

Environment Classification and Semantic Grid Map Building Based on Laser Range Finder Data

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

Abstract – Human robot interaction has become an important area of research in the robotics community. High level abstractions, which are commonly used by humans, can be learnt by robots to effectively communicate with humans. In this paper, we propose a Semantic Grid Map (SGM) to represent an environment. SGM is similar to an Occupancy Grid (OG) map, however with high level information as environment type labels. We use a robot-mounted laser range finder (LRF) data to learn and classify an environment into various area types. Then the classification results are combined probabilistically to update the semantic grid map. The classification accuracy is further improved by outlier rejection and topological correction. Finally we present a labeling strategy while a robot is exploring an unknown environment. Experimental results of a robot exploring in a university environment are presented to assess the performance of the algorithm.

I. INTRODUCTION

ENVIRONMENT classification enables a robotic device to build a human understandable, more meaningful representation, and it is the basis of building a Semantic Grid Map. This high level information can be shared between robots and humans, or used to support the robots to perform complex, advanced tasks.

Some popular sensors that have been used in environment classification schemes are cameras and laser range finders. For example, many prior works propose environment classification algorithms by matching lines and landmark objects extracted from vision data [1][2]. Some other researchers have been working on machine learning approaches to classify environments based on laser range finder data [3][4]. Multi-sensor based approaches for environment classification are also not uncommon [5]. However, our belief is that the potential use of the laser range finder data has not yet been well exploited. Therefore, in this paper, we have restricted our sensory module to laser range finder.

In this research area, Poncela *et al.* [6] adopted Principal Component Analysis to classify the objects as walls or doors based on laser range data. Buschka *et al.* [7] proposed a rectangular-fit algorithm to extract room-like nodes and

thereby segmented the space into room and corridor regions. Both approaches rely on the invariant width and length parameters of certain objects or space. Tapus *et al.* [8] proposed a Bayesian approach for topology recognition and door detection by considering many detailed topological situations like X-crossing, T-crossing etc. Mozos *et al.* [3] extracted variety of features from laser range data and used AdaBoost to label indoor environments as rooms, corridors, doorways and halls. In a similar approach, Sousa *et al.* [4] classified places using Support Vector Machines (SVM). In both cases, only the positions of the robot were labeled but not the environment itself.

In this paper, we firstly evaluated some common features extracted from laser range data and select four dominant features for our application using Differential Evolution algorithm [21]. Then we classify places by using logistic regression as a multi-class classifier. The classification results can be used to label either the observer's positions or the obstacle points (environmental features). Then we adopt the independent opinion pool approach to fuse the probabilities assigned to certain obstacle points detected by different observations, resulting in a semantic grid map. The labeling accuracy of the semantic grid map is improved in two ways: through outlier rejection, and through topological constraints. Finally, these techniques are combined to construct an exploration and labeling framework, which enables a robot to navigate in a relatively unknown environment and construct accurate labeled maps.

Remainder of the paper is organized as follows: Section II discusses the details about the classifier and the feature selection method. Data fusion approach and topological information based correction are discussed in the Section III. The idea of exploration and labeling is proposed in Section IV. In Section V, experimental results are presented, and Section VI concludes the paper.

II. CLASSIFICATION

A. Semantic Labels

In a typical university building, some commonly observed areas are *office rooms, lecture rooms, corridors* and *doorways*. All of them can be used as semantic labels.

However, as a doorway can be considered as a transition between other areas, it is not adopted as an independent semantic class in the application discussed in this paper.

This work is supported by the ARC Centre of Excellence programme, funded by the Australian Research Council (ARC) and the New South Wales State Government.

Lei Shi, Sarath Kodagoda, and Gamini Dissanayake are with the ARC Centre of Excellence for Autonomous Systems (CAS), The University of Technology, Sydney, Australia

B. Feature Selection

In supervised machine learning problems, feature selection issue directly affects the generalization ability, overhead and overfitting of the system. As is widely accepted that a small subset of features is sufficient to approximate some problems well [10]. Finding the dominant features becomes a key issue for most applications.

In general, office rooms, lecture rooms and corridors have different gross shapes and sizes. However due to the presence of various furniture, a laser range finder may not observe the gross shapes [20]. This causes the laser range/bearing data to appear in complex shapes and sizes. Therefore, environment classification based on few trivial features, such as the gross shape of a place will be erroneous.

