Laser Range Data Based Semantic Labeling of Places

L. Shi, S. Kodagoda and G. Dissanayake, *Member, IEEE*

Abstract – **Extending metric space representations of an environment with other high level information, such as semantic and topological representations enable a robotic device to efficiently operate in complex environments. This paper proposes a methodology for a robot to classify indoor environments into semantic categories. Classification task, using data collected from a laser range finder, is achieved by a machine learning approach based on the logistic regression algorithm. The classification is followed by a probabilistic temporal update of the semantic labels of places. The innovation here is that the new algorithm is able to classify parts of a single laser scan into different semantic labels rather than the conventional approach of gross categorization of locations based on the whole laser scan. We demonstrate the effectiveness of the algorithm using a data set available in the public domain.**

I. INTRODUCTION

SEMANTIC labels of places enables a robot to perceive the environment at a level more conceptual than what is feasible with purely metric maps. This results in a common representation of high level information, which can be effectively and efficiently shared between humans and robots. Hence, the humans could have a better understanding of the robots and the robots can eventually be designed with the capability of carrying out complex tasks cooperatively with a human.

Robots perceive the environment through information gathered from a variety of sensors. Most popular perceptual sensors that have been used in robotics are cameras and laser range finders. Either of these sensors can be used in semantic labeling of places. For instance, Shi et al [1] and Viswanathan et al [6] proposed environment classification algorithms using lines or objects extracted from vision data as features. Mozos et al [7] and Sousa et al [8] classified the environment based on data collected from laser range finders. Although there are environment labeling techniques based on multiple sensory data reported in the literature [18], it is our belief that the potential of laser range finders has not yet been well exploited. This paper addresses this issue through a combination of classification and estimation techniques. Following describes the literature most relevant to the approach proposed in this paper.

Poncela et al [3] adopted Principal Component Analysis (PCA) to classify the environments into walls and doors, and

Buschka et al [2] proposed a rectangular-fit algorithm to incrementally extract room-like nodes and automatically segment the space into room and corridor regions. These approaches have limitations due to the strong assumptions of width and length of a certain space. Tapus et al [5] proposed a complex Bayesian approach for topology recognition and door detection. Mozos et al [7] extracted a variety of simple features from laser range data and employed the AdaBoost classifier to label an environment consisting of rooms, corridors, doorways and halls; and Sousa et al [8] obtained better classification results using Support Vector Machines (SVM). In both cases, the location of the robot from which the data was obtained was classified into semantic labels rather than labels for sections of the environment.

In this paper, we first extract a set of features from laser range/bearing data and use these in a binary classifier based on logistic regression. The outcomes are analysed to select three dominant features most effective during the classification process. Results of the classification are in probabilistic form which enables the use of the independent opinion pool approach to fuse temporal probabilities, resulting in a labeled semantic grid map. A method for removing outliers in each single scan for further improving the accuracy of labeling is also presented.

The rest of this paper is arranged as follows. Section II discusses the details about the classifier and the feature selection strategy. Data fusion and map building processes following the classification is discussed in the Section III. An outlier removal approach, which is used to improve the labeling accuracy of generated semantic grid map, is introduced in Section IV. In Section V, experimental results are presented and discussed; and Section VI concludes the paper.

II. CLASSIFICATION

A. Semantic Labels

Much of the existing literature classifies typical office environments using labels, "rooms", "corridors" and "doorways" [7]. It is challenging to robustly recognize doorways as they are observed infrequently and the status of the door (opened/closed) significantly influences the character of the data. Therefore, in our research, rooms and corridors are adopted as semantic labels, which has lead to a simpler binary classifier. Doorways are treated as transit areas between rooms and corridors.

B. Feature Selection

Feature selection plays an important role in supervised machine learning problems because it affects the ability of generalization, overhead and over-fitting of the system. As is widely accepted for some problems, a small subset of features is sufficient to approximate the target concept [11], thus finding the dominant features becomes a key issue.

