Data Paper

The International Journal of
Robotics Research
29(13) 1602–1607

DOI: 10.1177/0278364910384638
ijr.sagepub.com

The Marulan Data Sets: Multi-sensor Perception in a Natural Environment with Challenging Conditions

Thierry Peynot[1], Steve Scheding[1] and Sami Terho[2]

Abstract

In this paper we present large, accurately calibrated and time-synchronized data sets, gathered outdoors in controlled and variable environmental conditions, using an unmanned ground vehicle (UGV), equipped with a wide variety of sensors. These include four 2D laser scanners, a radar scanner, a color camera and an infrared camera. It provides a full description of the system used for data collection and the types of environments and conditions in which these data sets have been gathered, which include the presence of airborne dust, smoke and rain.

Keywords

Range sensing, computer vision, sensor data fusion, field robots

1. Introduction

In this paper we present a singular effort that has been made to constitute multi-sensor data sets to evaluate and compare perception algorithms for unmanned ground vehicles (UGVs), in particular in challenging environmental conditions. This data gathering was realized in a rural environment at the University of Sydney's test facility near Marulan, NSW, Australia.

This article is organized as follows. The first section is a short introduction to discuss the motivation for this work and the experimental design that was adopted. Section 2 presents the platform used for the data gathering, describing in particular its sensors and the calibration processes involved. Section 3 presents the collected data sets, illustrating in particular the environment types and the corresponding environmental conditions. Finally, Section 4 draws some conclusions, including suggestions for exploitation of these data.

1.1. Motivation

Public data sets are extremely useful to evaluate the performances of algorithms and to compare the results obtained by related work based on the same reference data. Some notable examples are:

- the Radish repository (Howard and Roy 2003–2008), featuring numerous logs of odometry, laser and sonar data, as well as maps, acquired mainly in indoor environments;
- the Victoria Park data set (Nebot 2000), previously collected and published by the Australian Centre for Field Robotics (ACFR), which has been extensively used since 2001 to evaluate the performance of simultaneous localization and mapping (SLAM) algorithms;
- the MIT Darpa Urban Challenge public data (Leonard et al. 2007–2008), containing the logs of the MIT vehicle, including camera images and Velodyne 3D Lidar point clouds, in an urban environment;
- the New College Vision and Laser Data Set (Smith et al. 2009), composed of color images and laser data collected in an urban environment (a campus and a park);
- the Rawseeds database (Matteucci et al. 2009), proposing data sets acquired mainly with cameras and laser range finders, outdoors in a campus and indoors, with ground truth.

[1] ARC Centre of Excellence for Autonomous Systems, Australian Centre for Field Robotics (ACFR), The University of Sydney, Australia
[2] Department of Automation and Systems Technology, Aalto University, Aalto, Finland

Corresponding author:
Thierry Peynot
ARC Centre of Excellence for Autonomous Systems, Australian Centre for Field Robotics (ACFR), The University of Sydney, Australia
Email:tpeynot@acfr.usyd.edu.au

However, there are too few examples of publicly available data sets gathered with a wide variety of sensors, in particular in outdoor natural environments, due to the time and financial cost involved in their acquisition.

Bringing such data sets to the public not only provides a common reference to numerous researchers to evaluate their algorithms, it also provides real experimental data to research teams which may not have the necessary equipment at hand.

The main differences between the Marulan data sets presented in this paper and the aforementioned data sets are twofold. First, although sensors such as the stereovision bench are not included, more sensing modalities are proposed than in most public data sets, in particular with the presence of a radar and an infrared camera in addition to the commonly used lasers and visual cameras. Second, the data were acquired outdoors in natural and semi-urban environments and numerous data sets specifically contain *challenging environmental conditions*, while most data sets only propose *nominal* conditions in which the performance of perception is reasonably well understood.