There are various features that could be used for semantic classification of places. Mozos *et al.* [3][13] derived two sets of simple features from laser range data. One set was derived from raw range data and the other from polygonal approximation of the observed area. They employed about 150 single-valued features (considering different thresholds) of 21 kinds, and fed these features into a multi-class AdaBoost classifier. In a similar manner, Sousa *et al.* [4] selected 14 single-valued features to train a binary SVM classifier. In this paper, we adopt all 21 kinds of single-valued features proposed by Mozos *et al.* [3][13] as candidate features.

Instead of using the total number of available features, we utilized the feature selection algorithm proposed by Khushaba *et al.* [21]. In this algorithm, Differential Evolution was used to implement a constrained feature selection algorithm. The author proposed to build statistical feature distribution factors and employ them to guide the search toward the vicinity of the optimal solutions.

As a result, four dominant features were selected and adopted considering their performance and computational overhead.

C. Logistic Regression

Some prominent classifiers used in semantic labeling of places and other general applications are AdaBoost [11] and Support Vector Machines (SVM) [12]. The former constructs a strong classifier as a linear combination of many weak classifiers, and the latter solves the classification problem in a high-dimensional feature space and has the advantage of dealing with sparse training samples.

However, in this paper, we adopt Logistic Regression as classifier because it directly generates probabilistic output, which can be conveniently processed in data fusion.

Logistic Regression is an approach to learn functions of the form $P(y|\vec{x})$ in the case where y is discrete-valued, and \vec{x} is any vector containing discrete or continuous variables [9]. It assumes a parametric form for the distribution $P(y|\vec{x})$ while directly estimating its parameters from the training data [9].

For binary classification, given data \vec{x} and parameters or weights \vec{w} , the parametric model is:

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

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

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

However, overfitting is a potential risk of logistic regression especially when data is with high dimension and training data is sparse [9]. Therefore, regularization approaches which encourage the fitted parameters to be small are usually employed to reduce over fitting [10].

Here, we use L2 regularization, which solves the following problem.

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

where $C > 0$ is a penalty parameter.

Logistic regression is originally a binary classifier, however it has been extended to deal with multi-class problems by applying strategies such as one-against-all and one-against-one [19].

III. DATA FUSION AND CORRECTION

In this section, we describe the process of environment classification and semantic grid map building.

A. Independent Opinion Pool

In 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. Three common approaches to combine these probabilistic evidences are linear opinion pool, independent opinion pool and independent likelihood pool [14][15].

In this application, observations are made independently from unduplicated positions and they are processed to produce a probabilistic opinion on their labels. As this happens while the observer is moving, temporal data can be fused. Based on the structure of this application, independent opinion pool is an appropriate method to solve the problem.

The independent opinion pool can be described as follows:

$$P(x|z_1, \dots, z_n) = \alpha \cdot \prod_i^n P(x|z_i) \quad (4)$$

where z_i is the i^{th} observation of an obstacle point, x is the state of the obstacle point, and α is a normalizing factor.

Equation (4) can be further rephrased in a recursive way:

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

B. Accumulative Map Building

With the availability of temporal data, it is possible to update and improve the grid map based on equation (5). This results in a semantic grid map. As the classifier grossly label the whole laser range/bearing scan as belonging to a particular class, the generated semantic grid map has poor labeling results. With opened doors, the laser can “see” features in multiple area types. Therefore, if we can further discriminate the obstacle points in a single scan into different regions, the accuracy of the final semantic grid map can be improved.

C. Outlier Rejection

Here inliers are defined as the obstacle points belonging to a particular area type and outliers are belonging to another area type. For example, if a robot in an office room detects obstacle points both in the same office room and in an adjacent corridor, then the former are called inliers and the latter are called outliers. By discriminating inliers and outliers, it is possible to improve the quality of the semantic grid map.

One way to discriminate inliers and outliers is to find the doors which could be considered as the transition between two area types. There are some door detection techniques proposed in the literature based on assumption of a fixed door width [16], dynamic door states [17] and fuzzy temporal rules [18]. Here, we have combined some of the above thoughts to build a door detector using the following heuristics.

