiate them.

A. HMM

The semantic terrain mapping problem can be stated using the HMM framework as follows. The points on our 3-D map correspond to the states we are estimating and each scan provided by the range sensor generates a single-state sequence. Each point can assume one of two possible states: navigable (NA) or nonnavigable (NN), which correspond to the output of the HMM. The real state of each point is not directly given by the range sensor. Navigable areas in our context are characterized by flat terrains. On those terrains, the points generated by the range

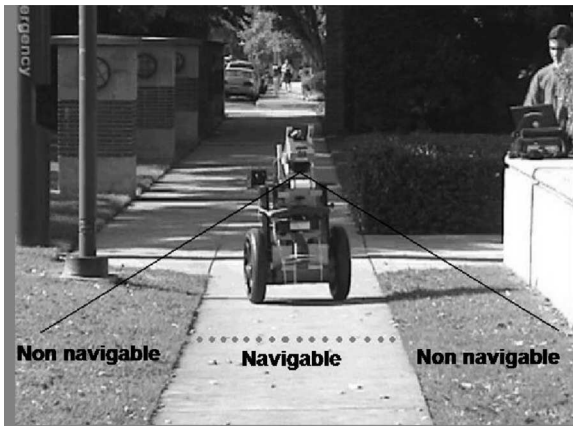


Fig. 1. Robot performing terrain mapping and identifying navigable areas.

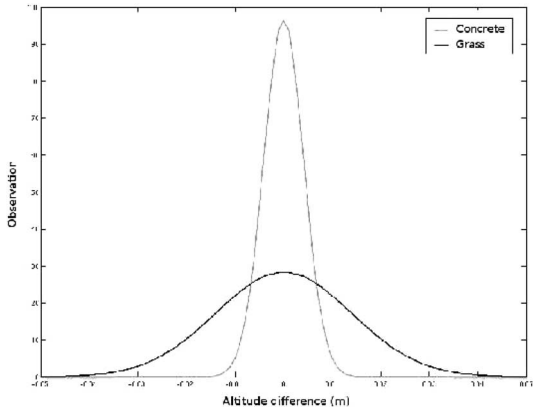


Fig. 2. Gaussian pdfs for the NA (black) and NN (grey) points of the walkway terrain. Flat terrain (walkway) points have smaller variance in altitude.

sensor are expected to be well aligned, with a minimal variance in altitude. Conversely, the nonnavigable areas are characterized by rougher terrains. The 3-D points that represent those areas are expected to be not well aligned, with some variance in altitude [40]. Fig. 1 shows a robot building a map and identifying navigable and nonnavigable parts of terrain at a location on the University of Southern California (USC) campus.

The information provided by the range sensor cannot be directly used as observation in our algorithm, since it is only a measurement of the distance between the sensor and the nearest object in a specific direction. Therefore, the data provided by the range sensor are represented as a sequence of points in 3-D Cartesian space. Given this sequence of points, the observation (HMM inputs) for a specific point is based on the difference in the altitude of a point compared to the altitude of its neighbors (in the same scan line). It is important to notice that instead of having a discrete set of observations for individual states, our approach uses continuous values.

It is possible to use the HMM to learn the model parameters λ (the third HMM problem), but in our specific case, learning can be done fairly easily through the use of examples. Given a set of range scans, it is possible to manually label the true state for each point on each scan. Given the labeled data, the calculation of the state transition probability distribution NA is straightforward. As we have only two possible states (NA and NN), it consists of counting the number of times that a point labeled NA is followed by a point labeled NA or NN. The same rule applies for points labeled NN. These numbers must be normalized so that the probability distribution sums up to 1. Calculating the initial state distribution π is also easy when labeled data is available (NA and NN are counted and the distribution is normalized).

Based on labeled data, it is also possible to calculate the variance in the observation for the points classified as NA and NN and use this information as an observation symbol probability distribution. In this case, it is necessary to calculate a mean for the altitude of the points that belong to a specific state. After that, it is necessary to calculate the amount of variation in the altitude of those points in comparison to the mean. Fig. 2 shows the Gaussian pdfs for the points NA and NN. We observe that points classified as NA have smaller variance in the relative altitude (observation) compared to the points classified as NN.

B. SVM

The terrain mapping problem can also be stated using the SVM framework as follows. Two properties of the space have been used as inputs of the SVM algorithm: 1) the altitude difference between a specific point and the robot and 2) the maximum difference in the altitude of a specific point and its neighboring points (same used in the HMM terrain mapping). As an output, SVM classifies each point in the 3-D terrain map as NA or NN.

As SVM is a supervised learning algorithm, previously classified data has to be provided in the learning phase. The same manually labeled points used for the HMM model learning have been used for the SVM learning. Depending on the kernel choice, some parameters have to be manually set during the learning phase. The cross-validation method has been used to tune these parameters.

V. Activity-Based Semantic Mapping

For the most part, mobile robotics mapping research has been concentrated on static environments. Few mapping algorithms are designed to build maps in the presence of dynamic obstacles, and in most cases, the dynamic entities are detected and filtered out. Some mapping approaches, like the one presented in [39], go further, detecting and representing moving obstacles in the map. There are also robot-based algorithms that focus on learning patterns of the activity in the environment [22]. The paper described in [34] presents a mapping approach that identifies areas occupied by dynamic entities to improve localization, focusing on dynamic indoor environments. Computer vision researchers have addressed the topic of activity modeling for many years [4], [28]. The use of video cameras provides richer information compared to laser range sensors but also demands considerably more computation to process the data. Usually, cameras may be placed in high vantage positions to better cover the space to be analyzed. Consequently, they are not suitable to be attached to ground robots for this purpose. Another fact to take into account when monitoring urban spaces is that the lasers

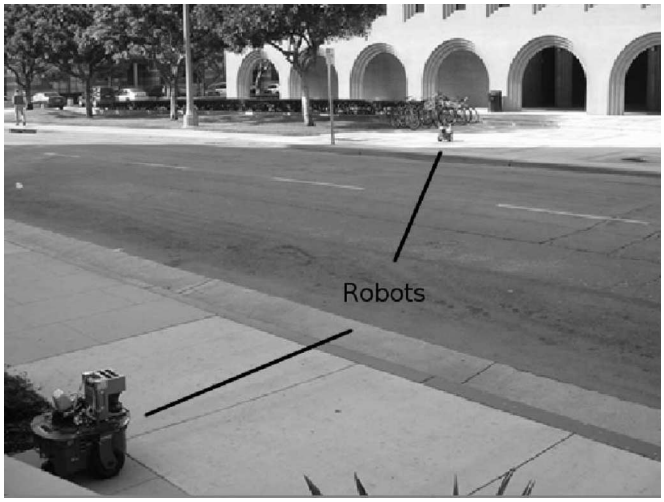


Fig. 3. Robots mapping the environment according to the activity of the dynamic entities (e.g., pedestrians walking by or cars driving by).

preserve the anonymity of the moving entities, which may be an issue for video cameras depending on the situation in which they are used.

We approach the problem of semantic mapping of the environment based on the usage of the space by dynamic entities. In urban environments, there are many different types of dynamic entities moving with different properties. Some entities are bigger than others, some are faster than others. We are particularly interested in obtaining semantic information from the environment based on how moving entities use space over time.

Our experimental scenario consists of an urban environment (a street on the USC campus) composed of a street and sidewalks, where people, cars, and bicycles are found regularly in traffic. We use an occupancy grid mapping algorithm [9] to represent the environment. Based on activity, we semantically classify the space (grid cells) into either street (S) or sidewalk (W). The applications for the activity-based semantic maps range from path planning to traffic modeling. Fig. 3 shows two robots collecting activity information in an urban environment.

A. HMM

Our semantic classification for the environment can be stated in the HMM framework as follows. We partition the environment into a 2-D grid of cells. At the end of the semantic mapping process, each cell is classified in one of the two categories (S or W).

As all cells are of the same size in the grid, it is possible to organize and address them by columns and rows. Each row of grid cells corresponds to a sequence of hidden states that we are estimating using the HMM framework.

The observations (HMM inputs) are obtained from the properties of the space sensed by the mobile robots over a fixed amount of discrete time. For example, one of the properties of the space is activity. Every time that a robot detects activity in a portion of the environment, the activity counter on the corresponding grid cell is increased by one. At the end of the data collection period, the amount of activity that occurred in the space represented by each cell is used as observation for the HMM classification.

