The 2010 IEEE/RSJ International Conference on
Intelligent Robots and Systems
October 18-22, 2010, Taipei, Taiwan

Semantic Map Partitioning in Indoor Environments using Regional Analysis

Carlos Nieto-Granda, John G. Rogers III, Alexander J. B. Trevor, and Henrik I. Christensen

***Abstract*— Classification of spatial regions based on semantic information in an indoor environment enables robot tasks such as navigation or mobile manipulation to be spatially aware. The availability of contextual information can significantly simplify operation of a mobile platform. We present methods for automated recognition and classification of spaces into separate semantic regions and use of such information for generation of a topological map of an environment. The association of semantic labels with spatial regions is based on *Human Augmented Mapping*. The methods presented in this paper are evaluated both in simulation and on real data acquired from an office environment.**

I. INTRODUCTION

Humans are constantly trying to make their lives easier. Service robots capable of operating in human environments have the potential to improve daily life by assisting humans in a variety of tasks. Endowing these robots with the ability to understand and reason about spatial regions such as individual rooms, as well as understanding the semantic labels of such spaces could facilitate tasks such as navigation and mobile manipulation in human environments.

Human environments are typically partitioned into discrete spaces, such as offices, corridors, living rooms, etc. Such a partitioning allows humans to organize and enable their everyday activities, and these spaces typically have specific purposes and labels. Service robots that understands the partitioning of human environments can utilize this information to better assist humans in everyday tasks. For example, if a robot is given the command "fetch the red mug from the kitchen", having an understanding of the location and extent of the region considered "kitchen" is beneficial.

In this paper, we present a method for building a semantic map partition in cooperation with a human guide. Given a metric map that the robot can localize in, the system creates a semantic map partition. The resulting semantic map partition provides a probabilistic classification of the metric map into a set of labels provided by the human guide. The robot can then use this semantic map partition to navigate to a region with a specific label, and can determine the maximum likelihood label for any point in the partition. Additionally, if the robot is not confident that it knows a likely semantic label for its current pose in the map, it will prompt the guide to provide one.

C. Nieto-Granda, J. Rogers III and A. Trevor and are Ph.D. students at Georgia Tech College of Computing {carlos.nieto, atrevor, jgrogers}@gatech.edu

H. Christensen is the Kuka Chair of Robotics at Georgia Tech College of Computing hic@cc.gatech.edu

The paper is organized as follows. We briefly describe related work in Section II, followed by the motivation of our research in Section III. We then present our Gaussian probabilistic regions approach in Section IV. In Section V, we present some experiments and results both in simulation and on a real robot. Finally, conclusions and future work are given in Section VI.

II. RELATED WORK

In recent years, we have seen important developments in service and assistive robots for domestic applications and tasks. These works focus on the understanding of the environment using semantic information in order to create a synergistic interaction between humans and robots. Dellaert and Bruemmer[3] proposed extending FastSLAM to add semantic information of the environment to each particle's map. Several approaches have been presented for map partitioning, using topological and geometric representations of the environment. For example, Oberländer [11] proposed a SLAM algorithm based on FastSLAM 2.0 [9] that maps features representing regions with a semantic type, topological properties, and an approximate geometric extent. The resulting maps enable spatial reasoning on a semantic level and provide abstract information allowing efficient semantic planning and a convenient interface for human-machine interaction. Thrun [16] integrated grid-based maps to learn the environment using artificial neural networks and naïve Bayesian integration to generate a topological map by partitioning the latter into coherent regions.

