Object guided autonomous exploration for mobile robots in indoor environments

Carlos Nieto-Granda‡, Siddharth Choudhary‡, John G. Rogers III *, Jeff Twigg *, Varun Murali ‡ and Henrik I. Christensen‡

‡ Institute for Robotics & Intelligent Machines, Georgia Institute of Technology, Atlanta GA 30332, USA

*U.S. Army Research Laboratory (ARL), Adelphi, MD 20783, USA

ABSTRACT

Autonomous mobile robotic teams are increasingly used in exploration of indoor environments. Accurate modeling of the world around the robot and describing the interaction of the robot with the world greatly increases the ability of the robot to act autonomously. This paper demonstrates the ability of autonomous robotic teams to find objects of interest. A novel feature of our approach is the object discovery and the use of it to augment the mapping and navigation process. The generated map can then be decomposed into into semantic regions while also considering the distance and line of sight to anchor points. The advantage of this approach is that the robot can return a dense map of the region around an object of interest. The robustness of this approach is demonstrated in indoor environments with multiple platforms with the objective of discovering objects of interest.

1. MOTIVATION

Robots have shown a lot of potential in exploration, mapping and navigation of indoor environments. Robots are ideal for applications such as discovery of Improvised Explosive Devices (IEDs). Robotic teams can achieve this task faster, provide dense maps of regions around an object of interest and a model of the discovered object. The model of the discovered object can then be viewed by a human user and take the necessary actions. Semantic mapping and object discovery have gained interest independently. The intersection of these areas provides an interesting insight into using robots to discover objects of interest in indoor environments.

Given the intended application, an object guided approach to indoor navigation is ideal. Robots can use knowledge of the environment and the semantic labeling to guide their ability to autonomously navigate and discover objects of interest. To achieve this, we propose an algorithm that combines the knowledge of discovered objects into the mapping process. This can provide an insight into the environment and aid the process of modeling the discovered objects.

2. RELATED WORK

Semantic mapping has gathered a lot of interest from the robotics community. Benjamin Kuipers modeled the environment as a spatial semantic hierarchy where each level of the hierarchy expresses states of partial knowledge corresponding to different level of representations.[1] Ranganathan and Dellaert presented a 3D generative model for representing places using objects.[2] The object models are learnt in a supervised manner. Nüchter and Hertzberg described an approach to semantic mapping by creating a 3D point cloud map of the environment and labeling points using the different semantic categories like floor, wall, ceiling or door.[3] SLAM++ system proposed by Salas-Moreno et al. is the most closely related to our work.[4] They train some domain specific object detectors corresponding to repeating objects like chairs and tables. The learnt detectors are integrated inside the SLAM framework to recognize and track those objects resulting in semantic map. Similarly Kim et al. uses learnt object models to reconstruct dense 3D models from single scan of the indoor scene.[5] In contrast to these approaches, we do not train object detectors before hand. We discover objects and train on object representation in an online manner.

Further author information:
Send correspondence to Carlos Nieto-Granda E-mail: carlos.nieto@gmail.com

Unmanned Systems Technology XVI, edited by Robert E. Karlsen, Douglas W. Gage, Charles M. Shoemaker, Grant R. Gerhart, Proc. of SPIE Vol. 9084, 90840M · © 2014 SPIE
CCC code: 0277-786X/14/$18 · doi: 10.1117/12.2050818

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 06/24/2014 Terms of Use: http://spiedl.org/terms

Similarly in the area of object discovery, Karpathy et al. decomposes a scene into candidate segments and ranks them according to their objectness properties.[6] Collet et al. used domain knowledge in the form of metadata and use it as constraints to generate object candidates.[7] All of these approaches detect objects in a batch mode after the robot has scanned the environment and generated a 3D model or a video sequence. We discover objects in an incremental fashion where it gets discovered simultaneously as the robot moves around in an environment.