Due to the fast motion of the dynamic entities and the limited observability of robots' sensors, robots do not move during the data collection to avoid missing information. The mapped area corresponds to the area covered by the sensor of the robot. In order to decrease the effect of occlusion and increase the mapped area, two robots have been used during the experiments. The location of the robots is known *a priori* and it is important to mention that only range information is provided by the sensors.

Four properties have been observed in the environment: activity, occupancy, maximum size of the dynamic entities, and average size of the dynamic entities. *Activity* is detected every time a certain place in the environment is occupied (by a dynamic entity) and becomes free or vice-versa. *Occupancy* occurs when a certain location of the environment is occupied by a dynamic entity. For example, if a car stops in a determined place, the corresponding grid cells will show high occupancy and low activity. The third property of the environment is the *maximum size of the moving entities* that occupied the space during the data acquisition. The fourth property of the environment is the *average size of the moving entities* that occupied the space during the data acquisition.

We define as dynamic any entity of the space that has been seen to move at least once. This information is extracted from the data provided by the range sensors using the algorithm presented in [39]. The size of the dynamic entities is estimated to be based on grouping adjacent occupied cells. As in the semantic terrain mapping, the HMM model λ is learned based on manually labeled examples.

B. SVM

The activity-based semantic mapping problem can be stated using the SVM framework as follows. The same four properties used in the HMM classification have also been used as inputs of the SVM algorithm. As output, the SVM algorithm classifies each grid cell of the map into either street (S) or sidewalk (W). An important difference between the HMM and SVM approaches for this problem is that the grid cells with no activity during the experiments have not been considered for classification with the SVM algorithm (they have no inputs other than trivial ones).

VI. Map Segmentation

It may happen that parts of the map are not correctly classified due to sensor noise or other reasons. When those errors occur in small parts of the map (considered noise), segmentation techniques can be used to fix them and improve classification.

Segmentation techniques have been used for many years by the computer vision community [11]. Among several segmentation methods, Markov random fields (MRF) have been extensively used to that end. For a complete overview of the MRF theory, see [17].

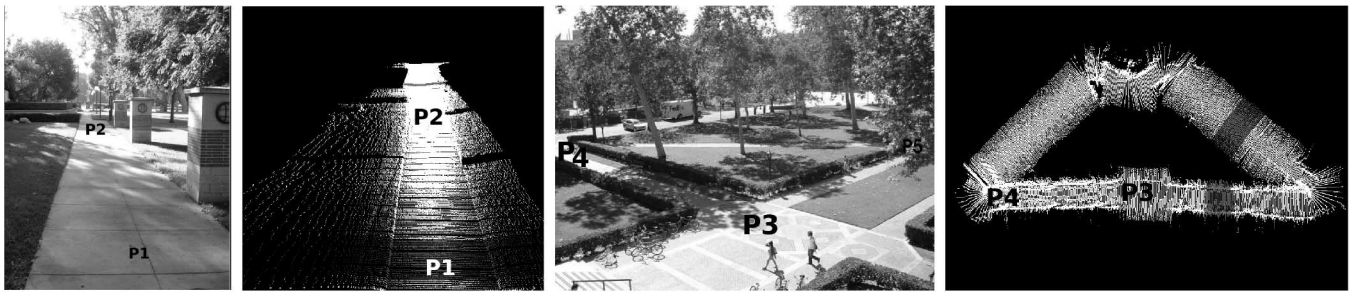


Fig. 4. Real environments [(a) and (c)] and corresponding 3-D models [(b) and (d)] with the ground truth areas in grey. Points P1 and P2 correspond to the same location in (a) and (b). Points P3 and P4 correspond to the same location in (c) and (d). (a) Walkway. (b) 3-D map of the walkway. (c) Doheny Library garden. (d) 3-D map of the garden.

In the terrain mapping context, it is unlikely that in an area with a large majority of cells NA, there are few cells NN. It is also unlikely that among several scans that contain both points labeled NA and NN, there is a scan that only contains points labeled NA. The MRF technique basically makes these points agree with their neighbors. As a result, NA and NN regions are well defined and clustered.

In order to use the MRF as a segmentation tool, we approximate our 3-D map to a 2-D grid. Each point is projected on the grid based on its x and y coordinates. Each cell in the grid is labeled as NA or NN according to the classification of the points projected on that cell. It is important to notice that not all the cells in the grid have a label, because the distribution of points on the plane is not uniform.

The basic idea of the MRF is that the probability distribution for each cell (and corresponding points) in the grid is specified conditionally on the probability distribution of its neighboring cells.

Let C_i be a random variable taking the values NA or NN, and denote by $n_i^{(k)}$ $(k = 1, 2, \ldots, n)$ the number of k neighbors of C_i that are labeled as NA. A simplified MRF model may be specified as

$$\frac{P(C_i = NA \mid \text{grid})}{P(C_i = NN \mid \text{grid})} = \exp\left(\alpha + \sum_{k=1}^{n} \left(\beta_k n_i^{(k)}\right)\right)$$

where α and β are, respectively, the prior about the number of NA cells and the importance ratio based on the distance to the cell C_i. As β is increased, the chance of each grid cell value agreeing with the value of its neighbors increases.

The same segmentation idea can be trivially applied to the activity-based mapping problem.

VII. Experimental Results

A. Semantic Terrain Mapping

In order to validate our terrain mapping approach, experimental tests have been performed. Our experiments have been done using both an ActiveMedia Pioneer and a Segway RMP robot. Both robots were equipped with SICK LMS200 laser range finders and a Microstrain IMU. A player [13] was used to interface with robots' sensors and actuators. On the Pioneer platform, the laser was mounted at a height of 42 cm and a pitch angle of 40° (pitched down). On the RMP, the laser was mounted at a height of 93 cm height and a pitch angle of 35°. This configuration allowed the robot to map a terrain approximately 1.3 m ahead of it.

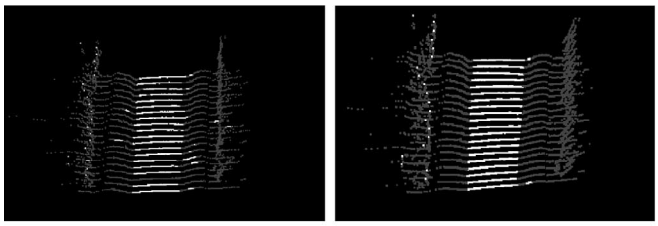
TABLE I
Accuracy of the HMM Terrain Semantic Classification

	Hist 10	Hist 50	HMM	HMM + MRF
Walkway section A	80.0%	84.2%	97.2%	97.4%
Walkway section B	65.6%	77.8%	96.2%	98.5%
Garden section A	64.7%	72.5%	93.1%	97.6%
Garden section B	66.5%	67.5%	97.8%	98.9%

Our experiments were performed in two different scenarios: a walkway and a garden, both on the USC campus. The walkway scenario was reasonably uniform, with a concrete passage in the center and short grass on both sides. The garden is more diverse than the walkway. It contains grass, bushes, garden seats, and a water fountain. The mapped walkway is approximately 50 m long and a complete loop around the garden is 200 m long.

In order to verify the accuracy of the HMM classification, we manually labeled 13 640 points in the walkway dataset and 12 987 points in the garden dataset. These were used for ground truth. For each dataset, the labeled points were extracted from two different sections (A and B) of the map to avoid biasing. Approximately 30% of the labeled points were used to learn the models and the other 70% to test them. Fig. 4 shows pictures of the real environments and the 3-D maps before the HMM semantic classification.

During the experiments, most of the maps were built when the robots were manually driven with a joystick. But some autonomous navigation experiments using the RMP were also performed. As our mapping and classification algorithms can be executed in real time, the robot used the information about the areas it should avoid and kept itself in the navigable areas while mapping the environment. For a better evaluation of the classification results obtained with the HMM method, histogram classification techniques have also been applied to the same data, serving as a reference for classification accuracy results [8]. Table I shows the semantic classification results obtained with histogram techniques, HMM, and HMM + MRF methods. The histogram classification technique has been tested with different number of classes and empirical tests showing that the most representative results were obtained with 10 and 50 bins.