1.2. Experimental Design

The data sets described in this paper have been collected especially for the general purpose of testing various perception algorithms (e.g. obstacle avoidance or terrain interpretation) for UGVs, with no very specific algorithm in mind, to limit the *experimenter's bias* (i.e. the bias towards results expected by the human experimenter, typically the algorithm developer, who has expert knowledge of the technique under evaluation). In particular, conditions that are known to be problematic for the perception of UGVs were not avoided. On the contrary, they were specifically included in this work (see Section 3.3), since they represent some of the most significant challenges for future work on perception. A common example of such challenging conditions is the presence of airborne dust. Indeed, it typically causes many state-of-the-art perception systems to fail, as noted in the CMU PerceptOR program outcomes (Kelly et al. 2006) and the DARPA Urban Challenge (Urmson et al. 2008). In that respect, this work can also be seen as a first step into promoting integrity in perception systems (Peynot et al. 2009b).

2. System Description

The vehicle used to collect the data is the eight-wheel skid-steering Argo platform (see Figure 1), that has been retrofitted with sensors and actuators to make it an UGV. Typical operating speeds of the Argo in the context of this work are 1–2 m s^{-1}. Its angular velocity is controlled through the brake pressures on both sides. This section describes the sensors mounted on this vehicle, the calibration of these sensors, and the time-synchronization of the collected data.

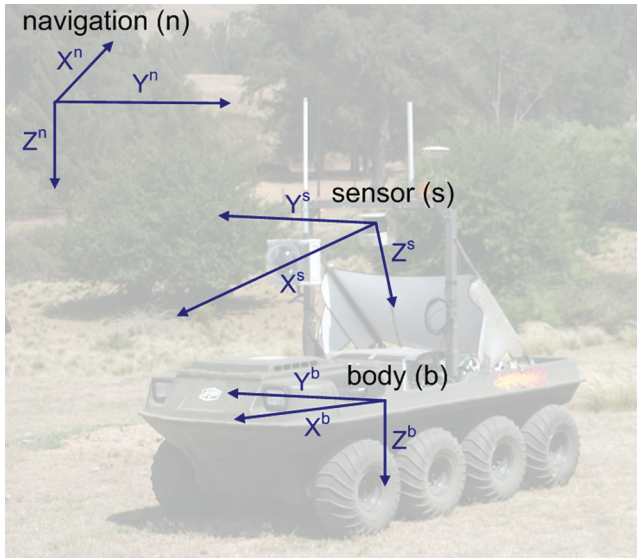

Fig. 1. The Argo and its *Sensor*, *Body* and *Navigation* frames.

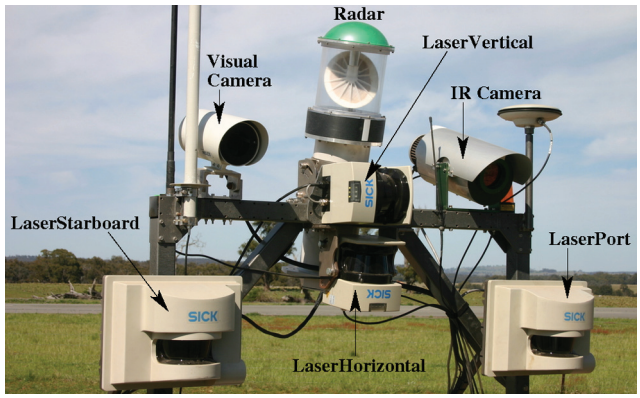

Fig. 2. Sensor frame on the Argo.

2.1. Sensors

The following exteroceptive sensors were mounted on a common frame of the vehicle (see Figure 2):

- Four 2D Sick laser range scanners, which specifications are given in Table 1. Referring to the names shown in Figure 2, the configurations of these lasers were as follows:
  - *LaserHorizontal* was horizontally centered on the sensor frame, slightly pointing down to the ground (a few degrees of pitch), with zero roll;
  - *LaserVertical* was horizontally centered on the sensor frame, with 90° roll (thus scanning vertically) and zero pitch;
  - *LaserPort* was located on the port side of the vehicle, slightly pointing down to the ground (a few degrees of pitch) and zero roll;
  - *LaserStarboard* was on the starboard side of the vehicle, with both pitch and roll angles close to zero;

Table 1. Range Sensor Specifications

Sensor	Model	Maximum range	Range resolution	Horizontal FOV	Angular resolution	Scanning rate[a]
Lasers	Sick LMS291/221	81 m	0.01 m	180°	0.25°	≈ 18 Hz
Radar	ACFR Custom built	40 m	0.2 m	360°	≈ 1.9°	≈ 3 Hz

[a]For the lasers, each full 180°scan is actually composed of four partial scans of resolution 1°,
i.e. the time stamping is four times more accurate.
FOV, field of view.