Another body of work focuses on extracting semantic spatial properties of the environment from 2D and 3D data. Donsung and Nevatia [7] introduced a new spatial representation, *s-map*, for an indoor navigation robot. This map represents the locations of visible 3D-surfaces of obstacles in a 2D space. O'Callaghan [12] developed a new statistical modeling technique for building occupancy maps by providing both a continuous representation of the robot's surrounding and an associated predictive variance employing a Gaussian process and Bayesian learning. Ekvall [4] applied an automatic strategy for map partitioning based on detecting borders between rooms and narrow opening to denote doors or gateways using different types of features (lines, points, SIFT). 'Rhino' [2] is an example of a service robot which integrates localization, mapping, collision avoidance, planning, and various modules concerned with user interaction telepresence giving tours on a museum. BIRON, a mobile Home Tour Robot [15], uses integrated vision based localization a modular architecture and extending a spoken dialog system for on-line labeling and interaction about different locations in a real, fully furnished home environment where it was able to learn the names of different rooms. The approach presented by Topp and Christensen [18] and [17], provides a separation of regions that relate to a users view on the environment and detection of transitions between them. They assumed an interactive setup for the specification of regions and showed the applicability of their method in terms of distinctiveness for space segmentation and in terms of localisation purposes.

III. Semantic SLAM

As robots have to cooperative with humans it is advantageous that they have a shared representation of the space, preferably a model that is simple for the human to use as part of commanding the robot and understanding feedback. Semantic mapping literature has focused on developing robotic mapping techniques capable of functionally supporting these types of interactions. To perform these tasks, one of the strategies that is used is to portray the relationship between a place and the knowledge that is associated with it e.g.(functionality, objective location), is semantic mapping. Kuipers [8] proposed the *Spatial Semantic Hierarchy* (SSH), which is a qualitative and quantitative model of knowledge of large-scale space consisting of multiple interacting representations. This map also informs the robot of the control strategy that should be used to traverse between locations in the map. This representation is based on the relationship of objects, actions and the dependencies from the environment. More recently, Beeson *et al.* [1] provided a more specific framework representation of spatial knowledge in small scale space. This framework is focused on the robot's sensory horizon e.g.(global and local symbolic, and metrical reasoning of the space), but also human interaction.

Existing approaches for robot indoor navigation build an occupancy grid map using range data from its sensors. These maps, however, only provide geometric information such as obstacles and open areas in the environment without a semantic understanding of it. Martínez-Mozos and Rottmann [10] [14] introduce a semantic understanding of the environment creating a conceptual representation referring to functional properties of typical indoor environments. Providing semantic information enables a mobile robot to more efficiently accomplish a variety of tasks such as human-robot interaction, path-planning, and localization. Ekvall [4] integrated an augmented SLAM map with information based on object recognition, providing a richer representation of the environment in a service robot scenario.

In this work, we focus on providing a semantic partition of a metric map using semantic labels provided by a human. We believe this representation could be used to support semantic reasoning for a variety of mobile robot tasks in indoor environments. As an example, we present navigation to the nearest point in a metric map that has a specific semantic label.

IV. Approach

Our goal was to design a system capable of reasoning about spaces. In contrast to work such as [10], which builds a topological map on top of a metric map, we instead provide a continuous classification of the metric map into semantically labeled regions.

The semantic map layer of our system is a multivariate probability distribution on the coordinates of our metric map to a set of semantic labels. This multivariate distribution is modeled as a Gaussian model. Each of the Gaussians in the model is based on the robot's sensor data when it was provided a label by a human guide. Each spatial region is represented using one or more Gaussians in our metric map's coordinate frame. So, a region with label L and n Gaussians, each with mean μ and covariance Σ, is represented as:

$$Region = \{L, \{\{\mu_1, \Sigma_1\}, \{\mu_2, \Sigma_2\}, ..., \{\mu_n, \Sigma_n\}\}\}$$

A semantic map is then just a collection of such regions, so a semantic map with m regions would be represented as:

$$Map = \{region_1, region_2, ..., region_m\}$$

Our system builds these maps partitions of our metric maps through human guidance. The human takes the robot on a tour of the space (either by driving the robot manually, or using a person following behavior), and teaches the robot typing the appropriate label for the space that it is currently in. The regional analysis technique is to take a laser scan measurement, fit a Gaussian to the resulting points, and store the mean and covariance in the map along with the label provided by the human.