(a) (b)

Fig. 5 Semantic classification of Section A of the garden terrain (white corresponds to NA and grey to NN). Some misclassified points are corrected based on the MRF segmentation technique. (a) HMM semantic classification. (b) HMM semantic classification + MRF.

TABLE II
ACCURACY OF THE SVM SEMANTIC CLASSIFICATION FOR THE WALKWAY ENVIRONMENT

Kernel	**Walkway section A**		**Walkway section B**	
	SVM	SVM+MRF	SVM	SVM+MRF
Linear	90.91%	90.51%	92.53%	93.18%
Polynomial	97.00%	97.06%	96.82%	97.35%
RBF	97.59%	97.76%	97.82%	98.15%
Sigmoid	90.18%	89.47%	91.85%	92.32%

Fig. 6(a) and (b) and Fig. 7(a) and (b) show the 3-D maps with the semantic classification using HMM in both scenarios. Parts of the terrain shown in white represent the navigable areas, while those in grey correspond to nonnavigable terrain. We notice that although the grass is level and uniform in several areas of the terrain, our method could successfully differentiate it from gravel.

Figs. 6(a) and 7(a) show that some portions of the terrain have been misclassified. This happens mostly due to the presence of small obstacles like leaves in the flat parts of the terrain or aligned regions in the rough terrain. Some of these errors were corrected afterward with MRF segmentation [Figs. 6(b) and 7(b)]. The improvement in the classification can also be noticed in the results presented in Table 1. Fig. 5 shows a closer view of Section A of the garden terrain where the correcting effects of the MRF segmentation can be seen clearly.

Table II shows the terrain classification results for Sections A and B of the walkway dataset using the SVM classification approach. As can be seen, the best semantic classification results were obtained with the RBF kernel (see [24] for more details on the RBF). In almost every case, the MRF segmentation method has contributed with at least a small improvement in the results. Fig. 6(c) and (e) shows the 3-D semantic terrain map of the walkway environment using the linear and RBF kernels. It is possible to notice that the MRF segmentation algorithm corrected a considerable part of the misclassified points in the RBF case [Fig. 6(d) and (f)]. The same did not happen to the linear kernel due to a larger number of misclassified points. Table III shows the results of the semantic classification for the garden dataset. The RBF kernel again presents the most accurate classification. Results obtained in Section A are noticeably better than those obtained in Section B. Compared to the walkway dataset, Sections A and B in the garden dataset are very different. The terrain data of Section A contains tall bushes, which are easier to classify than the short grass mapped in Section B. Fig. 7 shows

TABLE III
ACCURACY OF THE SVM SEMANTIC CLASSIFICATION FOR THE GARDEN ENVIRONMENT

Kernel	**Garden section A**		**Garden section B**	
	SVM	SVM+MRF	SVM	SVM+MRF
Linear	87.75%	87.84%	73.69%	73.78%
Polynomial	92.47%	93.87%	84.14%	85.58%
RBF	94.40%	95.34%	85.48%	91.21%
Sigmoid	47.52%	47.63%	47.20%	43.41%

TABLE IV
ACCURACY OF THE HMM ACTIVITY-BASED SEMANTIC CLASSIFICATION

Property	Hist 10	Hist 50	HMM	HMM + MRF
Activity	53.09%	51.77%	65.00%	69.87%
Occupancy	52.64%	47.32%	69.78%	76.73%
Average size	55.97%	54.11%	78.20%	83.01%
Maximum size	46.93%	31.70%	78.26%	82.72%

the semantic classification results obtained with linear and RBF kernels, with and without RMF segmentation.

The poor classification obtained by the linear kernel compared to the RBF and polynomial kernels can be explained by the fact that this data is not linearly separable. Fig. 8 shows the ground truth and the classification of the walkway data in the property space, where grey represents the NA points and black the NN points. As can be noticed in Fig. 8(a) and (b), there is some overlap between the two classes, which prevents completely correct classification results.

It is also important to mention that our approaches do not (presently) detect a step located in front (or in the back) of the robot. Our algorithm could be easily adapted to this situation, but it would require an extra laser pointing sideways

B. Activity-Based Semantic Mapping

The activity-based semantic mapping approach presented in the previous section has been tested with experimental data collected in the USC campus. Two Pioneer robots equipped with SICK LMS200 laser range finders have been positioned on opposite sides of a street to monitor the activity in the region. The area considered for mapping is approximately 16 m $\times$ 18 m with grid cells of 20 cm. The ground truth was obtained by measuring the width of the street and the sidewalks. Each data collection period lasted approximately 15 min with a sampling frequency of 10 Hz. One out of 80 rows of grid cells (from one sidewalk to the other) was manually labeled and used in the learning step of the HMM algorithm. It corresponds to 1.25% of the total grid cells. Fig. 9 shows a picture of the environment used for the experiments and a screenshot of the map generated by the robots at a particular time.

Based on the raw sensor data, the robots are capable of extracting the properties of the environment and semantically classifying each grid cell as either S or W. Each of the four properties has been individually tested. The accuracy results of the HMM classification can be seen in Table IV, which also includes classification results obtained with histogram techniques.

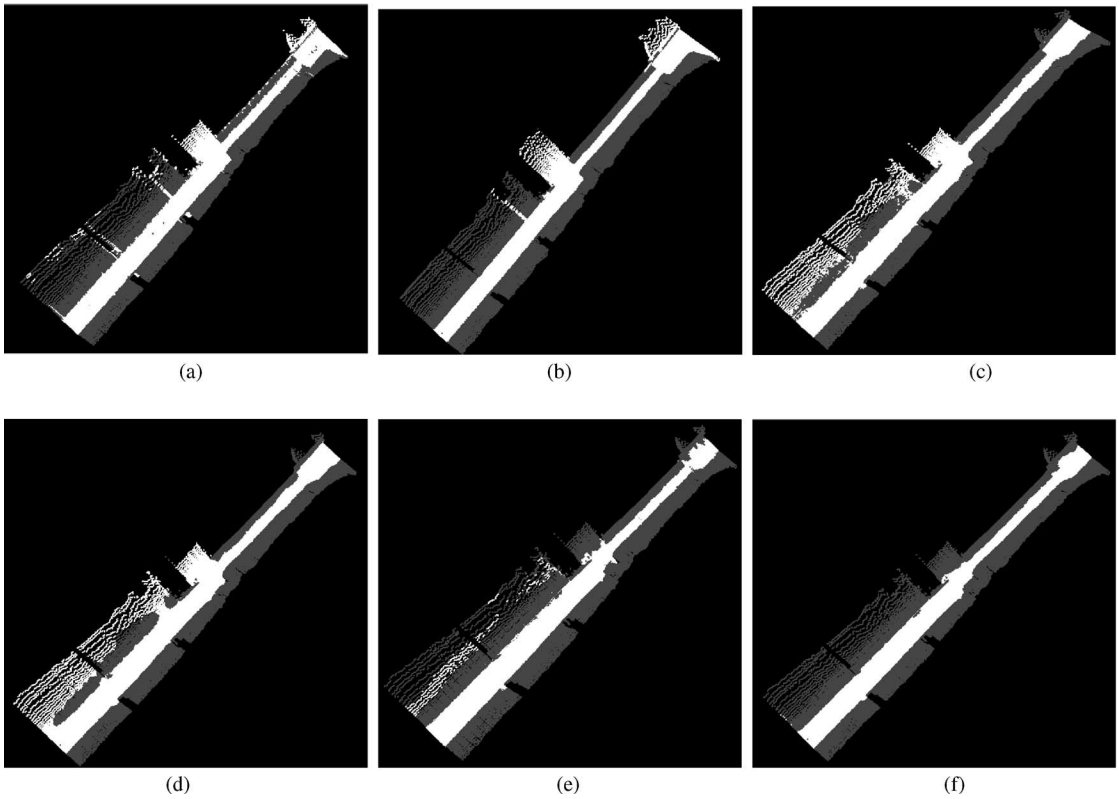


Fig. 6. Classification accuracy results for the walkway terrain. The walkway in the center (in white) is correctly identified as navigable. (a) HMM. (b) HMM + MRF. (c) SVM linear kernel (d). SVM linear kernel + MRF (e) SVM RBF kernel. (f) SVM RBF kernel + MRF.

