Literature Review of SLAM and DATMO

Ardhisha Pancham[1, 2], Dr Nkgatho Tlale[1], Prof Glen Bright[1, 2]
1) CSIR, Pretoria, South Africa
2) UKZN, Durban, South Africa

Abstract—**Simultaneous Localization And Mapping (SLAM) allows a mobile robot to be completely autonomous in an unknown environment and perform its tasks. The robot is able to create a map of its environment and at the same time locate itself. Real world environments however, are characterized by moving objects such as people, cars, robots and mobile furniture.**

In order for the robot to interact safely with these moving objects the robot would have to perform Detection And Tracking Of Moving Objects (DATMO). Moving object detection and tracking would eliminate errors in maps, resulting in a reliable map that would enable the robot to localize itself in the environment and execute its tasks. This paper provides a literature review of the techniques and sensors employed to allow a mobile robot to perform SLAM and DATMO.

Keywords-SLAM; DATMO; dynamic; detecting and tracking; Kinect

I. Introduction

The SLAM problem has been researched extensively in static environments. Applications have evolved from different environments such as indoor to outdoor, aerial, underwater and mining. Most of these applications are undertaken in static environments that do not account for dynamic objects.

Dynamic objects fall into two different categories, those that are always moving and those that are temporarily static and change their position over time.

Temporarily stationary objects can be incorrectly accounted for as static and lead to data association errors that reduce map accuracy. SLAM In Dynamic Environments (SLAMIDE), has been researched recently, but open questions still exist. These questions are:

- How to distinguish between static and dynamic objects,
- How to represent static and dynamic objects, and
- How to track dynamic objects and predict their position over time. [7, 13, 14, 17]

The paper is organized as follows: Section II introduces SLAM and DATMO, lists tracking and data association methods for DATMO, and explains the popular occupancy grip-based SLAM technique, and a free space strategy to identify moving objects. Section III consists of the literature review. Section IV compares the different techniques. Section V concludes the paper, and Section VI describes the intended application.

II. SLAM and DATMO

A. *SLAM and DATMO processes*

SLAM and DATMO provide a basis for the development of driverless cars such as those involved in the DARPA challenge [21, 22]. Autonomous cars can assist the physically disabled, reduce mundane and long transportation trips, and also prevent accidents that occur due to human errors such as speeding or distraction. [15]

SLAM and DATMO involve:

- Localization of the robot,
- Mapping of the environment,
- Detection of moving objects and
- Tracking of moving objects.

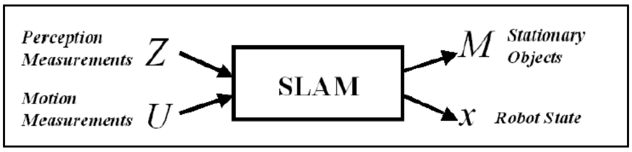
SLAM assumes the robot environment as static, i.e. having only non-moving objects. [14] Dynamic objects are regarded as noise sources. In some scenarios, this hypothesis is acceptable, but in most real world environments where dynamic objects cannot be avoided, these approaches succumb to errors reducing the overall map quality.

SLAM can be referred to as a process concerned with the state of the robot in a static environment. As illustrated in Figure 1.a the inputs are observations from exteroceptive sensors (e.g. laser scanners and cameras) and proprioceptive sensors (e.g. odometry and IMU). The outputs of SLAM are the location of the robot and a map of the static objects.

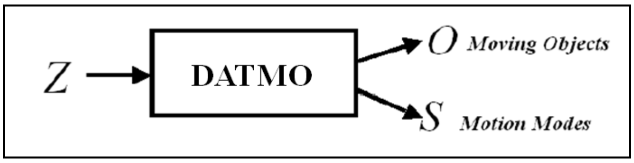
DATMO can be referred to as a process concerned with the states of objects that the robot can 'see' or perceives in a dynamic environment. If an accurate pose estimate is available, the inputs of DATMO are observations from exteroceptive sensors. The outputs are positions of dynamic objects and their respective trajectories (refer to Figure 1.b). Unlike SLAM, DATMO does not have proprioceptive data of the moving objects.