Using this semantic map partition, the robot can be asked for its belief of the name of the region it is currently occupying. This is done by evaluating the Mahalanobis distance of the robot's current pose x close by labels coded as Gaussian region models (Equation 1), and choosing the region that is closest using this metric.

$$D_M(x) = \sqrt{(x-\mu)^T \Sigma^{-1} (x-\mu)} \qquad (1)$$

This map representation allows for probabilistic classification of the map by region label. Additionally, while navigating through the environment, the robot continuously checks its position with respect to the semantic map partition. If it is not sufficiently confident (more than a certain threshold) that it is in a region with a known label, it prompts the user to input the name of the current region.

Once the robot has a semantic map partition, users can request that the robot navigate to one of the regions, such as "living room". The robot can then find the region in the map with label "living room", and calculate the Mahalanobis distance from its current position to the mean of each Gaussian in the region. The robot selects the closest of these as the goal, and sends this to its path planner in order to autonomously navigate to that region. While traveling, the robot continuously calculates its confidence of which region it is, and stops when it is confident that it is more likely to be in the goal region than any other region as follows:

$$\textit{Distance to goal} < \frac{1}{4} * \textit{Nearest distance to the non-goal}$$

This results in the robot entering a region, but not attempting to move to the region's center. Additionally, if the robot enters the region with the desired label at any point while navigating to its goal point, perhaps because the path to the closest point was blocked, it will recognize this and stop once it is sufficiently in the goal region.

In this work, we only use the semantic map partition for navigation tasks; however, we believe this map representation has a number of other applications, such as searching for objects. For example, if we have a mobile manipulation platform and ask the robot to "get the mug from the kitchen", our map representation can be used to give a spatial region of our metric map which should be search, by finding the area that is labeled as kitchen with at least a certain confidence level.

V. EXPERIMENTS

Our approach has been implemented and evaluated in several experiments. We designed two simulated environments in Stage[5][1] in which the robot can be taught locations and navigate between them. Preliminary experiments were also performed using our Segway RMP-200 mobile platform to verify our technique on a real robot platform.

A. Simulation Environment

In this simulated experiment, we tested the effectiveness of our method by providing labels for each of the rooms and hallways.

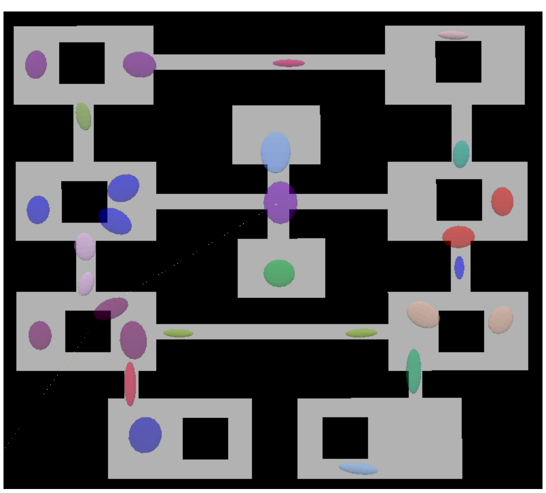

Fig. 1. A visualization of the Gaussian regions representing the rooms and hallways. All the Gaussians are colored in different colors for identification. This figure is best viewed in color.

The first experiment, shown in Fig. 1, consisted of labeling each room and each hallway and navigating between them in order to test the system's effectiveness at accepting this type of navigation command.

The robot is able to successfully navigate with the calculated trajectory, avoiding obstacles. One of the tests was to move from a room labeled "room 9" to "room 4" in the map as appears in Fig. 1.

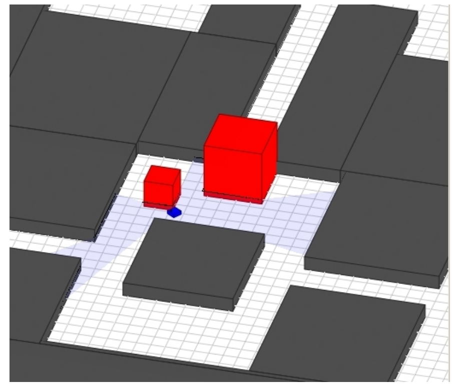

Fig. 2. The robot's path is obstructed with a simulated block. The robot is the smaller blue object and the obstacles are the large red cubes.