Rogers et al. recognize door signs and text describing the door to enable the robot to map complex environment and produce semantic map.[8] Another work by Trevor et al. used planar surfaces as landmarks in the mapping system.[9] Pronobis et al.[10] proposed a complete and efficient representation of indoor spaces including semantic information. They use a multi-layered semantic mapping representation to combine information about the existence of objects in the environment with knowledge about the topology and semantic properties of space such as room size, shape and general appearance

In,[11] the authors prove performance characteristics on a multi-robot collaboration strategy to perform adversary search. By representing the topological configuration of a map as a graph, the robots can guarantee that the adversarial search will prevent re-contamination of previously cleared nodes with an arbitrary sized team. Joyeux et al.[12] describe a distributed system for managing robot plans for performing high- level tasks. This architecture prevents conflicts between robot plans and can handle communication failures. These papers both present strategies and architectures for collaboration between robot agents to perform tasks.

3. METHODOLOGY

3.1 Object Discovery, Modeling and Mapping

Each RGBD frame is segmented using connected component segmentation to generate planes and non-planar segments.[13] We consider all the non-planar segments as the object segments. Each non planar segment from each frame is propagated across frames to generate object hypothesis. The proposed object hypothesis **O** are represented using 3D feature descriptors augmented with map information. Object representation helps in identifying loop closure when the robot sees the same object and for producing better object models when an object part undersegmented in a few frames matches to the full object model. Object recognition is done by finding the object representation with the maximum number of inlier correspondences with the current object. If an object does not find the minimum number of inlier correspondences with any of the saved representations, it saves the representation of the current object. This follows an online learning framework, where the robot identifies potential objects and matches to other objects in the map, if none of them matches it hypothesizes that the potential object is a new object and therefore saves the representation. Trevor et al. provide a detailed description of the mapping framework for integrating different sensor plugins.[14] Figure 1 shows the snapshot of the objects discovered using our pipeline.

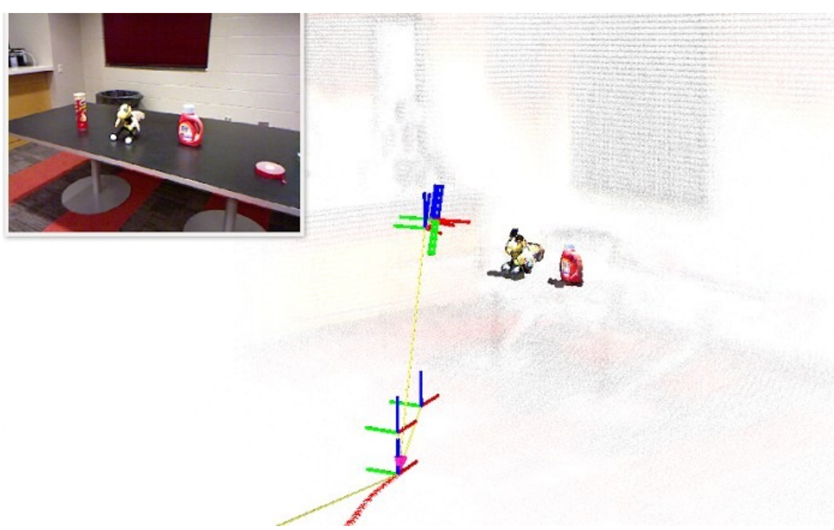


Figure 1. Snapshot of the objects discovered using our pipeline.

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 06/24/2014 Terms of Use: http://spiedl.org/terms

3.2 Region Decomposition

The region decomposition stage breaks the map into areas which are determined by distance and line of sight (LOS) to a specific point in a region. This point is called the region anchor. In indoor environments, we hypothesize that distance and line of sight are good heuristics for delineating regions.

Region decomposition is a well studied problem in robotics and often reduces to a variant of the Voronoi diagram computation. In,[15] the authors propose an algorithm to update a Voronoi decomposition based on input from a dynamic occupancy grid representation of the environment. In,[16] Voronoi information is used to extract higher-level semantic information about the environment in order to identify features such as hallways, doors, and rooms. This strategy also operatver a kmeans with a (LOS) awareness. Our approach continues in this vein and adapts Boris Lau's code to serve a prior for the clustering algorithm. This prior is used to sort and determine potential anchor points for creating the regions. As a result the decomposition is different from a voronoi based decomposition in that smaller rooms and regions can expand into hallways. Larger rooms are not form continuous regions and hallways are also divided up by distance. In a voronoi based decomposition, the region would be bounded by the intersection with the hallway. However, this decomposition provides for situations where the portion of the hallway might be more strongly correlated to the entrance of that room than the a section of hallway which is farther away.