As illustrated in Figure 1.c SLAM and DATMO combined, can be regarded as a process in a dynamic environment where the inputs are similar to the SLAM process, but the outputs include the results of both processes, i.e. the map of the environment, the robot pose, the positions

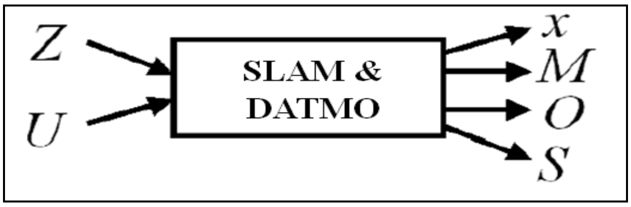
Presented at the 4th Robotics and Mechatronics Conference of South Africa (ROBMECH 2011)
23-25 November 2011, CSIR Pretoria South Africa.



a) The SLAM process



b) The DATMO process



c) The SLAM and DATMO process

Figure 1. a) The SLAM process, b) the DATMO process and c) the SLAM and DATMO process. Where *Z represents* exteroceptive measurements, *U*, motion measurements, *x*,the true robot state, *M*, static object locations, *O*, moving objects states and *S*, motion modes of the moving objects. Diagram adapted from [16].

and trajectories of dynamic objects. SLAM and DATMO require algorithms that reduce the computational complexity of both processes to produce optimal results. [14, 16, 17]

Wang [15] defined requirements for the DATMO algorithm. The algorithm must:

- Detect and initiate new dynamic objects;
- Model dynamic object trajectories;
- Perform data association;
- Combine two or more dynamic objects that correlate with each other;
- Omit dynamic objects that are no longer in the sensor's field of view;
- Account for objects that may be occluded;
- Attempt to rectify incorrect measurements;
- Operate robustly over long observation sequences.

B. Filtering and data association

After moving objects are detected they need to be tracked. The tracking of multiple moving objects in dynamic environments involves filtering and data association. Filtering is concerned with tracking one particular object and processing the data over time to compute a state estimate for the single object. Data association deals with tracking multiple objects and determines which data correlates with which object. Filtering methods then compute object state estimates with the respective data.

The most popular data association techniques are the Global Nearest Neighborhood (GNN) combined with filtering, Joint Probabilistic Data Association Filter (JPDAF) and the Multiple Hypothesis Tracking (MHT). Tracking techniques frequently used include Kalman Filters (KF), Particle Filters (PF) and Interacting Multiple Models (IMM). [2, 12, 13]

C. Grid-based SLAM

The occupancy grid based technique is described as it is applied in the majority of SLAMIDE applications for map representation and object differentiation presented in section III. It is an efficient method for representing uncertainty, combining many sensor measurements, and explicitly modeling free space for navigation purposes. The latter characteristic makes it ideal for SLAMIDE.

The occupancy grid technique was first introduced by Elfes [4]. It is regularly used as it is versatile in environment representation. It can be applied outdoors where feature identification and extraction are difficult to perform due to sensor noise.

In grid-based mapping, the occupancy of each grid cell is estimated when new sensor measurements are obtained and updated by filtering methods.

Bayesian filtering, through the use of Extended Kalman Filters (EKF), is the most common grid-based SLAM method applied in section III. [12, 18, 20]

D. Grid based Moving Object Detection

A motion-based method to differentiate between moving and stationary objects in occupancy grids is explained. The main concept is based on the occupation of space in a local grid map i.e. if the space is free (containing no objects), occupied (containing objects initially assumed static) and unknown (where it is not certain whether the objects are static or dynamic).

The reasoning follows that if an object is observed in a previously defined free space, then it is considered to be a dynamic object. If an object is detected in a previously defined occupied space then it is regarded as static. If an object appears in a previously defined unknown space, then its mobility state remains undefined.

In addition, any object found in a space where several objects are moving, should be classified as a potential moving object. Stationary objects are modeled in the local static grid map (S) and moving objects are modeled in a local dynamic grid map (D). Both maps have equal pose, size and resolution. Each cell in the dynamic map contains a value relating to the number of times a moving object has been detected in the cell.

In this way the static and dynamic parts are ascertained and errors that deteriorate localization and map quality are prevented. [12, 18, 20]

III. LITERATURE REVIEW

In the proceeding work, the following aspects have been discussed: The SLAM technique employed, the map representation, the data association and tracking methods to achieve SLAM and DATMO. The sensors utilized, environment and objects detected have also been named.

Wang [16] was the first author to integrate SLAM and DATMO. A Bayesian formula was introduced to solve SLAM and DATMO. EKFs were utilized and separate posteriors for the static and dynamic objects were maintained. A scan matching technique and the Iterative Closest Point (ICP) algorithm were used to represent data in grid-maps. Data association was achieved with the Multiple Hypothesis Tracking (MHT) algorithm. Moving objects were modeled and tracked using the IMM algorithm.