As a general observation, it could be seen that the laser range/bearing data plotted in a Cartesian coordinates referring to a corridor is of a long rectangular nature and that of a room is of a short rectangular/square form. However, object in the environment, such as furniture will make a classification system based on such simple feature erroneous.

There are various features that could be used for semantic labeling of places. Mozos et al [7] extracted two sets of simple features from laser range data. One set was from raw range data and the other was from polygonal approximation of the observed area. They employed about 150 single-valued features (considering different thresholds) of 22 categories, and used a multi-class AdaBoost classifier to process these. Similarly, Sousa et al [8] selected 14 single-valued features and a binary SVM classifier to perform the same task.

We evaluated the performances of different feature combinations using a binary L2-regularized logistic regression classifier, which led us to choose three dominant features out of the 150 features used in [7].

C. Logistic Regression

As stated in section II.B, two main classifiers used in the literature are AdaBoost and Support Vector Machines (SVM). While AdaBoost constructs a strong classifier with a linear combination of many weak classifiers, SVM performs classification in a high-dimensional feature space and has the advantage that training sample set can be small [16][17]. In this work, we have chosen the Logistic Regression as a classifier because the probabilistic nature of its output provides us with a seamless way of integrating with the mapping process described in section III.B.

Logistic Regression is an approach to learn functions of the form $P(y | \vec{x})$ in case where y is discrete-valued, and $\vec{x}$ is any vector containing discrete or continuous variables. It assumes a parametric form of the distribution, $P(y | \vec{x})$, while directly estimating its parameters from the training data [9]. Although it is a binary classifier, it can be extended to deal with multi-class classification as well [12].

The parametric model for binary classification, given data $\vec{x}$ and parameters or weights $\vec{w}$ is:

$$P(y = \pm 1 | \vec{x}; \vec{w}) = \frac{1}{1 + \exp\left(-y \cdot \vec{w}^T \cdot \vec{x}\right)} \quad (1)$$

Let the ith training sample be $(y_i, \vec{x}_i)$, where $\vec{x}_i$ is the feature set and $y_i \in \{+1, -1\}$ is the label of the ith training sample. Then the objective of the training task is to minimize the negative log-likelihood,

$$\min_{\vec{w}} \left(\sum_{i=1}^{m} \log\left(1 + \exp\left(-y_i \cdot \vec{w}^T \cdot \vec{x}_i\right)\right) \right) \quad (2)$$

However, over fitting is a potential risk of logistic regression, especially with sparse high dimensional data [9]. Hence regularization which encourages the fitted parameters to be small is usually employed to reduce over-fitting [11].

Both L1 and L2 regularization algorithms are commonly used for this purpose. The former encourages the sum of absolute values of the parameters to be small and the latter encourages the sum of squares of the parameters to be small [11]. In this paper, the L2 regularization is used.

L2-regularized algorithm solves the following problem.

$$\min_{\vec{w}} \left(\frac{1}{2} \vec{w}^T \vec{w} + C \sum_{i=1}^{m} \log\left(1 + \exp\left(-y_i \cdot \vec{w}^T \cdot \vec{x}_i\right)\right) \right) \quad (3)$$

where $C > 0$ is a penalty parameter.

III. DATA FUSION

In this section, we propose a way of labeling the environment based on laser range/bearing data. The labeling process becomes complex when open doors are present causing a robot to perceive multiple environment types from one location. We propose a probabilistic method to update the temporal information to solve this problem.

A. Independent Opinion Pool

Consider laser range/bearing observations made independently from a set of unique positions of the robot, we propose to use the independent opinion pool method to probabilistically update a semantic grid map of the robot's environment.