- Door is presented by two break points (door frame points), which are supported by “walls”
- The distance between the two break points should be within a certain range (a reasonable door width)
- All obstacle points seen through doors are belonging to another area type

Another simple way is to filter the laser data based on a range threshold. Although it is not always true, intuitively it could be perceived that most probably the short range data belongs to inliers. In addition, these near ranges are more reliable and accurate. Therefore, we only use obstacle points which are within a certain range to the observer to build the semantic grid map.

In our application, statistics show that only 1% of outliers lie within 2.5m of the laser ranges.

D. Correction Based on Topology

The semantic grid map could also be improved based on the topological data. For example, room to corridor transition is more common than that of room to room or corridor to corridor transitions.

In this application, we define following heuristics:

- Corridors: Enclosed by obstacles (like walls) and at least one other environment types (like office rooms)
- Office rooms / lecture rooms: Enclosed by obstacles (like walls or furniture) and at least one corridor

The algorithm initially segments the semantic grid map to generate a “segment and neighborhood table”. All segments that do not comply with above heuristics are merged iteratively with their largest neighbor (which must not be an “obstacle”).

IV. EXPLORATION AND LABELING

The complete process of exploration and labeling is shown in Fig. 1. It consists of two sub-modules, which are actual and virtual exploration stages.

In the actual exploration stage, the robot explores an unknown space based on any suitable exploration algorithm to generate an occupancy grid map (OG map). Then a virtual exploration is carried out on the OG map visiting all the possible robot locations while obtaining laser range data. Orientation of the robot is not significant here as the laser data provides a 360° field of view. These laser data and position of the robot are used to generate a labeled robot position map. The quality of the labeled robot position map is further improved by incorporating topological knowledge. Corrected labeled robot positions along with laser range finder data are probabilistically fused to generate a semantic grid map.

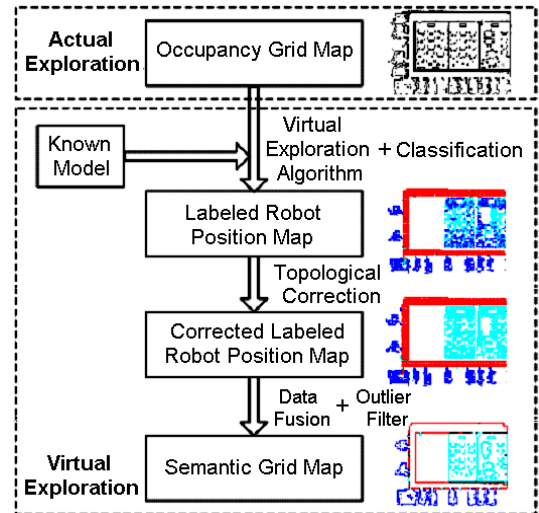


Fig. 1. Block diagram of exploration and labeling

V. RESULTS

A. Data Set

A simulated data set collected using a robot equipped with laser range finders operating in an indoor environment (Level 6, Building 2 of the University of the University of Technology, Sydney) are used for the analysis. As shown in Fig. 2, the space is consisted of 3 long corridors, 3 lecture rooms with tables and chairs (middle part in the figure) and 15 office rooms of different shapes with furniture (outer most areas in the figure). Two laser range finders were attached back to back to provide 360° laser range scans.

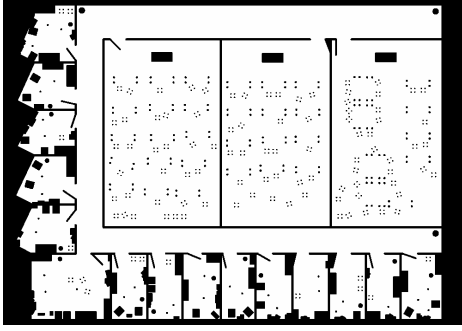


Fig. 2. Map of the environment

B. Feature Selection

All 21 single-valued features defined in [3][13] was used and Differential Evolution algorithm [21] was employed to find the optimal combinations. The evaluation result is shown in Fig. 3, and the corresponding time performance on testing and feature extraction are shown in Fig. 4 and Fig. 5. It is to be noted that the most dominant times are related to feature extraction (Fig. 5 in milliseconds) and not the classification (Fig. 4 in microseconds).