Tests were performed on the Navlab11 vehicle at high speeds in crowed urban environments. The Navlab11 was equipped with movement sensors (IMU, GPS, differential odometry, compass, inclinometer, and angular gyroscope) and perception sensors (video sensors, a light-stripe rangefinder, one SICK LMS221 and two SICK LMS291 laser range finders.

The lasers were mounted in different positions on the Navlab11 vehicle, to perform horizontal or vertical profiling and produce 3D (2.5D) maps. Pedestrians, cars, bikes and buses were detected and tracked. [16]

Hähnel, Schulz and Burgard [5] presented a probabilistic technique for forming maps in environments containing multiple persons. Sample-based JPDAFs (SJPDAFs) were implemented to track people from the laser range scans obtained. Moving people and stationary objects were identified in occupancy probability grids obtained from consecutive scans. The robust method produced better pose estimates, and reduced spurious objects in maps.

Several experiments were conducted on different robotic platforms in various environments for producing 2D and 3D maps. The Pioneer 2 robot was tested in an empty exhibition hall of the Byzantine Museum in Athens, Greece with fifteen people walking around. RWI B21 robot Rhino was tested in a large corridor environment of the Computer Science Department in Bonn with five people walking around. Both robots had a 2D laser range finder to obtain data indoors and create 2D maps.

The Pioneer 2 AT platform was equipped with two laser range finders, one to track people and the other to obtain the 3D structure of the environment. The latter laser was positioned on an AMTEC wrist module. Experiments occurred outdoors in the university campus with several people walking around. [5]

Montesano, Minguez and Montano [8] integrated the identification of moving and non-moving objects within the estimation process. A maximum likelihood incremental approach was utilized to estimate a map, robot pose and dynamic objects. An extended Iterative Dual Correspondance (IDC) algorithm jointly estimated the robot pose and differentiated the static and dynamic objects. The static parts were represented in a probabilistic grid map. Independent EKFs performed tracking of the dynamic objects. The data association was executed using the nearest neighbor rule.

Tactical planning and obstacle avoidance tests were conducted on a robotic wheelchair with a 2D laser range finder and odometric sensors, in a laboratory with moving people and doors, to construct a 2D map.

Following the work of Wang [16], Vu, Burlet, and Aycard [13] produced a reliable vehicle sensing system with an affordable 2D laser range scanner. A fast laser-based incremental localization method to correct robot positions from odometry was introduced. Laser measurements were integrated to construct a reliable grid map based on the occupancy grid framework formulated by Elfes [4]. Dynamic objects were detected by their inconsistencies with the existing grid map. Data association and tracking were achieved with MHT, combined with an adaptive IMM filter. Radar data was merged with laser data to validate the laser data results.

The DaimlerChrysler demonstrator car was equipped with a camera, two short-range radars and a 2D laser scanner. The vehicle was tested on city streets, country roads and highways at high speeds. Pedestrians and cars were detected and tracked. [13]

Vu [12] implemented an occupancy grid-based approach for SLAM with DATMO. SLAM was solved locally by maximum likelihood of the occupancy grid maps, and globally by EKF with feature-based maps, where each local grid map was depicted by a feature.

Vehicle positions from odometry where corrected with a fast incremental scan matching method similar to [13]. When adequate vehicle position was achieved, the map was updated incrementally. Dynamic objects were sensed without priori target knowledge based on free and occupied space.

The algorithm was tested on a Mercedes-Benz E-Class as part of the PreCrash collision avoidance safety application. The vehicle was equipped with Ibeo’s ALASCA laser scanner and two M/ACOM SRS100 24 GHz short range radars. Radar data was fused with laser data for more reliable results as in [13].

Dynamic objects were detected by comparing new scan measurements with the existing local grid map. Greedy Nearest Neighbor (GNN) and KF were applied for data association and tracking. The vehicle detected possible collisions in traffic on highways, rural roads and in urban areas.