Shown in Fig. 2, the shortest path was obstructed with a simulated block and the robot replanned a new trajectory to reach the goal as can be seen in Fig 3. The laser hits on the obstacle that blocks the shortest path can be seen near the robot in Fig. 3.

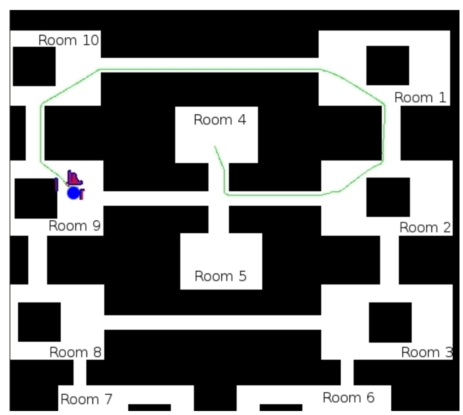

Fig. 3. Robot path replanned to navigate from room '9' to central hallway and arrived in room '4'. The blue circle represents the robot, the green line the robot path and in red the obstacles detected by the laser scan.

For the second experiment, we labelled twenty seven rooms and a hallway in the map, and left one unknown area unlabeled, as can be seen in the upper right corner of Fig. 4. The main purpose was to simulate an office environment, where the transition region between rooms is a hallway. The robot is able to continuously provide the current location and navigate from one region to another.

Based on this map classification, we created several different scenarios to test our system. One of the test scenarios, shown in Fig 5, involved the robot navigating between two regions with different labels, a room and a hallway,

[1] Stage is a 2D multiple-robot simulator from the Player project. http://playerstage.sourceforge.net

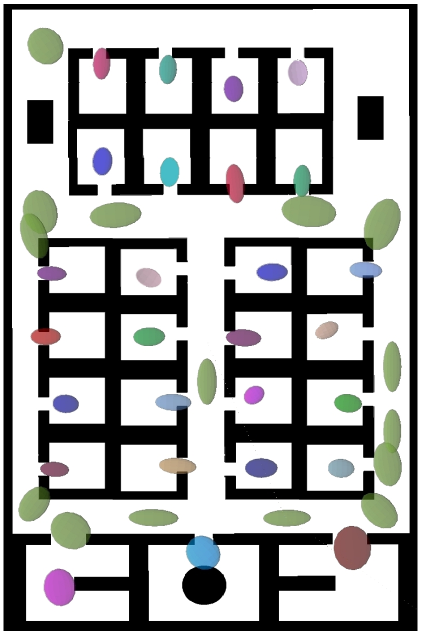

Fig. 4. The second simulated office environment used for our experiments. Colored ellipses represent the Gaussians in our model, and different colors represent different spaces. This figure is best viewed in color.

represented by Gaussians with means very close to each other. The robot was requested to navigate from the room to the hallway and back.

Our system calculated the Mahalanobis distance to find the closest Gaussian region, but the regions are very close to each other that the robot only turn and move only a short distance. This demonstrates that our system will cause the robot to move into a region until it has a certain confidence level that it is in the region, rather than stopping on the equiprobable decision boundary between the regions as shown in Fig 6.

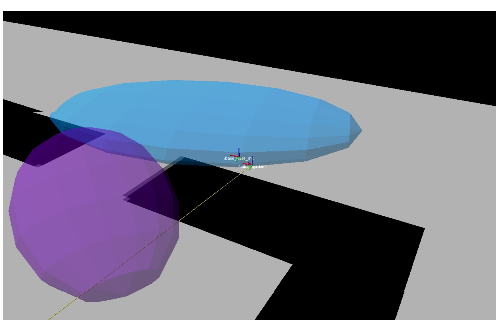

Fig. 5. Visualization of the robot navigating between to regions in the environment.

Another scenario, shown in Fig. 7 is to drive the robot through a hallway to an unknown area which has never been assigned any label. When the robot reached a location that was not likely to be part of any previously labeled region, it displayed *"I do not know where I am."* and requested that the user provide the current location's label displaying *"Please tell me where I am"*.