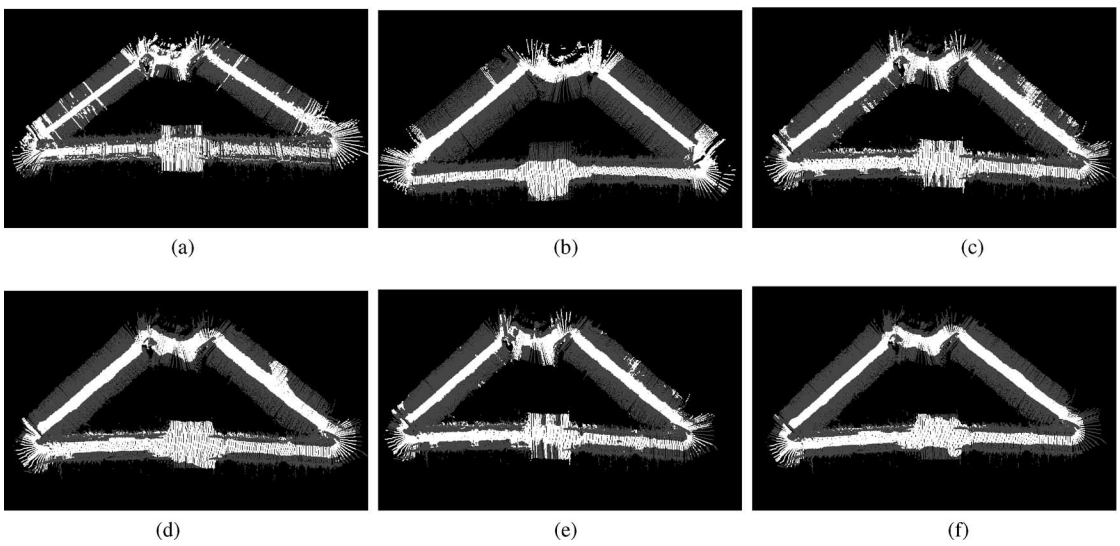


Fig. 7. Semantic classification results for the garden terrain. The walkway in the center (in white) is correctly identified as navigable. (a) HMM. (b) HMM + MRF. (c) SVM linear kernel. (d) SVM linear kernel + MRF. (e) SVM RBF kernel. (f) SVM RBF kernel + MRF.

Fig. 10 shows the semantic classification results with and without the use of the MRF segmentation algorithm. Parts of the map colored in light grey correspond to the W areas, while dark grey colored areas correspond to the S areas. The two black lines are the ground truth and the space between them corresponds to the street, while the side spaces correspond to the sidewalks.

As can be noticed in Fig. 10, the properties activity and occupancy cannot correctly differentiate the street from the sidewalks. The two wide dark grey lines in the center of the black lines in Fig. 10(a) and (c) do correspond to the parts of the street used by traffic, but when the semantic classification algorithm tries to generalize the learned information, it also classifies the most active parts of the sidewalks as S. After the MRF

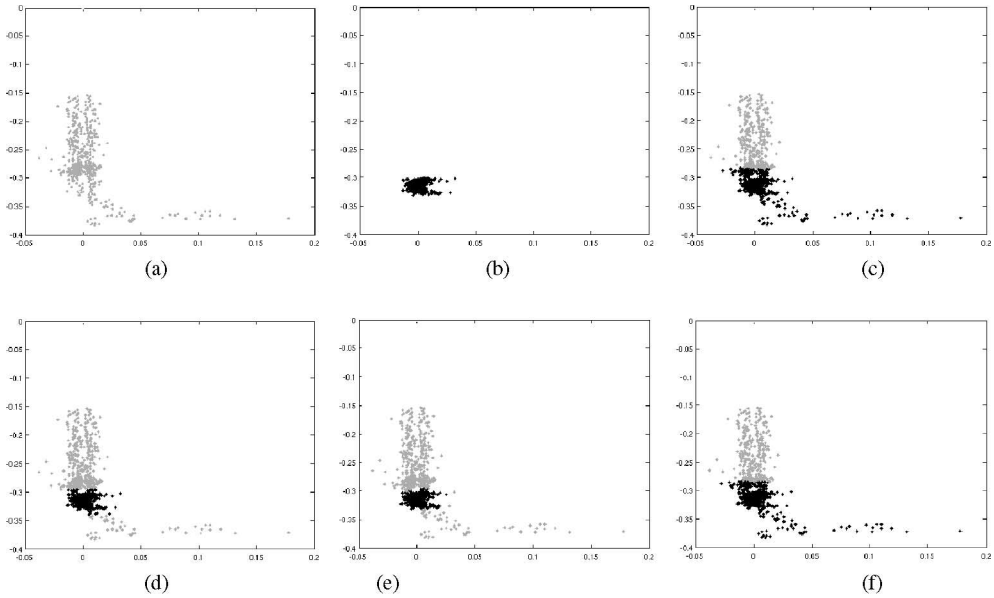


Fig. 8. Data points in the property space (NA in grey and NN in black), and the classification performed by linear, polynomial, sigmoid, and RBF kernels. (a) Ground truth—NN. (b) Ground truth—NA. (c) Linear kernel classification. (d) Polynomial kernel classification. (e) RBF kernel classification. (f) Sigmoid Kernel classification.

Fig. 9. Environment used for the activity-based semantic mapping and the space representation created by the mobile robots. The black square corresponds to the mapped area, the black parallel lines inside the square divide the street from the sidewalks, and the small black points are the locations of the robots. (a) Real environment. (b) Screenshot of the 2-D map. (c) Ground truth map.

segmentation, the right sidewalk is entirely classified as street, which is obviously wrong. In fact, based on the data collected during our experiments, it is not possible to determine whether a grid cell belongs to the street or sidewalks just by observing the activity and occupancy, as they are very similar in both the S and W areas.

The semantic classification based on the properties average size and maximum size shows better results, as can be seen in Fig. 10. We notice in Fig. 10(e) and (g) that most of the area colored in dark grey is between the two black lines, which matches the ground truth information. Some parts of the space between the two black lines are misclassified as W. This happens because these areas (which are close to the sidewalks and in the center of the street) are not used by cars. On the right side of Fig. 10(g), it is possible to see a dark grey area in the place that corresponds to the sidewalk. The explanation for this misclassification is that during the experiments, at a specific moment, a crowd of people stopped in front of the robot. As most of the space around the robot was occupied by moving entities, the range sensors detected a large-sized obstacle at that location. When the average size of the moving entities that occupied that space is used for the semantic classification, the effect of the crowd of people is attenuated, as can be seen in Fig. 10(e). Fig. 10(f) and (h) shows the classification results after the MRF segmentation.

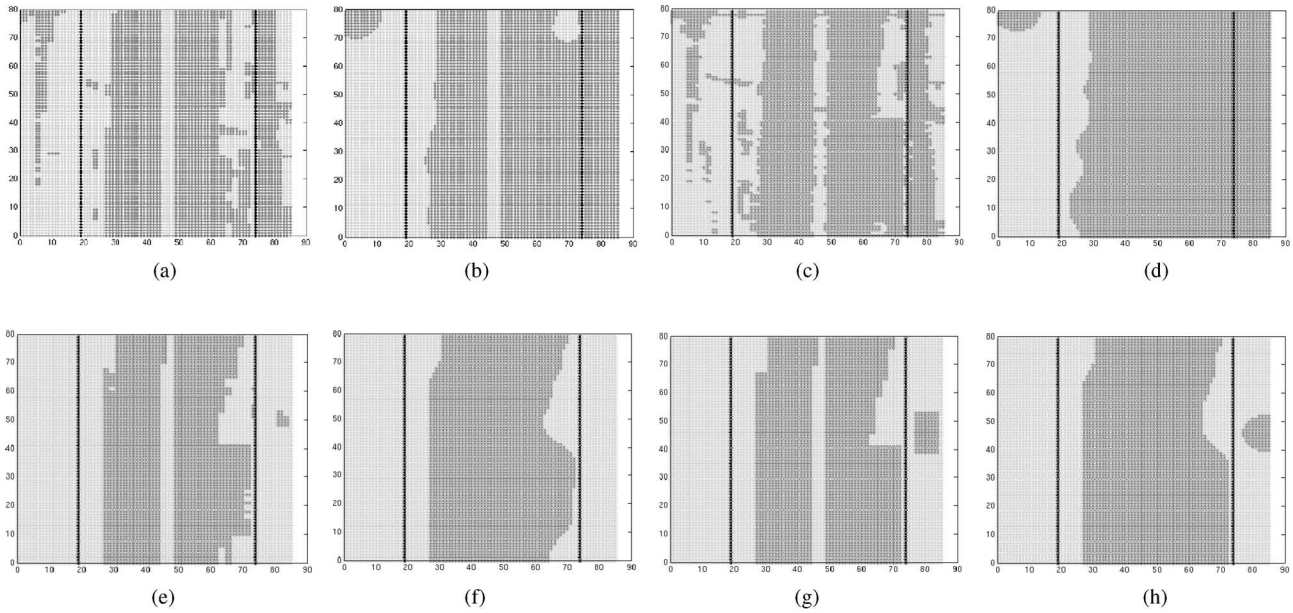


Fig. 10. Activity-based semantic classification based on different properties of the space. Dark grey corresponds to the areas identified as street (S). (a) Classification based on activity. (b) Classification based on activity + MRF. (c) Classification based on occupancy. (d) Classification based on occupancy + MRF. (e) Classification based on average size. (f) Classification based on average size + MRF. (g) Classification based on maximum size.(h) Classification based on maximum size + MRF.