By comparing the performance of classification in terms of accuracy and computational complexity, it can be seen that four features arguably provided the best performance. Those four features were identified as:

- Standard deviation of the distances between the centroid and the shape boundary
- Ma / Mi (Ma and Mi are Major axis and Minor axis of the ellipse that approximates $P(z)$)
- Normalized circularity of $P(z)$
- Kurtosis of the laser range sequence

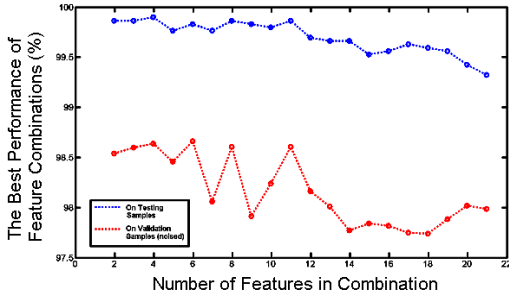


Fig. 3. Performance of best N ($N = 2, 3, \dots, 21$) features on testing data set (blue) and noised validation data set (red)

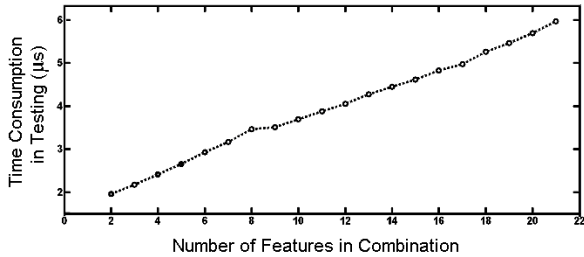


Fig. 4. Time consumption of of best N ($N = 2, 3, \dots, 21$) features in testing

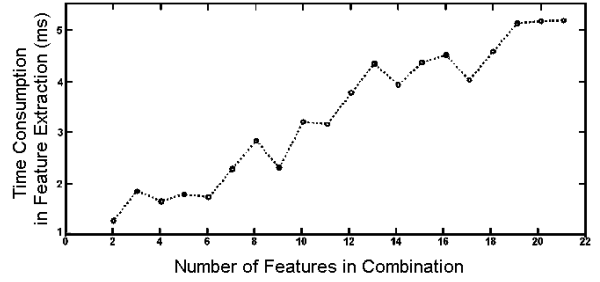


Fig. 5. Time consumption of of best N ($N = 2, 3, \dots, 21$) features in feature extraction

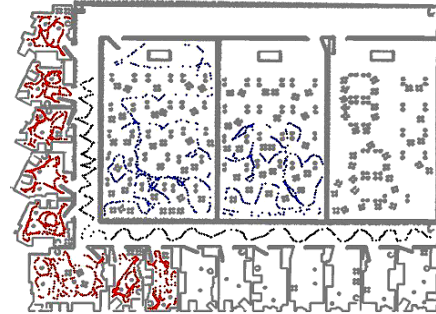
C. Classification Results

In this experiment, observer's positions are classified into three semantic labels. For this purpose, 2957 scans have been used as the training data set (as identified in Fig. 6 (a)), and another 2956 scans have been used as the testing data set.

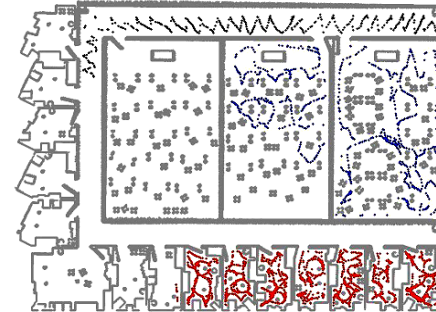
Classification is carried out using L2-regularized logistic regression as a multi-class classifier and the output is in the form of probability estimation. Performance of the classifier is shown in TABLE I, and the testing result on entire testing data set is visualized as in Fig. 6 (b).

TABLE I
PERFORMANCE OF CLASSIFIER

Item	Error
Training Error (2957 cases, mixed)	0.47 %
Testing Error (401 cases, Corridor)	< 0.01 %
Testing Error (746 cases, Lecture Room)	0.61 %
Testing Error (1809 cases, Office Room)	< 0.01 %



(a)



(b)

Fig. 6. (a) Training data set and (b) Testing data set. The grey points depict the background map as a reference. Red, black and blue points are observer's positions which are labeled as in *office room*, *corridor* and *lecture room* respectively (training data set is manually labeled and testing data set is labeled by classifier).