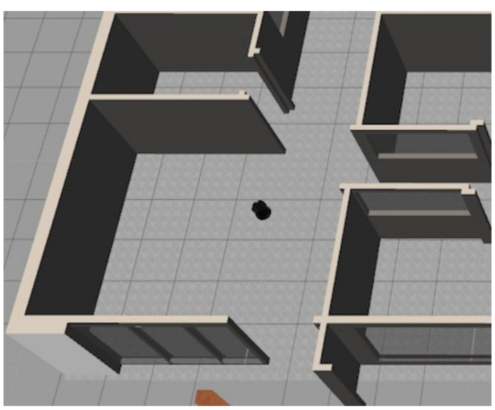
In the decomposition algorithm, A possible regions are considered given a Map M. The cells of the map that are considered part of the voronoi decomposition in Boris Lau's code are the possible anchor points A. The algorithm uses the initial position within the map to perform a brushfire computation on A. As a result, points which are not burned are dropped from A. This forces the decomposition to be over a continuous space. Afterwards A is sorted by distance from each point in A to its the nearest obstacle.

4. EVALUATION

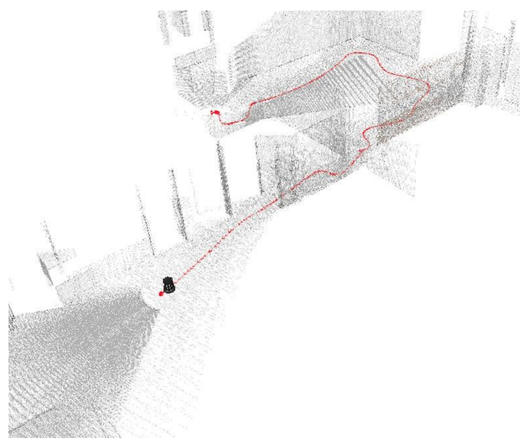
For the evaluation of the methods considered in the previous section we performed a series of experiments in simulation and real environments. In this paper we are focus on indoor and outdoor environments.

4.1 Simulation Environment

In these experiments, our testing scenarios are office environments. The simulations are run in the multi-robot simulator Gazebo *. The environments have hallways, rooms and some objects that there robot need to find meanwhile is navigating and exploring a floor. An example of a running simulation is shown in figure 2.



(a) Our mobile platform exploring a 3D office environment.



(b) Visualization of the map acquired by the robot in the simulation environment.

Figure 2. Data acquired by the robot on the user's interface.

*iGazebo multi-robot simulation currently develop at the Open Source Robotics Foundationhttp://gazebosim.org

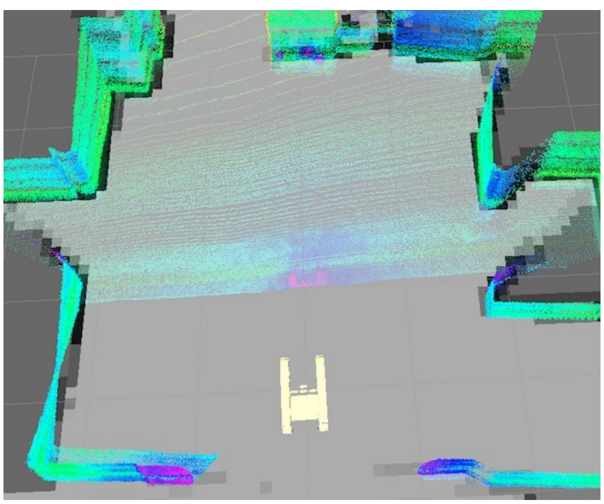
Downloaded From: http://proceedings.spiedigitallibrary.org/ on 06/24/2014 Terms of Use: http://spiedl.org/terms

4.2 Real Environments

In these experiments, our navigation system was tested in a Military Operations on Urbanized Terrain (MOUT) site. The mobile platform is a Packbot equipped with a payload (intel core i7). Onboard a Hokuyo UTMM-30LX in a dynamixel PTU with an ASUS camera and an IMU device.