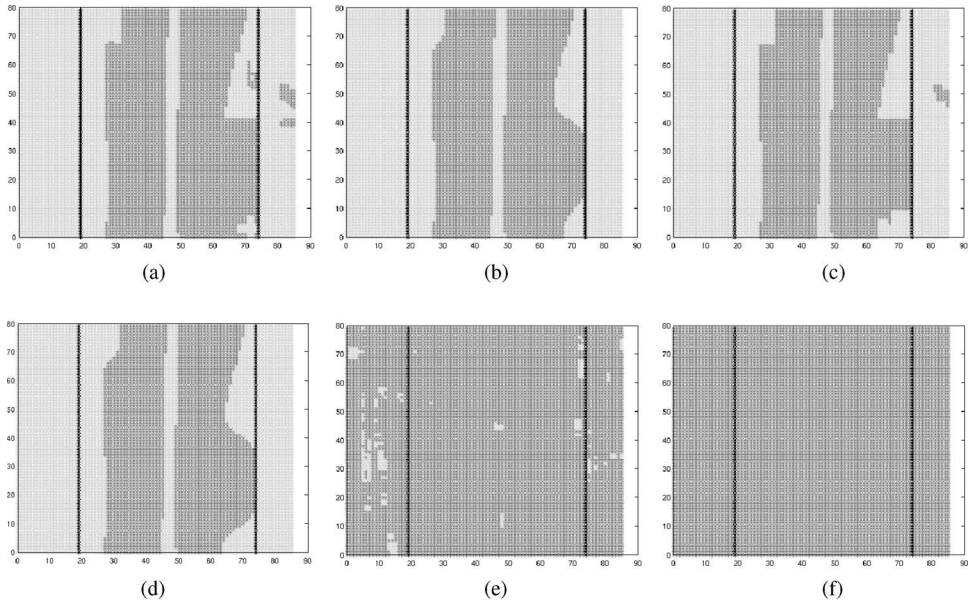


Fig. 11. Results of the SVM semantic classification (W in grey and S in black). The linear kernel correctly distinguishes the street from the sidewalks, while the RBF kernel overfits the learning datasets and presents poor classification results. (a) SVM linear kernel. (b) SVM linear kernel + MRF (c) SVM polynomial kernel. (d) SVM polynomial kernel + MRF. (e) SVM RBF kernel. (f) SVM RBF kernel + MRF.

An important difference between the HMM and SVM approaches for this problem is that the grid cells with no activity during the experiments have not been considered for classification with the SVM algorithm. However, they have been classified in the MRF segmentation step. Table V shows the semantic classification accuracy results using the four standard kernels and the four properties of the space. Fig. 11 shows the classification results for the linear, polynomial, and RBF kernels.

It is interesting to notice that although the sigmoid kernel classified every grid cell as S (which is obviously wrong), it still got 65.12% correct classification results. This happens because the ground truth data indicates that a larger part of the map is indeed supposed to be classified as S. In this case, the visual results have to be taken into account when one is evaluating the performance of the classifiers.

As opposed to the results obtained in the terrain mapping classification, where the RBF kernel obtained the best results, Table V shows that the best classification performance for the activity-based mapping was obtained with the linear kernel. The reason for the poor classification of the RBF kernel can be

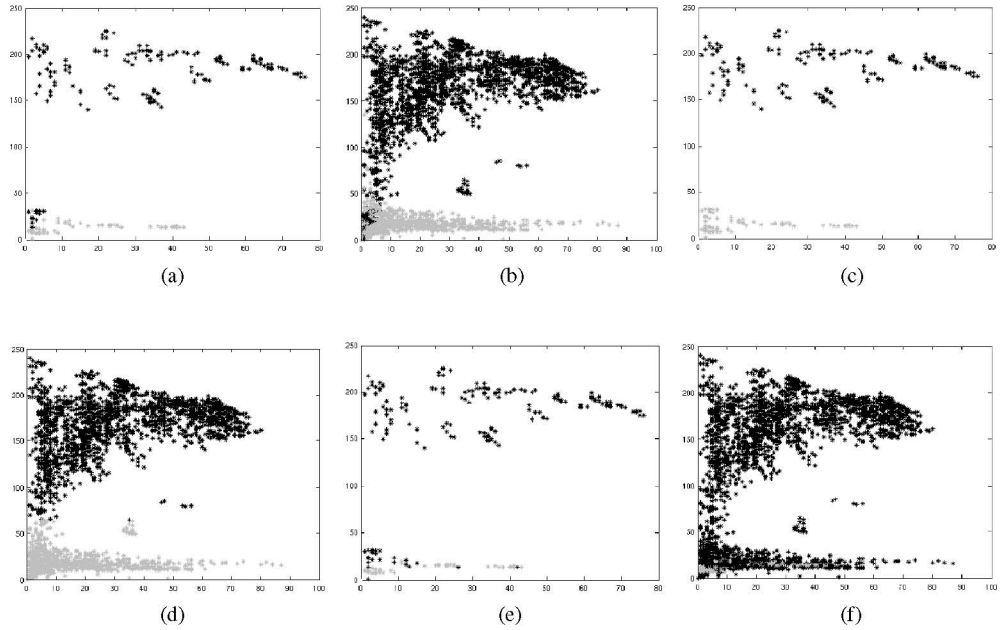


Fig. 12. Results of the SVM semantic classification for the learning and complete datasets using the properties (1) activity and (4) average size (S in dark grey, W in light grey). Learning dataset overfits and performs poor general classification. (a) Learning dataset ground truth. (b) Complete dataset ground truth, (c) Learning dataset classified using linear kernel. (d) Complete dataset classified using the linear kernel. (e) Learning dataset classified using the RBF kernel. (f) Complete dataset classified using the RBF kernel.

TABLE V
ACCURACY OF THE SVM ACTIVITY-BASED SEMANTIC CLASSIFICATION USING THE FOUR PROPERTIES OF THE SPACE

Kernel	SVM	SVM+MRF
Linear	79.96%	80.12%
Polynomial	78.88%	79.19%
RBF	66.69%	65.12%
Sigmoid	65.12%	65.12%

TABLE VI
ACCURACY OF THE SVM SEMANTIC CLASSIFICATION FOR THE LEARNING AND TESTING DATASETS USING THE PROPERTIES (1) ACTIVITY AND (4) AVERAGE SIZE

Kernel	Learning dataset	Testing dataset
Linear	92.25%	79.78%
Polynomial	92.25%	77.57%
RBF	97.06%	69.27%

explained with an analysis of the data presented in Table VI, which shows the classification accuracy results for the learning and complete datasets using only properties (1) and (4). The reasons that only these two properties have been chosen for analysis are that the classification results obtained from them are very similar to the ones obtained using all four properties and it is possible to visualize the classification results in a 2-D graph (Fig. 12). As can be noticed in Table VI, the classification results for the learning dataset using the RBF kernel are better than the ones obtained with the linear kernel. However, the same performance is not obtained with the complete dataset. This strongly suggests an overfitting of the learning dataset, resulting in a poor classification of the complete dataset. This fact can be confirmed in Fig. 12, which shows the classification in the property space. For the learning dataset, the RBF kernel [Fig. 12(e)] obtained very accurate classification results compared to the linear kernel [Fig. 12(c)]. However, when the classification is generalized to the complete dataset, the linear kernel [Fig. 12(d)] is considerably more efficient than the RBF kernel [Fig. 12(f)]. The results obtained with the polynomial kernel were similar to the ones obtained with the linear kernel for the learning dataset, but they were not as good as the ones with the complete dataset.