- A 94-GHz frequency modulated continuous wave (FMCW) radar (custom built at ACFR for environment imaging), whose specifications are given in Table 1.
- A mono-CCD color camera equipped with a 6.5-mm lens, acquiring images at a nominal frame rate of 15 frames per second (fps) in *static*[1] data sets and 10 fps in *dynamic* data sets[2] (see Table 2).
- A thermal infrared (IR) camera, with a spectral response range of 7–14 μm. The infrared images were obtained through a frame grabber (see the details in Table 2).

The vehicle was also equipped with a number of proprioceptive sensors, providing information such as wheel angular velocities, engine rotation rate and brake pressures.

The navigation solution was provided by a commercial off-the-shelf Novatel RTK DGPS/INS SPAN[3] unit, composed of a Novatel ProPak-G2plus GPS receiver and a Honeywell HG1700 AG17 Inertial Measurement Unit (IMU). This usually provides a 2-cm accuracy localization, with a constant update of the estimated uncertainties on this solution. This navigation solution was output at 50 Hz, in double precision.

2.2. Coordinate Frames

The frames used in this work are illustrated in Figure 1. They are defined as follows:

- The *Navigation* frame is a fixed global frame defined by the three axes: $X^n = North$, $Y^n = East$ and $Z^n = Down$ in which positions are expressed in UTM (Universal Transverse Mercator) coordinates.
- The *Body* frame is linked to the body of the vehicle. Its center is located at the center of the IMU, approximately at the center of the vehicle. The axes are X^b pointing towards the front of the vehicle, Y^b pointing to the starboard side of the vehicle, and Z^b pointing down.
- A *Sensor* frame is linked to a particular sensor. Its axes are defined in a similar way as the previous one (i.e. X^s forward, Y^s starboard, Z^s down), but it is centered on the considered sensor.

Note that in the rest of the document, the term navigation (or localization) will correspond to the six-degree-of-freedom (DOF) global positioning of the *Body* frame in the *Navigation* frame.

2.3. Sensor Calibration

The spatial transformations between sensors and reference frames have been estimated using thorough calibration methods. Consequently, the sensor data are ready to be used to build 3D representations of the world and to achieve accurate multi-sensor data fusion. Two categories of calibration have been made:

- a *Range Sensor Calibration*, to estimate the transformations between the *Sensor* frame associated with each range sensor and the *Body* frame;
- a *Camera Calibration*, to estimate the intrinsic geometric parameters of each camera, and the extrinsic parameters of the transformations between cameras and lasers.

2.3.1. Range Sensor Calibration The estimation of the transformations between the frame associated with each range sensor (laser scanner or radar) and the *Body* frame was made using a technique detailed in Underwood et al. (2010). It allows for joint calibration of multiple range sensors, to minimize systematic errors in individual sensors as well as the systematic contradiction between sensors. This is needed to perform low-level data fusion between all of these sensors. For that purpose, a data set was acquired in an open area with key geometric features such as a flat ground, a vertical wall and a few vertical poles (see Peynot et al. 2009a). The calibration routine estimates the transformations between the *Sensor* frames and the common *Body* frame that minimizes a global cost function representing how accurately the laser point clouds match the known geometry of the key features.