Figure 3. A mobile robot (packbot) during an autonomous exploration found an IED.

Our navigation system was tested across different environments. The goal of each experiment was to explore an area and look for interesting objects such as an improvised explosive device (IED) as shown in figure 4. When the human viewer is interested on an object, the robot will acquire a very dense 3D point cloud data where the object of interest was found (figure 5).

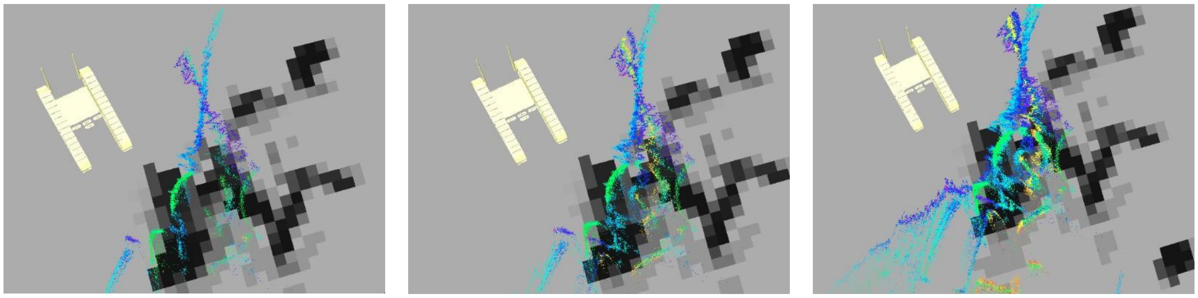


(a) 3D map built by the mobile robot.

(b) Image from the camera on board of the mobile robot.

Figure 4. Data acquired by the robot on the user's interface.

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 06/24/2014 Terms of Use: http://spiedl.org/terms



(a) Normal resolution of the 3D Point cloud building a map. (b) Acquisition of a higher resolution as the normal 3D point cloud. (c) Increasing the resolution of the 3D point cloud.

Figure 5. Sequence of how the 3D dense point cloud is acquired by the robot.

The robot is able to acquire a dense 3D point cloud which will be use to identify the size and type of an IED when an image is not a good reference for object recognition. The monocular camera will fail in very bright light conditions, when there is grass on the image, etc.

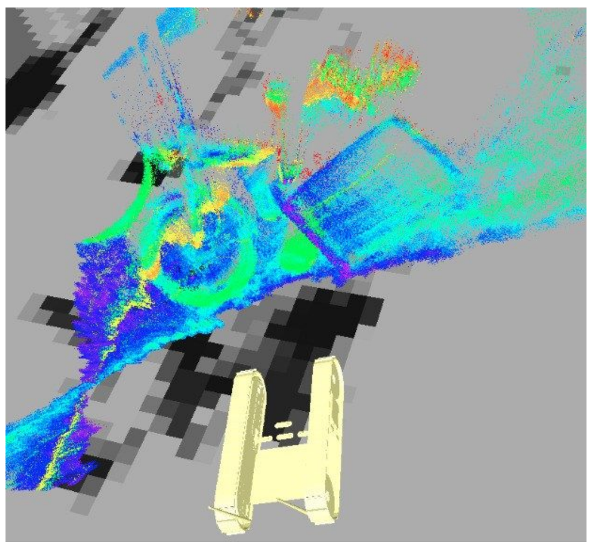


Figure 6. 3D dense point cloud acquisition to identify IEDs on the environment.

5. CONCLUSIONS AND FUTURE WORK

In this paper we have demonstrated a navigation and exploration system based on object guidance. This system has been tested in different urban environments to guarantee the efficiency and performance for future operations. Also, the system runs in different robot mobile platforms.