TABLE VII
ACCURACY RESULTS OF THE SVM SEMANTIC CLASSIFICATION USING COMBINATIONS OF PROPERTIES OF THE SPACE (1—ACTIVITY, 2—OCCUPANCY, 3—MAXIMUM SIZE, AND 4—AVERAGE SIZE)

Properties	SVM	SVM+MRF	Category
1	65.12%	65.12%	C
2	65.12%	65.12%	C
3	78.53%	78.88%	B
4	79.64%	79.77%	A
1,2	65.12%	65.12%	C
1,3	79.20%	78.88%	B
1,4	79.78%	79.88%	A
2,3	79.40%	79.19%	B
2,4	79.85%	79.90%	A
3,4	79.48%	79.41%	A
1,2,3	79.14%	79.01%	B
1,2,4	79.87%	79.90%	A
1,3,4	79.87%	79.90%	A
2,3,4	79.81%	79.96%	A

Besides the experiments with all four properties of the environment, different combinations of properties have also been tested. The accuracy results are shown in Table VII. We grouped the classification results in three categories such that the difference between results in the same category is very minor. Analyzing property combinations that belong to each group, it is

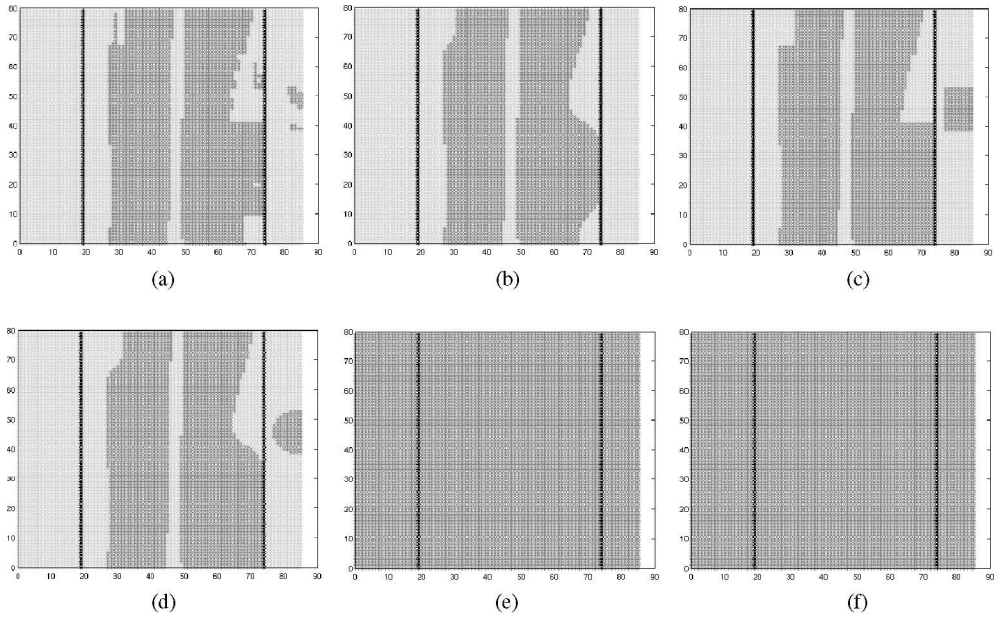


Fig. 13. Results of the SVM semantic classification for the three categories (S in dark grey, W in light grey). Results of the categories A and B have correctly distinguished the street and the sidewalks. (a) Category A: SVM. (b) Category A: SVM + MRF. (c) Category B: SVM. (d) Category B: SVM + MRF. (e) Category C: SVM. (f) Category C: SVM + MRF.

possible to notice how each property contributes to the classification results. Fig. 13 shows the classification results for the three categories, where the environment properties have been numbered as follows: (1) activity, (2) occupancy, (3) maximum size, and (4) average size. The results are based only on the linear kernel as it obtained the most accurate classification among the kernels tested.

Category A presents the most accurate results. Notice that after the MRF segmentation, all the space corresponding to the sidewalk has been correctly classified [Fig. 13(b)]. There are some classification errors in the space that corresponds to the street, mainly in the region close to the black line. This can be explained by the fact that almost no activity happens in that region. In most cases, cars (which characterize the streets) occupy the center of the map. In Fig. 13(a), it is even possible to notice the space that divides the two lanes of the street, which is also not used by most of the cars. All the property combinations that belong to category A (and only these) exploit the average size (4) as a property. Similar to the results obtained with the HMM classification approach, this property leads to the most accurate classification results.

Category B presents good classification results with considerable similarities to Category A, except for a classification error on a small area in the right side of the map. Category B includes property combinations that include maximum size (3) but not property average size (4). The results are also similar to the ones obtained by the HMM approach using maximum size as a property. This can be explained by the fact that at a specific moment during test, a crowd of people stopped in front of the robot, which was placed at that location. As most of the space around the robot was occupied by moving entities, the range sensors detected a large-sized obstacle at that location, which matches with the large obstacles that run in the streets (cars).

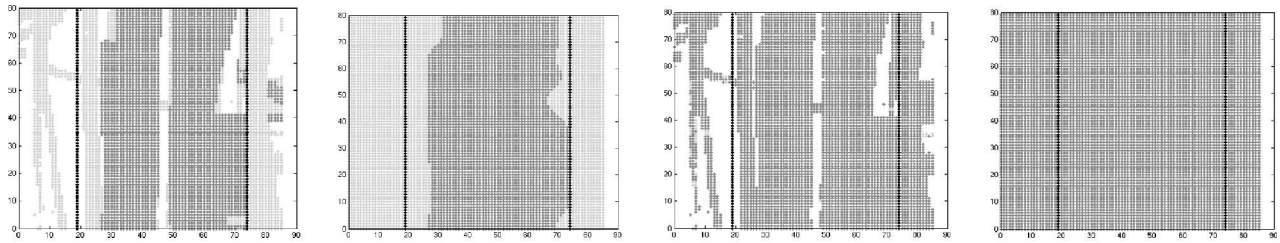
Category C, which includes only the properties activity (1) and occupancy (2) presented very poor classification results. It could not correctly distinguish between street and sidewalks and wrongly classified the entire space as street.

The accuracy results presented in Table VII suggest how each property of the space contributes to the overall classification. The properties average size and maximum size lead to reasonable results while occupancy and activity do not provide enough information to correctly differentiate the environment into street and sidewalk. The data collected during our experiments show that activity and occupancy of the moving entities that occupy the street and the sidewalk are very similar. This is a typical case where there is no association between the input properties and desired classification patterns. Both HMM and SVM methods failed when only these two properties were available.

During the experiments, there was some space in the environment that did not register any activity. In the classification results presented earlier, these areas were also classified as street (S) or sidewalk (W). If we consider those particular regions as a separate class (e.g., nonused space), better classification results can be obtained. The classification results are presented in Table VIII and in Fig. 14, where the MRF segmentation

TABLE VIII
ACCURACY OF THE SVM SEMANTIC CLASSIFICATION EXCLUDING THE NONUSED SPACE

Kernel	SVM	SVM+MRF
Linear	86.14%	85.96%
Polynomial	79.93%	80.04%
RBF	76.09%	67.53%
Sigmoid	75.40%	67.47%



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

Fig. 14. Results of the SVM semantic classification (W in light grey, S in dark grey, and nonused space in white). (a) SVM linear kernel. (b) SVM linear kernel + MRF. (c) SVM RBF kernel. (d) SVM RBF kernel + MRF.

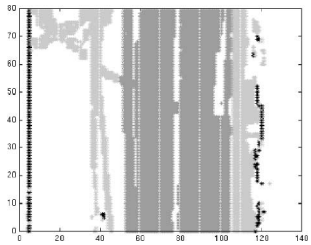


Fig. 15. Multiclass classification ground truth (S in dark grey,w in light grey, and E in black).

technique is used to estimate the classification of the nonused space after the initial SVM classification.