The outputs of this calibration are the estimation of the three rotation angles (*RollX*, *PitchY* and *YawZ*) around the frame axes and three translation offsets (*dX*, *dY*, *dZ*) that fully describe the transformation from the *Body* frame to the *Sensor* frame. Table 3 shows the results obtained after combined calibration of all four range sensors, i.e. *LaserHorizontal* (or *LaserH*), *LaserVertical* (or *LaserV*), *LaserPort* (or *LaserP*), *LaserStarboard* (or *LaserS*) and the *Radar*. All angles are expressed here in degrees for convenience, and distances in meters.

Table 2. Camera Specifications

Camera	Model	Image size	Horizontal FOV	Vertical FOV	Rate	
					Static	Dynamic
Visual	Prosilica GC1380CH	1,360 × 1,024	68.2°	53.8°	15 fps	10 fps
Infrared	Raytheon Thermal-Eye 2000B	640 × 480	35.8°	27.1°	12.5 fps	12.5 fps

FOV, field of view.

Table 3. Transformations of the *Body* Frame to the *Sensor* Frame

Sensor	RollX	PitchY	YawZ	dX	dY	dZ
LaserH	−0.7328	−8.5869	−1.6313	0.1090	0.0083	−0.9197
LaserV	88.5630	−0.1180	−1.1231	−0.0003	−0.0823	−1.1268
LaserP	−0.5002	−2.6162	−1.8059	0.1909	−0.5488	−0.7638
LaserS	−0.6082	−0.4311	−2.3500	0.1987	0.5343	−0.8495
Radar	−0.1516	191.1617	173.2781	−0.0258	−0.0472	−1.3991
		in degrees			*in meters*	

2.3.2. Camera Calibration For each type of camera (IR and visual) a two-step calibration was executed. It is described as follows. The first step is to estimate the intrinsic (geometric) parameters of each camera using the *Camera Calibration Toolbox for Matlab* (Bouguet 2008). This requires a set of images featuring a chess board with squares of known dimensions. The second step then estimates the extrinsic transformations between cameras and lasers, using a method adapted from Pless and Zhang (2003). This method exploits the relation between the planar chess board surface as seen by the camera and the laser scanline on this same planar pattern. The same process was used for both color and IR cameras. The only difference concerned the chess board. Indeed, for the calibration of the IR camera, a chess board had been printed on thick paper and stuck on a planar isolating material (8-mm thick corrugated PVC sheet) using adhesive tape on the borders. It was then heated by exposing it to direct sunlight during the acquisition of the calibration images. The sizes of the black and white squares of these chess boards were as follows:

- 114.8 mm on both sides for the IR camera;
- 74.9 mm on the *left–right* axis as can be seen in the images and 74.7 mm on the axis corresponding to the direction *up–down* for the Visual camera.

Note that the data sets (laser scans and images) that were used for these calibrations are provided, next to the multi-sensor data sets, so that the user can perform any other calibration method that relies on similar input data (same type of calibration features, in particular). The images in these data sets feature the chess board exposed with various orientations in space, and at various distances, as appropriate for the Matlab camera calibration toolbox that was used.

Table 4. Lasers to Visual Camera Transformations

***LaserHorizontal* to visual camera:**					
δX_c	δY_c	δZ_c	ϕX_c	ϕY_c	ϕZ_c
0.4139	−0.2976	−0.0099	−4.7341	−0.3780	−0.4230
***LaserVertical* to visual camera:**					
δX_c	δY_c	δZ_c	ϕX_c	ϕY_c	ϕZ_c
0.5045	−0.0905	−0.208	−13.2030	−0.5851	−88.3628
***LaserPort* to visual camera:**					
δX_c	δY_c	δZ_c	ϕX_c	ϕY_c	ϕZ_c
0.9592	−0.5011	−0.0867	−10.6026	−0.0747	−0.5791
***LaserStarboard* to visual camera:**					
δX_c	δY_c	δZ_c	ϕX_c	ϕY_c	ϕZ_c
−0.1343	−0.4976	−0.0532	−12.6652	0.2409	−0.5293
in meters			*in degrees*		

The results of the intrinsic calibration of both cameras can be found in Peynot et al. (2009a). The estimated extrinsic parameters are given in this section. The offset translations (δX_c, δY_c,δZ_c) and rotations (ϕX_c, ϕY_c,ϕZ_c), indicated in Table 4 (visual camera) and Table 5 (IR camera), describe how to move each laser so that it aligns with the camera. They are expressed in the camera frame, using the Matlab Toolbox convention (i.e. $+X_c$ to the right, $+Y_c$ down, $+Z_c$ forward, Figure 3). In these tables distances are expressed in meters and angles are expressed in degrees.