Several more scenarios consisted of teaching the robot with different orientations and locations inside of the rooms.

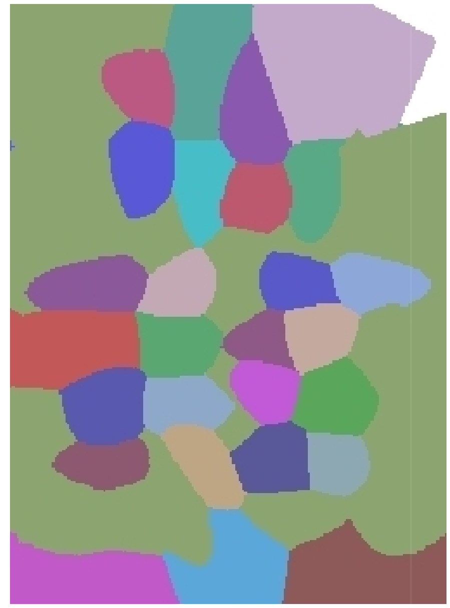

Fig. 6. A visualization of the decision boundaries of the regions representing the rooms and hallways on the map shown in 4. All the Gaussians are colored in different colors for identification. This figure is best viewed in color.

Localizing the robot between rooms and hallways worked better when the user taught new locations to the robot in the middle of the room as opposed to in the doorways because the laser hits were more representative of the room's extent.

Finally, we tested the robot by starting it with no semantic map information, so that it would prompt the user immediately for a label, which the human guide provided. The robot then navigated around the environment, and would stop whenever it was not confident that it knew the label for the current pose. Upon being provided with a label for the current location by the human guide, the tour resumed until it again needed a new label. This resulting map is shown in Fig. 8. This experiment demonstrates our method's effectiveness at determining when a new label is required. The robot began in the left middle room, and ended in the upper right.

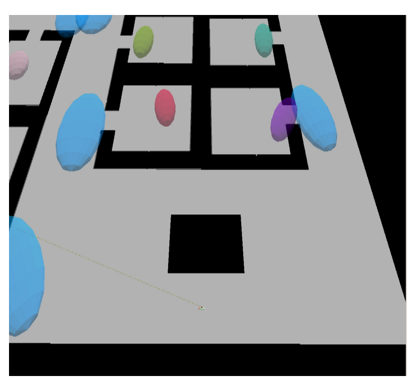

Fig. 7. The robot is driven in an unknown area which has not been assigned any label.

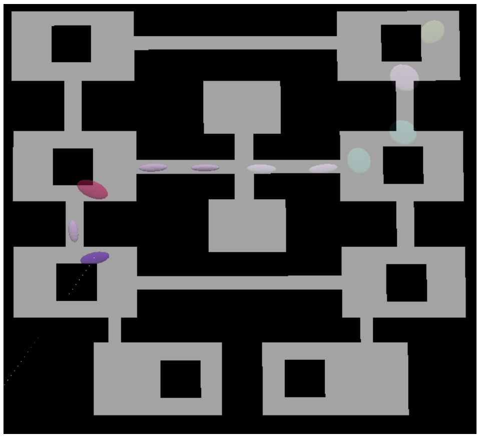

Fig. 8. Result from starting the robot with no known locations in the semantic map, and prompting the user when the robot was not confident of the appropriate label.

B. Real Environment

Our approach has been implemented on a Segway RMP-200 mobile platform (Fig 9). It is equipped with a SICK LMS-291 laser scanner, which is used for localization, mapping and obstacle avoidance, and is controlled by an on-board Mac mini computer (2.26GHz/Core 2 Duo). We conducted an experiment in our real office environment (Fig. 10), teaching the robot with 5 different rooms and several different points in the hall.

Fig. 9. Our robot platform used in our experiments.

The first step was to drive our robot through the environment while collecting laser data and odometry. The SLAM gmapping module included with ROS[13], which is based on the Rao-Blackwellized particle filter technique by Grisetti *et. al* [6], was then used to build a metric grid map of the environment. The map shown in Fig. 11 was then used for localization in our experiments.