In the viewpoint of data fusion, the output of a sensor could be either a likelihood $P(z | x)$ or an opinion $P(x | z)$, where z is an observation and x is a state of the target (e.g. point being in a room). There are three common approaches to combine these probabilistic evidences: linear opinion pool, independent opinion pool and independent likelihood pool [19][20]. The structure of this application leads us to make use of the independent opinion pool method. An example of a similar scenario is analyzed in detail in the literature [19][21].

The independent opinion pool method can be described as follows:

$$P(x | z_1, ..., z_n) = \alpha \cdot \prod_{i}^{n} P(x | z_i) \quad (4)$$

where z_i is the ith observation of a particular grid point, x is the state (semantic label) of that grid point, and α is a normalizing factor. In our application, all possible states (semantic label) of grid points are exclusive, i.e. the posterior probabilities of the point belonging to a certain state sums up to one.

Equation (4) can be rewritten in a recursive way as follows:

$$P(x | z_1, ..., z_n) \propto P(x | z_1, ..., z_{n-1}) \cdot P(x | z_n) \quad (5)$$

B. Map Update

Using equation (5), semantic grids can be updated as independent observations accumulatively. Note that the semantic grid map is analogous to its counterpart, occupancy grid map [23].

IV. PREPROCESSING FOR MAP BUILDING

The semantic map that is built using the independent opinion pool method is not capable of handling the problems due to open doors. Therefore a further processing step is required to improve the labeling accuracy.

A. Inliers and Outliers

In the following, inliers are defined as the laser range/bearing data in a particular scan belonging to a specific semantic label. Outliers are the remainder of the data. For example, if a robot in an office room detects obstacle points both in the same office room and in a corridor outside the door, then the former are called inliers and the latter are called outliers. If we can discriminate inliers and outliers before performing data fusion, the detected semantic grid points can be updated selectively, which leads to a higher labeling accuracy.

B. Door Detection

One approach to discriminate inliers and outliers in a single laser scan is by detecting doors. Many door detection techniques can be found in the literature. ElKaissi et al [13] devised a door detection system using laser range data based on the assumption of a door width. Anguelov et al [14] employed laser and vision sensors to detect doors by their width, color, and dynamic states. Carinena et al [4] applied fuzzy temporal rules to detect doors. In this paper, we propose to combine some of these ideas to realize a simple door detector using the following heuristics.

1) Door is represented by two break points (door frame points), which are supported by "walls";
2) Distance between the two break points should be within a certain range (a reasonable door width);
3) All laser range data that belong to a detected door opening is considered to be associated with a different labels or an unknown label.

With regard to the implementation of this algorithm, firstly the laser range/bearing data is processed to detect discontinuity points using Extended Kalman Filter (EKF) based segmentation [15]. Then the discontinuity points are combined in pairs to be tested using the door detector to find potential doors.

C. Selective Update

With the detection of potential doors, the laser range/bearing data of a single scan can be updated selectively in the accumulatively labeling process described in section III.B. For the inliers, the labels are updated as usual. However, for the outliers, the labels are kept unchanged as they belong to an "unknown" class of labels; or apparently, their new label information $P(x | z_1, ..., z_n)$ is the same as their historical labels $P(x | z_1, ..., z_{n-1})$.

V. RESULTS

A. Environment and Dataset

Dataset including raw laser range/bearing data and associated feature for a robot operating in an office like environment are used for the analysis (the dataset published by Oscar [10]). This environment consists of a long corridor with several rooms of different sizes (with different furniture) as shown in Fig. 1. The data is collected by a robot equipped with a frond and a rear two dimensional (2D) laser range finders, which have a field of view of 360 degrees. Training samples are chosen from the left half of the environment, whereas the right half of the environment is used for testing.

B. Feature Selection

As mentioned before (Section II.B), Mozos et al [7][22] derived two sets of features from raw range/bearing data (which is called the B series) and from a polygonal approximation of the observed area (which is called the P series). The definitions of these features are described in the literature [7][8][22].