2.4. Time-synchronization

As the UGV platform used in this work was equipped with several computers, the sensor data were coming from various sources with their own clocks. For example, all

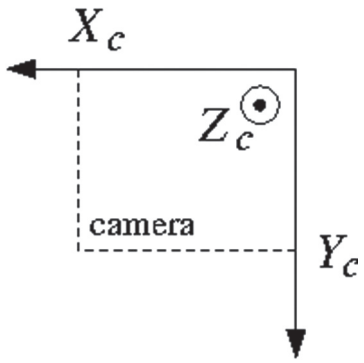

Fig. 3. Camera frame in the Matlab calibration toolbox.

Table 5. Lasers to Infrared (IR) Camera Transformations

***LaserHorizontal* to IR camera:**					
δX_c	δY_c	δZ_c	ϕX_c	ϕY_c	ϕZ_c
−0.3391	−0.3278	0.0975	−6.5307	−1.2671	−2.1308
***LaserVertical* to IR camera:**					
δX_c	δY_c	δZ_c	ϕX_c	ϕY_c	ϕZ_c
−0.2485	−0.1207	−0.0115	−14.9996	−1.4742	−90.0505
***LaserPort* to IR camera:**					
δX_c	δY_c	δZ_c	ϕX_c	ϕY_c	ϕZ_c
0.2090	−0.5400	0.0194	−12.7686	−1.0343	−2.3348
***LaserStarboard* to IR camera:**					
δX_c	δY_c	δZ_c	ϕX_c	ϕY_c	ϕZ_c
−0.8772	−0.5652	0.0584	−15.7179	−0.8259	−3.3619
in meters			*in degrees*		

proprioceptive and navigation data were collected by the low-level control computer, running the real-time operating system QNX, while all laser data were gathered by the "Sensor Server" computer running Linux and the radar data were logged on a separate computer running QNX as well. To account for timing differences, all data were systematically time-stamped at the time of their acquisition, and NTP (Network Time Protocol) was used to reduce the time-synchronization errors between the computers to less than 3 ms[4].

3. The Data Sets

This section focuses on the actual data collected, describing in particular the two main types of data sets (*static* and *dynamic*), the natures of the perceived areas, and the controlled environmental conditions.

3.1. Static Tests

The static tests consisted of sensing a fixed "reference" terrain, containing simple known objects, from a motionless vehicle (Figure 4). The positions and sizes of all objects within the pre-defined test area being known (their measurements are given in Peynot et al. (2009a)), the actual geometry of the environment can be used as a *ground truth* (within the hand measurement error) to evaluate the ability of each sensor, or sensor combination, to accurately represent the environment.

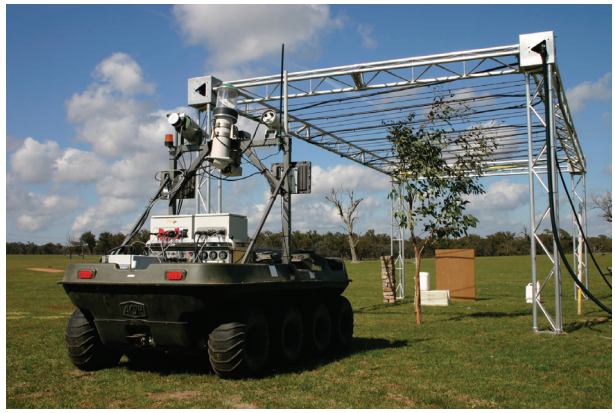

Fig. 4. The Argo UGV sensing the *static* trial area.