Beginning with the map, the human tour started from the hallway to one of the three cubicles. When the human guide stopped, the robot was provided with the name of

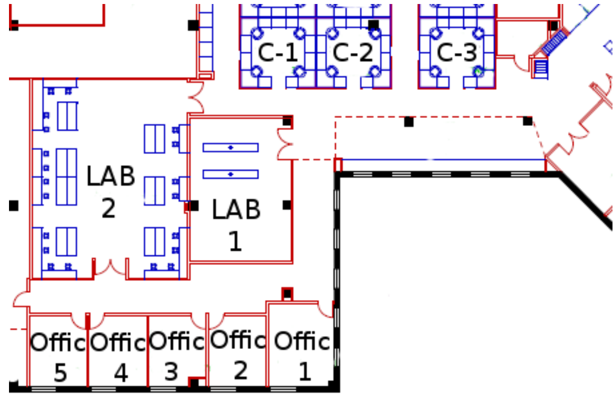

Fig. 10. Robotics and Intelligent Machines Laboratory Map.

the location. During the tour, the robot could be queried for its current location, then the robot would calculate the Mahalanobis distance to the Gaussians regions and report the label of the nearest one. If the robot pose was not near to a previously labelled location, the robot report the location as "unknown". Also, the robot can be asked to move to a specific known location, for example move from "*C-3*" to "*LAB 1*". Then, the robot calculates a safe trajectory to the room using the global map and continuously runs a local planner to avoid obstacles throughout the environment. When the robot successfully completed the task, it reported that it arrived at the current location's label.

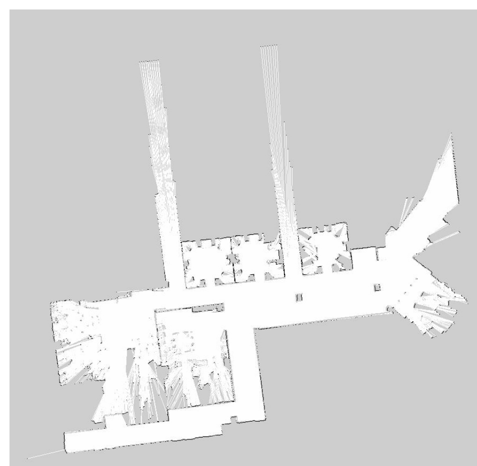

Fig. 11. Generated Occupancy Grid Map for localization used in our experiments.

VI. Conclusion and Future Work

We presented a technique for partitioning metric maps into labeled spatial regions using a Gaussian model. We then used this representation to perform navigation tasks in the map, as demonstrated by our preliminary experiments both in simulation and on a real robot platform.

In future work, we plan to investigate techniques for automatically labeling regions. One example of how to do this could be by reading signs that are present in office environments. Objects detected by the robot could also provide information as to an appropriate label for a space, for example, if a microwave and a toaster are detected, then "kitchen" might be an appropriate label.

Also, we would like to investigate additional uses for this map representation such as constraining the search space when processing a request such as "fetch the red mug from the kitchen".

VII. ACKNOWLEDGMENTS

This work was made possible through the KORUS project, the Boeing corporation, and the ARL MAST CTA project 104953.

REFERENCES

[1] Patrick Beeson, Matt MacMahon, Joseph Modayil, Aniket Murarka, Benjamin Kuipers, and Brian Stankiewicz. Integrating multiple representations of spatial knowledge for mapping, navigation, and communication. In *Symposium on Interaction Challenges for Intelligent Assistants*, AAAI Spring Symposium Series, Stanford, CA, March 2007.

[2] Wolfram Burgard, Armin B. Cremers, Dieter Fox, Dirk Hähnel, Gerhard Lakemeyer, Dirk Schulz, Walter Steiner, and Sebastian Thrun. Experiences with an interactive museum tour-guide robot. In *Artificial Intelligence*, volume 114, pages 1–2, 1998.