In this experiment, a L2-regularized logistic regression classifier is employed to segment the environment into three categories based on different combinations of features. Many different feature combinations were evaluated and the most prominent results are listed in TABLE I.

TABLE I
PERFORMANCE OF DIFFERENT FEATURE COMBINATIONS

Feature Combination	Training Error	Testing Error	Total Error	Execution Time
All 150 features	2.21 %	1.97 %	4.18 %	15292ms
All 21 single-valued features	2.37 %	2.09 %	4.46 %	2517 ms
All 11 single-valued P series	2.65 %	2.40 %	5.05 %	1175 ms
All 10 single-valued B series	2.97 %	2.57 %	5.54 %	1243 ms
3 selected features	2.62 %	2.12 %	4.73 %	361 ms

In the table above, the three selected features extracted from a single laser scan are:
- The standard deviation of the difference between the lengths of consecutive ranges
- The standard deviation of ranges
- The area of polygonal approximation

TABLE I indicates that feature selection is a tradeoff between accuracy and overhead (or computational complexity). Given the need for real-time performance, the three selected features described above are chosen for illustrating the labeling process.

C. Classification Results

In this section, laser range/bearing data is classified into two categories: "room" and "corridor". For this purpose, 51353 laser range scans (43957 taken from rooms and 7396 taken

from corridors) were used as the training dataset. To evaluate the performance of the classifier, a testing dataset comprising of 38307 laser range scans (32214 taken from rooms and 6093 taken from corridors) from right hand side of the environment (Grey area in Fig. 1) were used.

L2-regularized logistic regression is used as the binary classifier as implemented in LIBLINEAR toolkit, which is an open source library designed for linear classification [12]. The performance of the classifier is shown in TABLE II, with the test result of entire testing dataset visualized in Fig. 2. It is to be noted that the output of the classifier is the probability of the robot's location in a particular type of environment.

TABLE II
PERFORMANCE OF LOGISTIC REGRESSION CLASSIFIER

Training Error	Testing Error	
0.43%	in Rooms:	< 0.01%
	in Corridors:	1.40%

CONFUSION MATRIX [24]

		Predicted Class	
		Corridor	Room
Actual Class	Corridor	6008	85
	Room	0	32214

Fig. 1. Training dataset: the grey points depict the background map as a reference. Red and blue points are observer's positions which are manually labeled as in *room* and *corridor* respectively.
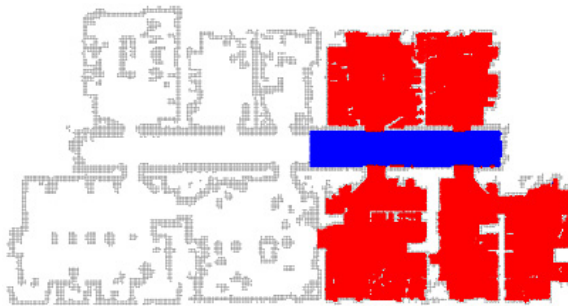

Fig. 2. Testing results: the grey points depict the background map as a reference. Red and blue points are observer's positions which are labeled as in *room* and in *corridor* respectively by the classifier.

According to the TABLE II, accuracy of the classifier is higher when laser observations are made while the robot is in rooms is used. On the other hand laser observations made while the robot is in the corridor results in a slightly lower accuracy.

The simulation data considered above does not contain sensor errors. In order to test the robustness of the classifier,

another experiment was carried out on noisy data generated by adding ±10% range uncertainty to the raw range. Results which are shown in TABLE III show that the classification accuracy has not degraded substantially.