In addition, Vu [12] introduced a technique for simultaneous detection, classification and tracking of dynamic objects. A model-based method was used to decipher consecutive laser scans within a sliding time window by dynamic object path hypotheses. The Data-Driven Markov Chain Monte Carlo (DDMCMC) method was applied for data association and tracking moving objects. The algorithm was simulated with Wang's Navlab dataset [16] in urban traffic. Models assisted in detecting and tracking buses, cars, bikes and pedestrians. The DDMCMC was not tested for PreCrash as it was not available at the time. [12]

Sola [10] introduced a probabilistic and geometric technique known as BiCamSLAM, where a MonoSLAM algorithm was applied to a stereo camera system. Sola divided the SLAM and tracking algorithms. Feature detection and matching methods were combined to generate a map and camera pose. A different EKF was used for each detected dynamic object, and its path approximated.

A robot with a stereo head and odometry was tested in the robotics laboratory at LAAS (Laboratory for Analysis and Architecture of Systems). Movable objects in the lab included a table, a bin, a small box, a trunk, a fence and, a white-board (which served as the target at the far end of laboratory). A 3D map was built from the sensor data. [10] Table 1 summarizes the techniques reviewed in section III.

IV. DISCUSSION

Similarities and differences of the techniques in section III are discussed in this section.

Wang [16], Hähnel, Schulz and Burgard [5] did not account for robot motion uncertainties when objects were detected. If several observations were identified inaccurately, problems may occur when there are high odometry errors, decreasing the precision and the convergence of the algorithms [9]. Vu, Burlet, Aycard [13] and Vu [12], implemented a fast laser-based incremental localization method to amend robot positions from odometry.

Montesano, Minguez and Montano [8] incorporated identification in the estimation process, whereas [5, 12, 13, 16] separated the problem, such that identification was carried out before the estimation process. According to Montesano [8], the method allowed for better object classification and increased robustness of the algorithm.

Unlike the methods in [5, 8, 12, 13, 16] that performed SLAM and DATMO with laser range finder(s), Sola [10] presented a technique that utilized cameras. Both laser range finders and cameras have their advantages and disadvantages.

Laser range finders provide reliable and accurate distance measurements but cannot detect certain materials. Transparent materials such as glass cannot be detected because the laser beam passes through the material. Black objects may not be detected because the light from the laser is absorbed. Objects with surfaces that do not diffuse sufficient light, may reflect the laser beam out so that it is not returned to the laser range finder. [9, 16]

Montesano [8] mentioned that some static objects were classified as dynamic due to misclassifications caused by the laser range finder. This occurred generally when the laser beam was parallel to the surface it was reflected off or when the beams overlooked an object because it was the same height as the laser scan. The system later determined the objects were static and classified them as such. [8]

In the method introduced by Vu, Burlet, and Aycard [13] there were more tracks than the number of real objects detected. This was caused by objects which moved across or near the laser range boundary. In this region the laser decreased in precision and the efficiency with which objects were detected diminished. Also, if an object left the laser range and later returned, it was regarded as a new object by the tracker. The tracker was able to handle a high number of non-detections and false alarms.

Wang [16] and Vu [12] proposed heterogeneous sensor fusion of laser and camera data to overcome the disadvantages of using laser range finders alone and improve performance.

Cameras produce information of a higher quality for data association or object classification. They are generally smaller, less expensive and have lower power consumption. However, signal-to-noise ratio of gained images decreases under inferior lighting and poor weather, making the use of this sensor alone, inadequate in outdoor environments. [9, 12]

The self-calibration solution proposed by Sola [10] for the BiCamSLAM system experienced reduced observability and inconsistency complications. Drifts in the yaw angle led to drifts in the map and vice-versa. Improvement was needed to enable constant, real-time calibration to allow more robustness. The method however, demonstrated that 3-D observability can be attained from uncomplicated analysis on the image plane for mapping. [10]

Methods [5, 8, 13, 16] are model-free approaches i.e. they have the ability to detect any type of dynamic object with no a prior knowledge about the object. Vu [12] utilized a model-based approach to eliminate the disadvantages of object segmentation when using laser scanners for tracking. Models