[3] Frank Dellaert and David Bruemmer. Semantic slam for collaborative cognitive workspaces. In *In AAAI Fall Symposium Series*, 2004.

[4] Staffan Ekvall, Danica Kragic, and Patric Jensfelt. Object detection and mapping for service robot tasks. *Robotica: International Journal of Information, Education and Research in Robotics and Artificial Intelligence*, 25(2):175–187, March/April 2007.

[5] B. Gerkey, R. Vaughan, and A. Howard. The player/stage project: Tools for multi-robot and distributed sensor systems. In *11th International Conference on Advanced Robotics (ICAR 2003)*, Coimbra, Portugal, June 2003.

[6] Giorgio Grisetti, Cyrill Stachniss, and Wolfram Burgard. Improved techniques for grid mapping with rao-blackwellized particle filters. In *IEEE Transactions on Robotics*, volume 23, 2007.

[7] Dongsung Kim and Ramakant Nevatia. A method for recognition and localization of generic objects for indoor navigation. *Image and Vision Computing*, 16:729–743, 1994.

[8] Benjamin Kuipers, Rob Browning, Bill Gribble, Mike Hewett, and Emilio Remolina. The spatial semantic hierarchy. In *Artificial Intelligence*, volume 119, pages 191–233, 2000.

[9] Michael Montemerlo, Sebastian Thrun, Daphne Koller, and Ben Wegbreit. Fastslam 2.0: An improved particle filtering algorithm for simultaneous localization and mapping that provably converges. In *In Proc. of the Int. Conf. on Artificial Intelligence (IJCAI*, pages 1151–1156, 2003.

[10] Oscar Martinez Mozos, Rudolph Triebel, Patric Jensfelt, Axel Rottmann, and Wolfram Burgard. Supervised semantic labeling of places using information extracted from sensor data. *Robotics and Autonomous Systems (RAS)*, 55(5):391–402, May 2007.

[11] Jan Oberländer, Klaus Uhl, Johann Marius Zöllner, and Rüdiger Dillmann. A region-based slam algorithm capturing metric, topological, and semantic properties. In *In Proc. of the IEEE Intl. Conf. on Robotics and Automation (ICRA)*, pages 1886–1891, 2008.

[12] Simon O'Callaghan, Fabio T. Ramos, and Hugh F. Durrant-Whyte. Contextual occupancy maps using gaussian processes. In *In Proc. of the IEEE Intl. Conf. on Robotics and Automation (ICRA)*, pages 1054–1060. IEEE, 2009.

[13] Morgan Quigley, Ken Conley, Brian Gerkey, Josh Faust, Tully B. Foote, Jeremy Leibs, Rob Wheeler, and Andrew Y. Ng. ROS: an open-source robot operating system. In *International Conference on Robotics and Automation*, Open-Source Software workshop, 2009.

[14] Axel Rottmann, Óscar Martínez Mozos, Cyrill Stachniss, and Wolfram Burgard. Semantic place classification of indoor environments with mobile robots using boosting. In *AAAI'05: Proceedings of the 20th national conference on Artificial intelligence*, pages 1306–1311. AAAI Press, 2005.

[15] Thorsten Spexard, Shuyin Li, Britta Wrede, Jannik Fritsch, Gerhard Sagerer, Olaf Booij, Zoran Zivkovic, Bas Terwijn, and Ben Kröse. Biron, where are you? - enabling a robot to learn new places in a real home environment by integrating spoken dialog and visual localization. In *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*. IEEE, October 2006.

[16] S. Thrun. Learning metric-topological maps for indoor mobile robot navigation. In *Artificial Intelligence*, volume 99, pages 21–71, 1998.

[17] E. A. Topp and H. I. Christensen. Detecting region transitions for human-augmented mapping. *Robotics, IEEE Transactions on*, pages 1–5, 2010.

[18] Elin A. Topp and Henrik I. Christensen. Topological modelling for human augmented mapping. In *Intelligent Robots and Systems, 2006 IEEE/RSJ International Conference on*, pages 2257–2263, Oct. 2006.