C. Multiclass SVM Classification

For all the semantic classification problems presented in this paper, the environment has been divided into two categories: navigable and nonnavigable in the terrain mapping context, and street and sidewalk for the activity-based mapping. A natural extension of approaches presented in this paper would be multiclass semantic classification. In fact, the SVM algorithm was originally developed for binary classification, but it has been extended to cases in which there are more than two classes [6]. This approach has been tested in the activity-based semantic mapping context. As opposed to the experiments previously described, where only the information obtained from the dynamic entities has been taken to account, in the multiclass case, the static entities are also considered. The environment has been divided into three classes: street (S), sidewalk (W), and static entities (E). Building structures, trees, and all the other parts of the environment that do not move over time are considered part of E. The main characteristics of the elements in this class are high occupancy and very low activity. The ground truth map can be seen in Fig. 15, where the black color corresponds to the class E. Grid cells with no activity have not been considered for classification. Table IX shows the multiclassification results. Only combinations of properties that include the properties 1 and 2 have been presented. All the other property combinations could not classify elements in class E correctly. Only the linear and polynomial kernel obtained reasonable classification results. Fig. 16 shows the classification results with and without MRF segmentation. The best classification results were obtained after the MRF segmentation with the combination of properties (1), (2), and (4). This can be clearly noticed in Fig. 16(c).

TABLE IX
ACCURACY OF THE SVM MULTICLASS SEMANTIC CLASSIFICATION

Kernel	Properties	SVM	SVM+MRF
Linear	1,2,3,4	85.37%	84.34%
Linear	1,2,3	85.20%	86.13%
Linear	1,2,4	85.13%	88.93%
Polynomial	1,2,3,4	86.15%	85.62%
Polynomial	1,2,3	87.71%	89.29%
Polynomial	1,2,4	89.17%	95.67%

VIII. CONCLUSION AND FUTURE WORK

This paper approached the problem of semantic mapping, which consists of creating robot-based maps that go beyond representing the metric structure of the environment. Semantic maps can represent other properties of the environment, allowing a more complex and complete description of the space. The semantics present in the maps also allow robots to more easily share information with people and ultimately perform more complex tasks.

Two scenarios have been used as a test bed for our mapping approaches: terrain mapping and activity-based mapping. The terrain mapping problem consists of creating 3-D representations and classifying terrain into either navigable or nonnavigable areas. The activity-based mapping problem consists of creating 2-D maps that classify the environment according to the usage of the space by dynamic entities.

Two approaches for the semantic mapping problem have been presented. The first one is based on HMMs and the second on SVMs. A fundamental difference between the HMM and the SVM semantic classification methods is that, in the HMM approach, each data sequence is considered at once, while in the SVM algorithm, each point is individually classified. In the terrain mapping domain, the data sequences correspond to the 3-D points in the range scans. In the activity-based mapping context, the rows in the grid of cells correspond to the data sequences. In order to classify each point in the data sequence, the previously made classification is taken into account. This characteristic does not necessarily lead to better classification results; it depends on the nature of the data to be classified. In most cases, when the data is divided into well-defined clusters, the HMM method tends to be more efficient (exploring the effect of locality), while the SVM approach is theoretically better for

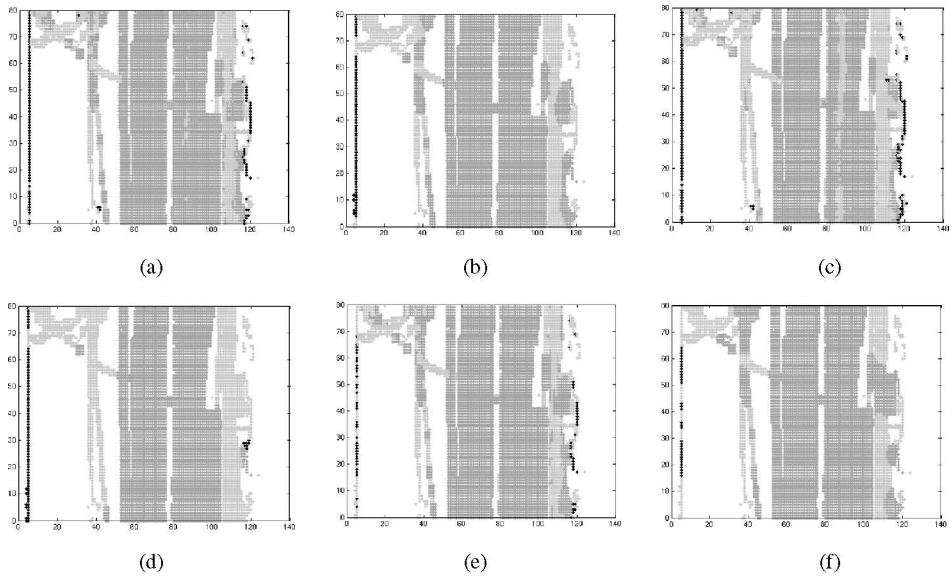


Fig. 16. Results of the SVM multi class semantic classification using the polynomial kernel. Most part of the static entities in the environment have been identified (S in dark grey, W in light grey, and E in black). (a) Properties 123 (b) Properties 123 with MRF (c) Properties 124 (d) Properties 124 with MRF (e) Properties 1234 (f) Properties 1234 with MRF.

nonclustered data. Another important difference between these two learning methods is the fact that the SVM can handle several input properties while only one can be used in our particular implementation of the HMM. For most experiments performed in this paper, the classification results of the two methods are very similar and noticeably better than the standard histogram-based classification algorithm. Another important conclusion obtained from the data analyzed is that not all properties of space lead to the desired classification. In the activity-based mapping problem, both HMM and SVM failed when only the properties activity and occupancy were available.

As semantic mapping is still a young research topic, there are many open problems and interesting directions for new research in the field. Differentiating entities (e.g., people) in an urban environment, understanding and possibly predicting the behavior of dynamic entities, and improving the complexity of the map representation by fusing information provided by other types of sensors (e.g., cameras) are all part of the future work to be addressed on this topic.

ACKNOWLEDGMENT

The authors thank D. Fox and W. Burgard for the initial ideas and discussions on terrain mapping using an HMM.

REFERENCES

[1] D. Anguelov, R. Biswas, D. Koller, B. Limketkai, S. Sanner, and S. Thrun, "Learning hierarchical object maps of non-stationary environments with mobile robots," in *Proc. Annu. Conf. Uncertainty Artif. Intell.*, 2002, pp. 10–17.

[2] J. Borenstein and Y. Koren, "The vector field histogram—Fast obstacle-avoidance for mobile robots," *IEEE J. Robot. Autom.*, vol. 7, no. 3, pp. 278–288, Jun. 1991.

[3] B. E. Boser, I. Guyon, and V. Vapnik, "A training algorithm for optimal margin classifiers," in *Proc. Comput. Learn. Theory*, 1992, pp. 144–152.

[4] N. Brand, N. Oliver, and A. Pentland, "Coupled hidden Markov models for complex action recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 1997, pp. 994–999.

[5] C. J. C. Burges, "A tutorial on support vector machines for pattern recognition," *Data Min. Knowl. Discov.*, vol. 2, no. 2, pp. 121–167, 1998.

[6] K. Crammer and Y. Singer, "On the algorithmic implementation of multiclass kernel-based vector machines," *J. Mach. Learn. Res.*, vol. 2, pp. 265–292, 2001.

[7] F. Dellaert and D. Bruemmer, "Semantic slam for collaborative cognitive workspaces," presented at the AAAI Fall Symp. Ser., Arlington, VA, 2004.

[8] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification, 2nd ed.* New York: Wiley, 2001.

[9] A. Elfes, "Sonar-based real-world mapping and navigation," *IEEE Trans. Robot. Autom.*, vol. RA-3, no. 3, pp. 249–265, Jun. 1986.

[10] G. D. Forney Jr., "The Viterbi algorithm," *Proc. IEEE*, vol. 61, no. 3, pp. 268–278, Mar. 1973.

[11] D. A. Forsyth and J. Ponce, *Computer Vision: A Modern Approach, 1st ed.* Englewood Cliffs, NJ: Prentice-Hall, 2002.