TABLE III
PERFORMANCE OF LOGISTIC REGRESSION CLASSIFIER
ON NOISY RAW DATA

Training Error	Testing Error	
0.63%	in Rooms:	< 0.01%
	in Corridors:	1.51%

CONFUSION MATRIX

		Predicted Class	
		Corridor	Room
Actual Class	Corridor	6001	92
	Room	1	32213

D. Door Detection

Results of door detection algorithm are shown in Fig. 3. It could be noted that the door detection is more accurate while the robot is in corridors (Fig. 3 (a)) than that of in rooms (Fig. 3 (b)). This is attributed to various furniture that is present in room environments, which give rise to door looking features in a 2D laser range/bearing data. Although this problem could be solved by adopting a more complex approach, current results are acceptable for our application with the availability of temporal updates.
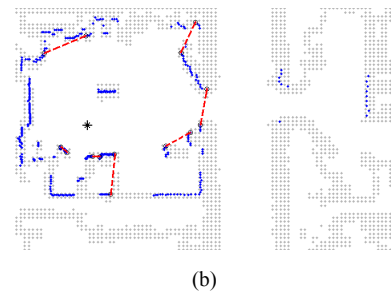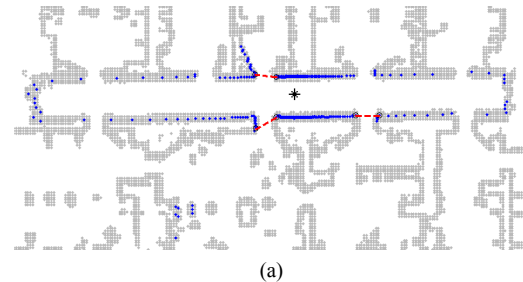

(a)

(b)

Fig. 3. The grey points depict the background map as a reference. The observer marked as a star is in a *corridor* (a)/*room* (b) and the blue points are those detected by the laser sensor. Dashed lines in red indicate the potential doors.

E. Map Update

In this section, we use classification results to label the obstacle points rather than the observer's positions, and fuse temporal information to build a semantic grid map. Fig. 4 shows the labeled grid map built without (Fig. 4(a))/with (Fig. 4(b)) using the outlier removal strategy based on 100 samples of range scans randomly taken from rooms and

corridors. Simulation results show that the overall labeling accuracy of the semantic grid map has been improved from about 93% to about 98% with the outlier removal strategy.
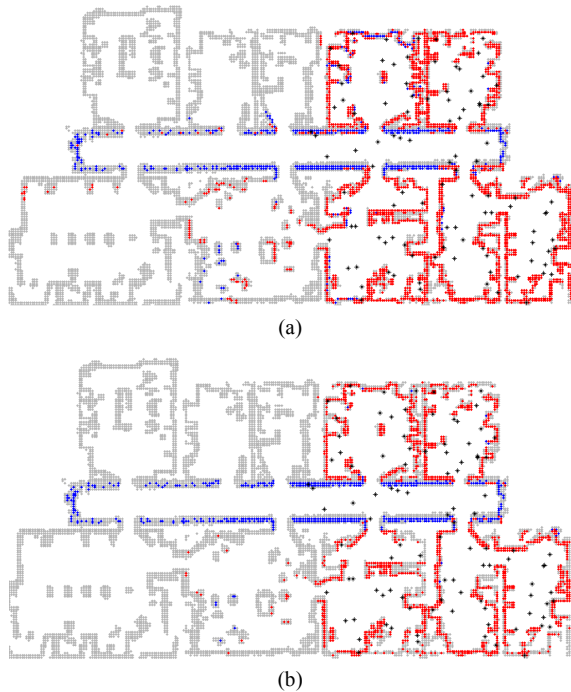

(a)

(b)

Fig. 4. Labeled semantic grid map based on randomly taken 100 scans from rooms and a corridor, (a) with outliers with (b) without outliers. The grey points depict the background map as a reference. The observer's positions are marked as black stars. The red and blue points are detected obstacle points labeled as belonging to rooms and corridor respectively.

The map building algorithm is tested under different sample sizes, with and without door detections. Plots on the relationships between labeling error (relative to the number of all labeled points) and number of samples are shown in Fig. 5.