D. Outlier Rejection

In this section, two approaches to remove outliers mentioned in (Section III.C) are tested and compared.

The door detection approach is carried out using heuristic rules, and the results are shown in Fig. 7. In the corridor scenario (Fig. 7, (b)), the door detection is reasonable due to the lack of foreign objects like furniture. However, there are few possible candidate doors (false detections) detected in the office room scenario (Fig. 7, (a)). This is due to the presence of various furniture providing similar features to doors.

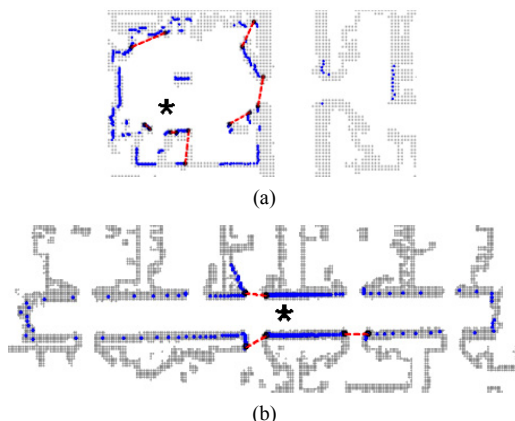


Fig. 7. Performance of door detection: The observer marked as a star is in (a) an office room and (b) a corridor. The blue points are the laser based observations. Red lines indicate the potential doors.

The second approach, which only labels near obstacle points, is not visualized here because it is relatively simple. Compared with the other approach, this method is the fastest and no additional implementation is required, even though there is a significant information loss due to discarding long range readings. In the following experiments, this approach is adopted.

E. Virtual Exploration and Labeling

This experiment implements the idea of exploration and labeling (Section IV) of an unknown environment.

For an actual implementation, any complex exploration algorithm can be used. Here we adopt right-wall following algorithm as an example, to actually explore an environment and generate an OG map (as shown in Fig. 8). Localization of the robot is assumed to be known.

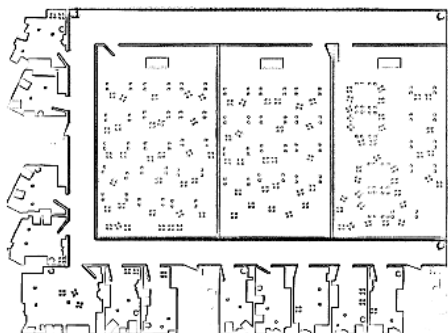


Fig. 8. Occupancy Grid Map generated by right wall following exploration

Once the OG map is generated, a virtual exploration is carried out. It is based on a cell (a different concept, not a grid in OG map) lay over the environment. The size of a cell is determined by the size of the robot.

It is assumed that the robot can move to the all four adjacent cells (front, back, left, and right) if there are no obstacle cells. Two variable-length buffers are maintained by the robot: one buffer is used to store unvisited cells, and another buffer is used to store visited empty cells. After placing the robot randomly in any empty cell, the robot moves to neighboring four cells (while obtaining laser range/bearing data) and they are marked as visited cells. The rest of the cells are maintained as unvisited cells. This process is to be continued until there are no unvisited empty cells.

During virtual exploration, the robot labels its positions and builds a labeled robot position map as shown in Fig. 9. As can be seen in the figure, it has many wrong robot position classifications (error rate 4.25%).

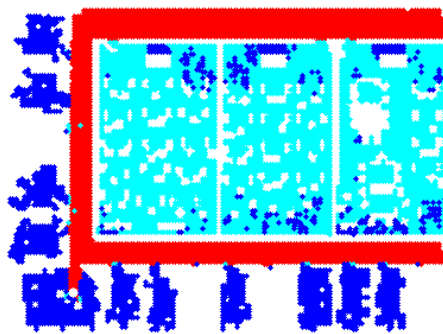


Fig. 9. Labeled robot position map that was built during virtual exploration. Blue, cyan and red points are either observer's positions labeled as belonging to office room, lecture room and corridor respectively.

This map can be further improved by topological correction. Fig. 9 is represented in a tabular form preserving the neighborhood relationship of segments (a segment means a connected grid cells with the same label). The table will be processed according to the topological heuristics given in Section III. D. It starts with the smallest segments and iteratively updates the table by merging segments with the biggest neighbors based on the heuristics. This is to be repeated until all the segments satisfy the topological heuristics.