[12] C. Galindo, A. Saffiotti, S. Coradeschi, P. Buschka, J. A. Fernandez-Madrigal, and J. Gonzalez, "Multi-hierarchical semantic maps for mobile robotics," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2005, pp. 3492–3497.

[13] B. P. Gerkey, R. T. Vaughan, K. Stoy, A. Howard, G. S. Sukhatme, and M. J. Mataric, "Most valuable player: A robot device server for distributed control," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2001, pp. 1226–1231.

[14] V. Harmandas, M. Sanderson, and M. D. Dunlop, "Image retrieval by hypertext links," in *Proc. Int. Conf. Res. Dev. Inf. Retr.*, 1997, pp. 296–303.

[15] T. Joachims. (2005, Dec.). *SVM Light* [Online]. Available: http://svmlight.joachims.org/

[16] D. Kim, J. Sun, S. M. Oh, J. M. Rehg, and A. Bobick, "Traversability classification using unsupervised on-line visual learning for outdoor robot navigation," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2006, pp. 518–525.

[17] R. Kindermann and J. L. Snell, *Markov Random Fields and Their Applications*. Providence, RI: American Mathematical Society, 1980.

[18] M. D. Klemen, E. R. Weippl, and A. M. Tjoa, "The semantic desktop: A semantic personal information management system based on RDF and topic maps," in *Proc. Workshop Ontol.-Based Tech. DataBases Inf. Syst.*, 2005, pp. 135–151.

[19] K. Konolige, M. Agrawal, R. C. Bolles, C. Cowan, M. Fischler, and B. Gerkey, "Outdoor mapping and navigation using stereo vision," presented at the Int. Symp. Exp. Robot., Rio de Janeiro, Brazil, Jul. 2006.

[20] B. Limketkai, L. Liao, and D. Fox, "Relational object maps for mobile robots," in *Proc. Int. Joint Conf. Artif. Intell.*, 2005, pp. 1471–1476.
[21] Y Liu, N. Lazar, and W. Rothfus, "Semantic-based biomedical image indexing and retrieval," in *Proc. Int. Conf. Diagn. Imag. Anal.*, 2002, pp. 18–20.
[22] A. Lookingbill, D. Lieb, D. Stavens, and S. Thrun, "Learning activity-based ground models from a moving helicopter platform," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2005, pp. 3948–3953.
[23] J. Modayil and B. Kuipers, "Bootstrap learning for object discovery," in *Proc. Int. Conf. Intell. Robots Syst.*, 2004, pp. 742–747.
[24] M. D. Buhmann and M. J. Ablowitz, *Radial Basis Functions: Theory and Implementations*. Cambridge, U.K.: Cambridge Univ. Press, 2003.
[25] O. M. Mozos, C. Stachniss, and W. Burgard, "Supervised learning of places from range data using Adaboost," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2005, pp. 1742–1747.
[26] NASA. (2005, Mar.) [Online]. Available: http://ranier.hq.nasa.gov/telerobotics_page/technologies/0240.html
[27] C. W. Nielsen, B. Ricks, M. A. Goodrich, D. Bruemmer, D. Few, and M. Walton, "Snapshots for semantic maps," in *Proc. IEEE Conf. Syst., Man, Cybern.*, 2004, pp. 10–13.
[28] W. Niu, D. H. Han, and Y. F. Wang, "Human activity detection and recognition for video surveillance," in *Proc. IEEE Multimedia Expo. Conf.*, 2004, pp. 719–722.
[29] A. Nuchter, O. Wulf, K. Lingemann, J. Hertzberg, B. Wagner, and H. Surmann, "3D mapping with semantic knowledge," presented at the RoboCup Int. Symp., Osaka, Japan, 2005.
[30] L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proc. IEEE*, vol. 77, no. 2, pp. 257–286, Feb. 1989.
[31] N. Rishe, "Efficient organization of semantic databases," in *Proc. Int. Conf. Found. Data Org. Algorithms*, 1989, pp. 114–127.
[32] A. Rottmann, O. Martinez Mozos, C. Stachniss, and W. Burgard, "Place classification of indoor environments with mobile robots using boosting," in *Proc. Nat. Conf. Artif. Intell. (AAAI)*, 2005, pp. 1306–1311.
[33] D. Schröter, "Region & gateway mapping: Acquiring structured object-oriented representations of indoor environments," Ph.D dissertation, Dept. Image Understanding Knowl.-Based Syst., Tech. Univ. München,, Munchen, Germany, 2005.
[34] C. Stachniss and W. Burgard, "Mobile robot mapping and localization in non-static environments," in *Proc. AAAI Nat. Conf. Artif. Intell.*, 2005, pp. 1324–1329.
[35] S. Thrun, "Robotic mapping: A survey," in *Exploring Artificial Intelligence in the New Millenium*. San Mateo, CA: Morgan Kaufmann, 2002.
[36] V. Vapnik, *Estimation of Dependences Based on Empirical Data*. Moscow, Russia: Nauka, 1979 [English translation. New York: Springer-Verlag, 1982].
[37] S. Vasudevan, V. T. Nguyen, and R. Siegwart, "Towards a cognitive probabilistic representation of space for mobile robots," in *Proc. IEEE Int. Conf. Inf. Acquis.*, Aug., 2006, pp. 353–359.
[38] C. Wellington, A. Courville, and A. Stentz, "Interacting Markov random fields for simultaneous terrain modeling and obstacle detection," presented at the Robot. Sci. Syst. Conf., Cambridge, MA, Jun. 2005.
[39] D. F. Wolf and G. S. Sukhatme, "Mobile robot simultaneous localization and mapping in dynamic environments," *Auton. Robots*, vol. 19, no. 1, pp. 53–65, 2004.
[40] D. F. Wolf, G. S. Sukhatme, D. Fox, and W. Burgard, "Autonomous terrain mapping and classification using hidden Markov models," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2005, pp. 2038–2043.
[41] D. F. Wolf, A. Howard, and G. S. Sukhatme, "Towards geometric 3D mapping of outdoor environments using mobile robots," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2005, pp. 1258–1263.
[42] C. Ye and J. Borenstein, "A new terrain mapping method for mobile robots obstacle negotiation," in *Proc. UGV Technol. Conf. 2003 SPIE AeroSense Symp.*, pp. 52–62.
[43] H. Zhuge, "Retrieve images by understanding semantic links and clustering image fragments," *J. Syst. Softw. Arch.*, vol. 73, no. 3, pp. 455–466, 2004.

Denis F. Wolf received the B.S. degree from the Federal University of Sao Carlos, Sao Carlos, Sao Paulo, Brazil, in 1999, the M.S. degree from the University of Sao Paulo, Sau Paulo, Brazil, in 2001, and the Ph.D. degree from the University of Southern California (USC), Los Angeles, in 2006, all in computer science.

He is currently an Assistant Professor in the Computer Systems Department in the Mathematics and Computer Science Institute at the University of Sao Paulo. He is the Director of the Mobile Robotics Laboratory, which he founded in 2007. He has published more than 20 technical papers. His current research interests include mobile robotics, embedded systems, and machine learning.

Gaurav S. Sukhatme (SM'05) received the M.S. and Ph.D. degrees in computer science from University of Southern California (USC), Los Angeles.

He is an Associate Professor of computer science (joint appointment in electrical engineering) with USC. He is the Codirector of the USC Robotics Research Laboratory and the Director of the USC Robotic Embedded Systems Laboratory, which he founded in 2000. His current research interests include multirobot systems, sensor/actuator networks, and robotic sensor networks. He has published extensively in these and related areas.

Prof. Sukhatme was a recipient of the National Science Foundation (NSF) CAREER Award and the Okawa Foundation Research Award. He has served as PI on several NSF, DARPA, and NASA grants. He is a Co-PI on the Center for Embedded Networked Sensing (CENS), an NSF Science and Technology Center. He is a member of the Association of Advacement of Artificial Intelligence (AAAI) and the Association for Computing Machinery (ACM), and is one of the founders of the Robotics: Science and Systems conference. He is a Program Chair of the 2008 IEEE International Conference on Robotics and Automation and the Editor-in-Chief of Autonomous Robots. He has served as an Associate Editor of the IEEE TRANSACTIONS ON ROBOTICS AND AUTOMATION, the IEEE TRANSACTIONS ON MOBILE COMPUTING, and on the Editorial Board of *IEEE Pervasive Computing*.