By observing the results, it could be noted that 1) in all cases, the labeling error rates converge to a steady state value with increasing number of samples; 2) the application of door detection algorithm remarkably reduces the error rate; 3) samples taken from corridor are the main source of labeling error, which implies that more outliers are observed in corridors than in rooms through open doors.
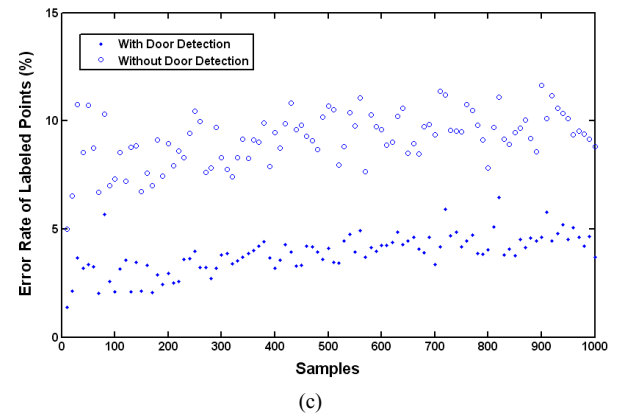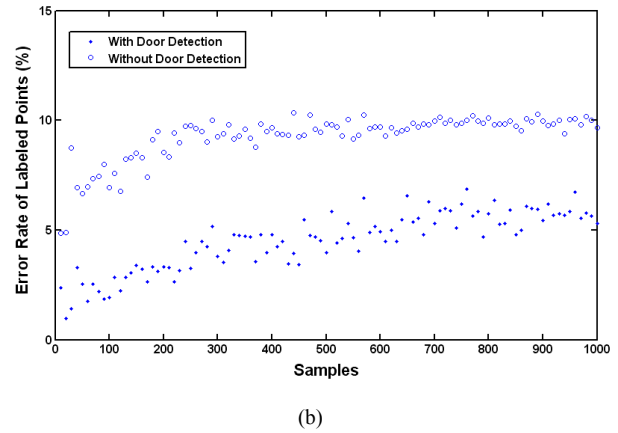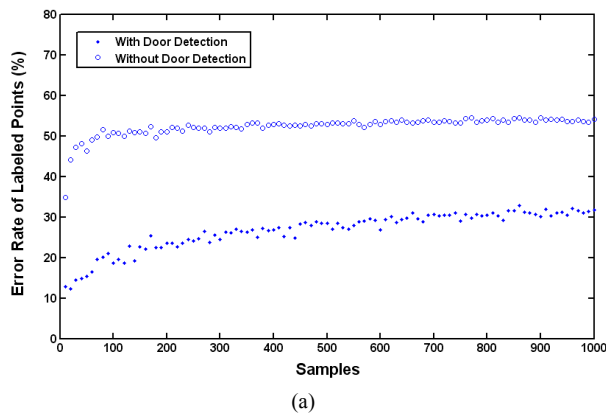

(a)

(b)

(c)

Fig. 5. Labeling error vs. number of samples. These samples are randomly selected from scans taken *in corridors only* (a)/*in rooms only* (b)/*in corridors and rooms* (c), and tested on updating algorithm with and without door detection.

As stated above, the application of door detection has increased labeling accuracy; however, the cost is the loss of information due to some discarded data points. Fig. 6 shows the relation between unlabeled data rate (relative to the number of all detected data points) and number of samples, according to the scenarios described in Fig. 5. Note that the unlabeled rate is constantly zero without the application of door detection.