The resulting corrected labeled robot position map is shown in Fig. 10 (a) (error rate 0.69%). Then the corrected labeled robot position map along with the respective laser range finder data can be used in independent opinion pool framework to generate a semantic grid map as shown in Fig. 10 (b) (error rate 1.25%). In this figure, it is to be noted that the obstacle points are labeled rather than the robot's positions.

Evaluation results show that by introducing the topological correction, the error rates in labeling the robot position map is reduced from 4.25% to 0.69% and error rates of semantic grid map have improved from 5.77% (if built based on uncorrected robot position map) to 1.25%. Therefore, it

shows that the topological correction approach dramatically improves the accuracy of the final semantic grid map.

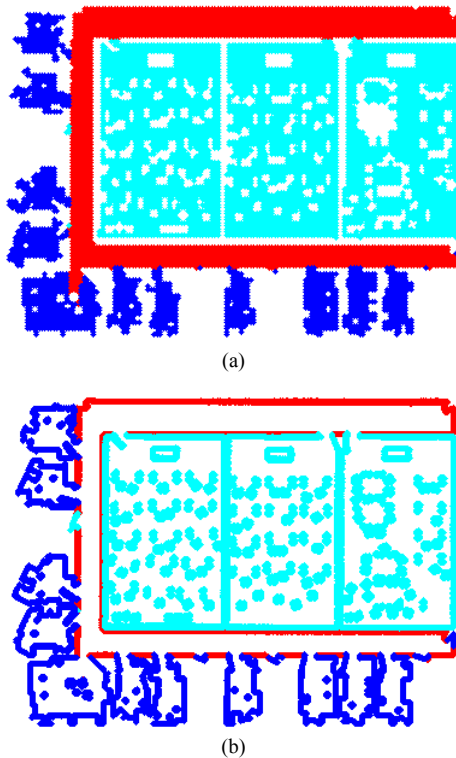


Fig. 10. (a) corrected labeled robot position map and (b) semantic grid map. Blue, cyan and red points are either observer's positions or obstacle points labeled as belonging to *office room*, *lecture room* and *corridor* respectively.

VI. CONCLUSIONS

In this paper, we have presented an approach to classify the robot positions in an environment based on laser range data. This is followed by a classification of obstacle points in the environment based on a probabilistic temporal update leading to a semantic grid map. The accuracy of semantic grid map was further improved based on outlier rejection algorithm and topological information. In addition, an idea of *virtual exploration and labeling* is proposed and implemented, which enables a robot to explore a relatively unknown environment and produce semantic grid maps. Experiments on university indoor environments with lecture rooms, office rooms and corridors show that the results are convincing with low labeling errors.

We are currently in the process of implementing the algorithms on an indoor robot platform.

ACKNOWLEDGMENT

This work is supported by the ARC Centre of Excellence programme, funded by the Australian Research Council (ARC) and the New South Wales State Government.

REFERENCES

- [1] WX. 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. 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
- [3] 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
- [4] 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
- [5] 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
- [6] 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
- [7] 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
- [8] 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
- [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] A.Y. Ng, "Feature selection, L1 vs. L2 regularization, and rotational invariance", *International Conference on Machine Learning*, 2004
- [11] 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
- [12] 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
- [13] O. M. Mozos, "Semantic Place Labeling with Mobile Robots", PhD thesis, Dept. of Computer Science, University of Freiburg, 2008.
- [14] J.O. Berger, *Statistical decision theory and Bayesian analysis*, 2nd ed., New York: Springer-Verlag, 1985. pp. 272-277
- [15] J.R. Raol, *Multi-Sensor Data Fusion: Theory and Practice*, Boca Raton: CRC Press, 2010, pp. 60-61
- [16] 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
- [17] 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
- [18] 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
- [19] 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
- [20] S.Thrun, W.Burgard, D.Fox, *Probabilistic robotics*, Cambridge, Mass: MIT Press, 2005, pp. 281
- [21] R.N. Khushaba, A. Al-Ani, A. Al-Jumaily, "Feature subset selection using differential evolution", *15th international conference on Advances in neuro-information processing*, vol. 1, 2008