TABLE 1. TABLE OF RESEARCHED SLAM AND DATMO TECHNIQUES

Summary of SLAM and DATMO methods									
Author-Paper	**SLAM Method**	**DATMO Method**		**Navigation sensors**	**Moving objects**	**Test environment**	**Experimental platform**	**2D/3D Map**	**Contribution**
		Data Association	**Tracking**						
Wang 2002-2004	Grid-based (Bayesian State Estimation:EKF)	MHT	IMM	Laser range finders, odometry (main sensors)	People, cars, bikes, buses	Outdoor (crowded urban areas.)	Navlab11	3D (2.5D)	Pioneer of SLAM & DATMO
Haehnel 2003	Grid-based (Bayes filter)	SBJPDA		Pioneer 2: 2D SICK laser range finder RW1 B21 Rhino: 2D SICK laser range finder	People	Indoor	Pioneer 2 RW1 B21 Rhino:	2D	Obtained better pose estimates and reduced spurious objects with a more robust algorithm
				Tilting 3D laser, 2D laser		Outdoor	Pioneer 2 AT platform	3D	
Montesano 2005	Grid map	NNR	EKF	SICK 2D laser rangefinder, odometry	People, doors	Indoor (office)	Robotic wheelchair	2D	Differentiated static & dynamic objects in estimation problem
Sola 2007 (PhD Thesis)	BiCamSLAM (through Monoslam algorithm)	EKF		2 cameras	People, Movable objects(boxes, table, bin, fence)	Indoor	Robot	3D	BiCamSLAM (through Monoslam algorithm)
Vu 2009 (Thesis)	Occupancy grid framework(Elfes 1989) (ML-SLAM & EKF)	DDMCMC		2D laser range finder	People, cars, bikes, buses	Outdoor	Simulation from Navlab dataset of Wang (2004)	2D	A new fast laser-based incremental localization method and DDMCMC
		GNN	KF	2D laser range finder, odometry 2 short range radars	Objects that may lead to collisions in traffic		Mercedes Benz E class		A new fast laser-based incremental localization method
Burlet, Vu, Aycard 2010	Occupancy grid framework(Elfes 1989) (Bayes) ICP (localization)	MHT	IMM	2D laser range finder, odometry, 2 short range radars, camera (visualization only)	People, cars	Outdoor	Diamler	2D	A new fast laser-based incremental localization method

were used to detect and classify dynamic objects. According to Vu [12] this resulted in more accurate tracking results. A model-based approach, however, is limited only to the set of predefined models and cannot identify objects that are not included in the model set. [12]

The methods presented by Wang [16], Vu [12], Vu, Burlet, and Aycard [13] may be utilized for outdoor applications while the methods described by Montesano [8], and Sola [10], may be used indoors. Hähnel, Schulz and Burgard's method [5] was demonstrated in both indoor and outdoor environments.

V. CONCLUSION

In order to prevent the error of referencing dynamic objects in the localization process, static and dynamic objects need to be differentiated between. Identification of moving and stationary objects will improve localization and reduce spurious measurements in maps. The robust tracking of moving objects will guarantee reliable maps and improve map accuracy.

This paper reviewed some of the techniques implemented to perform SLAM and DATMO. SLAM and DATMO were explained. Filtering and data association methods were defined. The well-known grid-based SLAM approach was described. SLAM and DATMO techniques were researched in terms of map representation, data association, tracking, and sensor and environment implementation. Differences in techniques were compared.

VI. FUTURE WORK

The literature review provides a foundation for future work concerning SLAM and DATMO. A.P intends to research the algorithms, filtering and data association techniques in more depth, and develop algorithms for the implementation of SLAM and DATMO using multiple Kinect sensors. (Microsoft Kinect sensors are inexpensive structured 3D light sensors that generate rich point cloud data.) The SLAM and DATMO algorithms will be tested, optimized and validated on a mobile robot in a dynamic environment. [19]

ACKNOWLEDGMENT

A.P wishes to thank her supervisors for their guidance and support and the CSIR for granting her the studentship opportunity.

REFERENCES

[1] T. Bailey and H. Durrant-Whyte. Simultaneous localization and mapping (SLAM): Part II. Robotics & Automation Magazine, IEEE 13 (3). pp. 108-117. 2006.

[2] Y. Bar-Shalom and T. Fortman. Recursive tracking algorithms. Academic Press, New York, 1988.

[3] H. Durrant-Whyte and T. Bailey. Simultaneous localization and mapping (SLAM): Part I the essential algorithms. Robotics and Automation Magazine 13 (2). pp. 99-110. 2006.

[4] A. Elfes. Using occupancy grids for mobile robot perception and navigation. Computer 22 (6). pp. 46-57. 1989.

[5] D. Hahnel, D. Schulz and W. Burgard. Mobile robot mapping in populated environments. Adv. Rob. 17 (7). pp. 579-597. 2003.

[6] M. Hebert, C. Thorpe, C. C. Wang, S. Thrun and H. Durrant-Whyte. Simultaneous localization, mapping and moving object tracking. The International Journal of Robotics Research. 26 (9). pp 889-936. 2007.