The result shown in Fig. 6 indicates that unlabeled rate decreases with increasing sample numbers. The specific trend of the plots can be attributed to the specific selection of the environment and the robot poses.
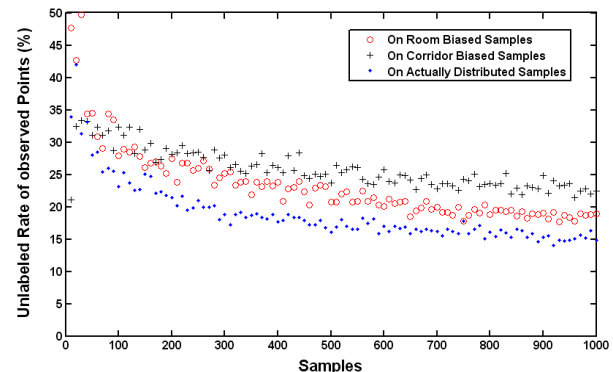

Fig. 6. Unlabeled data rate vs. number of samples, when door detection approach is applied.

In the final experiment, we consider a robot moving from point "A" to "B" following a trajectory as shown in Fig. 7. It is assumed that the robot localization is always known. While the robot is moving from one pose to the other, it generates a labeled semantic grid map as shown in Fig. 7. This shows the capability of the algorithm to explore and label an unknown environment. It also suggests that the robot needs to observe enough to overcome some ambiguous or wrong classifications.
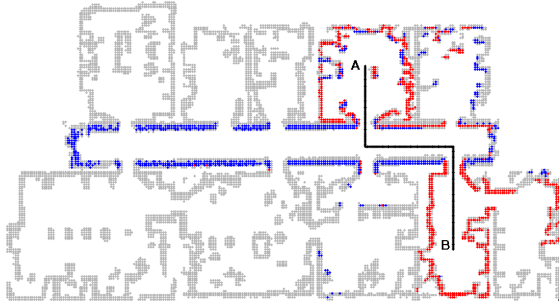

Fig. 7. Labeled semantic grid map built from a robot, which follows the trajectory from point A to point B. The grey points depict the background map as a reference. The red and blue points are detected points and finally labeled as belonging to rooms and corridor respectively.

VI. CONCLUSIONS

In this paper, we have presented an approach to classify the environment around a robot based on laser range/bearing data, and then applied the perceived information together with partial understanding of the geometric structure of the environment to accumulatively build a labeled semantic grid map. Independent opinion pool approach was adopted on fusing temporal information.

Experiments on previously published datasets showed that the classification results are convincing with low training and testing errors. In building the semantic grid map, obstacle points in a single scan were further discriminated as inliers and outliers by detecting doors. The semantic labels belonging to inliers were only updated. This has resulted in improving the labeling accuracy.

Currently, we are extending the algorithm to enable the robot to incorporate more semantic labels while exploring an unknown environment. In addition, we also plan to implement the algorithm on the LISA robot (Lightweight Integrated Social Autobot) to assess real time performance.

ACKNOWLEDGMENT

REFERENCES

[1] W. X. Shi, J. Samarabandu, "Investigating the Performance of Corridor and Door Detection Algorithms in Different Environments", *International Conference on Information and Automation*, pp. 206 - 211, 2006

[2] P. Buschka, A. Saffiotti, "A virtual sensor for room detection", *IEEE/RSJ International Conference on Intelligent Robots and Systems*, vol. 1, pp. 637 - 642, 2002

[3] A. Poncela, C. Urdiales, B. Fernandez-Espejo, F.Sandoval, "Place characterization for navigation via behaviour merging for an autonomous mobile robot", *The 14th IEEE Mediterranean Electrotechnical Conference*, pp. 350 - 355, 2008

[4] P. Carinena, C.V. Regueiro, A. Otero, A.J. Bugarin, S. Barro, "Landmark detection in mobile robotics using fuzzy temporal rules", *IEEE Transactions on Fuzzy Systems*, vol. 12(4), pp. 423 - 435, 2004

[5] A.Tapus, G. Ramel, L. Dobler, R. Siegwart, "Topology learning and recognition using Bayesian programming for mobile robot navigation", *IEEE/RSJ International Conference on Intelligent Robots and Systems*, vol. 4, pp. 3139 - 3144, 2004