We would like to perform experiments using a multi-robot heterogeneous teams (UAV and UMV) . The main goal will be to identify objects of interest and share the information between each member of the teams. The tasks of each team will be determined by the sensors and capabilities of each robot. In addition, we want to investigate an optimal exploration adaptation using a persistent pattern technique. Therefore, the robot will be able to learn and take advantage of it's on-board sensors. At the end, the robot network will provide support to soldiers and marines on the field.

ACKNOWLEDGMENTS

The research was performed as part of the Army Research Laboratory - Collaborative Technology Alliance on Micro Autonomous Systems Technology (MAST) . The support is gratefully acknowledged.

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 06/24/2014 Terms of Use: http://spiedl.org/terms

REFERENCES

[1] Kuipers, B., "The spatial semantic hierarchy," *Artificial Intelligence* **119**, 191 – 233 (2000).

[2] Ranganathan, A. and Dellaert, F., "Semantic modeling of places using objects," in [*Robotics: Science and Systems*], (2007).

[3] Nüchter, A. and Hertzberg, J., "Towards semantic maps for mobile robots," *Robot. Auton. Syst.* **56**, 915–926 (Nov. 2008).

[4] Salas-Moreno, R. F., Newcombe, R. A., Strasdat, H., Kelly, P. H., and Davison, A. J., "SLAM++: Simultaneous localisation and mapping at the level of objects," in [*The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*], (June 2013).

[5] Kim, Y. M., Mitra, N. J., Yan, D.-M., and Guibas, L. J., "Acquiring 3d indoor environments with variability and repetition," *ACM Trans. Graph.* **31**(6), 138 (2012).

[6] Karpathy, A., Miller, S., and Fei-Fei, L., "Object discovery in 3d scenes via shape analysis," in [*ICRA*], IEEE, ed. (May 2013).

[7] Collet Romea, A., Xiong, B., Gurau, C., Hebert, M., and Srinivasa, S., "Exploiting domain knowledge for object discovery," in [*IEEE International Conference on Robotics and Automation (ICRA)*], IEEE, ed. (May 2013).

[8] Rogers, J. G., Trevor, A. J. B., Nieto-Granda, C., and Christensen, H. I., "Simultaneous localization and mapping with learned object recognition and semantic data association," in [*IROS*], 1264–1270 (2011).

[9] Trevor, A., Rogers, J., and Christensen, H., "Planar surface SLAM with 3D and 2D sensors," in [*Robotics and Automation (ICRA), 2012 IEEE International Conference on*], 3041–3048 (2012).

[10] Pronobis, A. and Jensfelt, P., "Large-scale semantic mapping and reasoning with heterogeneous modalities," in [*Robotics and Automation (ICRA), 2012 IEEE International Conference on*], 3515 –3522 (may 2012).

[11] Hollinger, G., Singh, S., and Kehagias, A., "Improving the efficiency of clearing with multi-agent teams," *The International Journal of Robotics Research* **29**(8), 1088–1105 (2010).

[12] Joyeux, S., Alami, R., Lacroix, S., and Philippsen, R., "A plan manager for multi-robot systems," *International Journal of Robotics Research* **28**(2), 220–240 (2009).

[13] Trevor, A. J. B., Gedikli, S., Rusu, R. B., and Christensen, H. I., "Efficient organized point cloud segmentation with connected components," in [*Semantic Perception Mapping and Exploration (SPME)*], (May 2013).

[14] Trevor, A., Rogers, J., and Christensen, H., "Omnimapper: A modular multimodal mapping framework," in [*Robotics and Automation (ICRA), 2014 IEEE International Conference on*], (2014).

[15] Lau, B., Sprunk, C., and Burgard, W., "Improved updating of euclidean distance maps and voronoi diagrams," in [*Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*], 281 –286 (oct. 2010).

[16] Owens, J. and Fields, M., "Incremental region segmentation for hybrid map generation," in [*Army Science Conference*], 281 –286 (2010).

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 06/24/2014 Terms of Use: http://spiedl.org/terms