[7] G. Lidoris, D. Wollherr and M. Buss. Bayesian state estimation and behavior selection for autonomous robotic exploration in dynamic environments. Presented at IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). 2008.

[8] L. Montesano, J. Minguez and L. Montano. Modeling the static and the dynamic parts of the environment to improve sensor-based navigation. Presented at Proceedings of IEEE International Conference on Robotics and Automation (ICRA). 2005.

[9] L. Montesano. Detection and tracking of moving objects from a mobile platform. Application to navigation and multi-robot localization. PhD Thesis. University of Zaragoza. 2006.

[10] J. S. Ortega. Towards visual localization, mapping and moving objects tracking by a mobile robot: A geometric and probabilistic approach. PhD Thesis. National Polytechnic Institute of Toulouse. 2007.

[11] B. Thrun. Fox. Probabilistic robotics. The MIT Press, Cambridge, 2005.

[12] T. D. Vu. Vehicle perception: Localization, mapping with detection, classification and tracking of moving objects. PhD Thesis. Grenoble Institute of Technology. 2009.

[13] T. D. Vu, J. Burlet and O. Aycard. Grid-based localization and local mapping with moving object detection and tracking. Information Fusion 12, pp. 58-69. 2011.

[14] C. C. Wang and C. Thorpe. Simultaneous localization and mapping with detection and tracking of moving objects. Presented at Proceedings of 2002 IEEE International Conference on Robotics and Automation (ICRA). 3. pp. 2918-2924, 2002.

[15] C. C. Wang, C. Thorpe and S. Thrun. Online simultaneous localization and mapping with detection and tracking of moving objects: Theory and results from a ground vehicle in crowded urban areas. Presented at Proceedings of 2003 IEEE International Conference on Robotics and Automation (ICRA). 1. pp. 842-849. 2003.

[16] C.C. Wang, Simultaneous Localization, Mapping And Moving Object Tracking. PhD Thesis. Robotics Institute. Carnegie Mellon University. 2004.

[17] D. Wolf and G. S. Sukhatme. Online simultaneous localization and mapping in dynamic environments. Presented at Proceedings of 2004 IEEE International Conference on Robotics and Automation (ICRA). 2. pp. 1301-1307, 2004.

[18] D. F. Wolf and G. S. Sukhatme. Mobile robot simultaneous localization and mapping in dynamic environments. Autonomous Robots 19 (1). pp. 53-65. 2005.

[19] M. Wolfram. An integral mobile robot platform for research and experiments in the field of intelligent autonomous systems, MSc. Thesis, Graz University of Technology. 2011.

[20] M. Wu and J. Y. Sun. Simultaneous Localization, Mapping and Detection of Moving Objects with Mobile Robot in Dynamic Environments. Presented at 2010 2nd International Conference on Computer Engineering and Technology. IEEE. 1. pp 696 -701. 2010.

[21] M. Montemerlo, J. Becker, S. Bhat, H. Dahlkamp, D. Dolgov, S. Ettinger, D. Haehnel, T. Hilden, G. Hoffmann, B. Huhnke, D. Johnston, S. Klumpp, D. Langer, A. Levandowski, J. Levinson, J. Marcil, D. Orenstein, J. Paefgen, I. Penny, A. Petrovskaya, M. Pflueger, G. Stanek, D. Stavens, A. Vogt, S. Thrun, Junior: The Stanford entry in the Urban Challenge, Field and Service Robot Journal. 2008.

[22] C. Urmson, J. Anhalt, D. Bagnell, C. Baker, R. Bittner, M.N. Clark, J. Dolan, D. Duggins, T. Galatali, C. Geyer, M. Gittleman, S. Harbaugh, M. Hebert, T. Howard, S. Kolski, A. Kelly, M. Likhachev, M. McNaughton, N. Miller, K. Peterson, B. Pilnick, R. Rajkumar, P. Rybski, B. Salesky, Y. Seo, S. Singh, J. Snider, A. Stentz, W. Whittaker, Z. Wolkowicki, J. Ziglar, H. Bae, T. Brown, D. Demitrish, B. Litkouhi, J. Nickolaou, V. Sadekar, W. Zhang, J. Struble, M. Taylor, M. Darms, D. Ferguson, Autonomous driving in urban environments: Boss and the urban challenge, Journal of Field Robotics 25 (8). pp. 425–466. 2008.