[6] P. Viswanathan, D. Meger, T. Southey, J.J. Little, A.K. Mackworth, "Automated Spatial-Semantic Modeling with Applications to Place Labeling and Informed Search", *Canadian Conference on Computer and Robot Vision*, pp. 284 - 291, 2009

[7] O.M. Mozos, C. Stachniss, W. Burgard, "Supervised Learning of Places from Range Data using AdaBoost", *Proceedings of the IEEE International Conference on Robotics and Automation*. pp. 1730 - 1735, 2005

[8] P. Sousa, R. Araiijo, U. Nunes, "Real-Time Labeling of Places using Support Vector Machines", *IEEE International Symposium on Industrial Electronics*, pp. 2022 - 2027, 2007

[9] T. Mitchell. (2005). Generative and Discriminative Classifiers: Naive Bayes and Logistic Regression [Online]. Available: http://www.cs.cmu.edu/~tom/mlbook/NBayesLogReg.pdf

[10] O.M. Mozos, Semantic Place Labeling [Online]. Available: http://www.informatik.uni-freiburg.de/~omartine/place_data_sets.html File: fr52.tgz

[11] A.Y. Ng, "Feature selection, L1 vs. L2 regularization, and rotational invariance", *International Conference on Machine Learning*, 2004

[12] R.E. Fan, K.W. Chang, C.J. Hsieh, X.R. Wang, and C.J. Lin. "LIBLINEAR: A library for large linear classification", *Journal of Machine Learning Research*, vol. 9, pp. 1871-1874, 2008

[13] M. ElKaissi, M. Elgamel, M. Bayoumi, B. Zavidovique, "SEDLRF: A New Door Detection System for Topological Maps", *International Workshop on Computer Architecture for Machine Perception and Sensing*, pp. 75 - 80, 2007

[14] D. Anguelov, D. Koller, E. Parker, S. Thrun, "Detecting and modeling doors with mobile robots", *IEEE International Conference on Robotics and Automation*, vol. 4, pp. 3777 - 3784, 2004

[15] W.S. Wijesoma, K.R.S. Kodagoda, A.P. Balasuriya, "Road-boundary detection and tracking using LADAR sensing", *IEEE Transactions on Robotics and Automation*, vol. 20(3), pp. 456 - 464, 2004

[16] Y. Freund, R. Schapire, N. Abe, "A Short Introduction to Boosting", *Journal of Japanese Society for Artificial Intelligence*, vol. 14(5), pp. 771-780, 1999

[17] C.J.C. Burges, "A Tutorial on Support Vector Machines for Pattern Recognition", *Data Mining and Knowledge Discovery*, vol. 2(2), pp. 1-47, 1998

[18] A. Pronobis, O. M. Mozos, B. Caputo, and P. Jensfelt, "Multi-modal semantic place classification", *International Journal of Robotics Research*, vol. 29, pp. 298-320, 2010

[19] J.O. Berger, *Statistical decision theory and Bayesian analysis*, 2nd ed., New York: Springer-Verlag, 1985, pp. 272-277

[20] J.R. Raol, *Multi-Sensor Data Fusion: Theory and Practice*, Boca Raton: CRC Press, 2010, pp. 60-61

[21] A Telmoudi, S Chakhar, "Data fusion application from evidential databases as a support for decision making", *Information and Software Technology*, vol. 46(8), pp. 547 - 555, 2004

[22] O. M. Mozos, "Semantic Place Labeling with Mobile Robots", PhD thesis, Dept. of Computer Science, University of Freiburg, 2008.

[23] R. Siegwart, I.R. Nourbakhsh, *Introduction to autonomous mobile robots*, Cambridge: MIT Press, 2004, pp. 200-212

[24] I.H. Witten, E.Frank, *Data mining : practical machine learning tools and techniques*, Boston, MA: Morgan Kaufman, 2005, pp. 163