3.2. Dynamic Tests

For these tests, data were acquired from a moving vehicle in three different areas, representative of typical UGV operating environments. The first area, named the *Open area*, was mainly composed of a large, roughly flat ground, delimited by fences, a metallic shed, a static car and a trailer. The *Houses area* comprised a few barracks, a static car and a few trees and was delimited by fences. The *Dam area* featured trees, a static car, a couple of trailers and a small lake. Illustrations of these areas can be found with the data. The average linear speed of the vehicle was 0.5 m s^{-1}, with a maximum of 1.5 m s^{-1}. Variable yaw rates were achieved, with a maximum of 1.12 rad s^{-1} (i.e. 64° s^{-1}).

3.3. Controlled Environmental Conditions

For both categories, data have been gathered in controlled environmental conditions, which included the presence of airborne dust, smoke and rain. Dust clouds were generated by blowing air across dry soil using a high-power air compressor. Smoke clouds were generated using emergency smoke bombs that worked for about 1 minutes, and having the wind naturally carry them across the test area. Rain was generated using two different procedures. In the static tests, water was quite homogeneously and continuously spread in the test area using sprinklers attached to the top of the large metal frame seen in Figure 4. However, in the dynamic test, rain was simulated by spraying water with a hand-held hose in front of the vehicle throughout the corresponding data set.

3.4. Summary

Let a particular *data set* consist of the continuous acquisition of synchronized data from all sensors for a few minutes[5]. The data presented in this paper consist of a total of 40 separate data sets, in addition to three calibration-dedicated data sets, for a total amount of about 400 GB of raw data. They are published with a technical report (Peynot et al. 2009a), describing all of the details of the sensor characteristics, formats and content of files, at http://sdi.acfr.usyd.edu.au/.

4. Conclusion

In this work, large, accurately calibrated and synchronized, multi-modal data sets have been gathered in controlled environmental conditions (including the presence of dust, smoke and rain) by a representative UGV equipped with various types of sensors. These data sets have been made available to the public to test and compare perception algorithms. This is all the more possible thanks to tests in a static environment where the sensors perceived a "reference" scene with known objects geometry characteristics that may be used as "ground truth". In addition, while illustrating interesting and challenging cases for perception of on-board UGVs, these data were gathered with no very specific algorithms in mind, unlike most available data sets in the literature. This significantly reduced the experimenter's bias.

Possibilities of future work on perception exploiting these data sets are numerous. They include the promotion of sensor data integrity and reliable perception in outdoor environments, for 2D/3D terrain representation, obstacle detection or even pedestrian detection and tracking.

Funding and Acknowledgements

This work was supported by the US Air Force Research Laboratory through the Asian Office of Aerospace Research and Development (grant AOARD-08-4059) and the ARC Centre of Excellence programme, funded by the Australian Research Council (ARC) and the NSW State Government. The authors would also like to thank Craig Rodgers, Marc Calleja, James Underwood, Andrew Hill and Tom Allen for their valuable contribution to the gathering and processing of these data.

Notes

1. See Section 3 for a definition of *static* and *dynamic* in this context.
2. A slightly reduced frame rate was used for the *dynamic* data sets to avoid data loss due to factors such as the vibrations of the system which affect the efficiency of the hard disk writing.
3. Synchronised Position Attitude and Navigation.
4. Except for the radar data in sets 25 to 28, i.e. dynamic data sets at nighttime, where the synchronization error with all other sensor data was less than 10 ms.
5. The actual duration of each data set is available in Peynot et al. (2009a)

References

Bouguet, J.-Y. (2008). *Camera Calibration Toolbox for Matlab*. California Institute of Technology, http://www.vision.caltech.edu/bouguetj/calib_doc/.

Howard, A. and Roy, N. (2003–2008). *Radish: The Robotics Data Set Repository*. http://radish.sourceforge.net/.

Kelly A., Stentz A., Amidi, O., Bode, M., Bradley D., Diaz-Calderon A., et al. (2006). Toward reliable off road autonomous vehicles operating in challenging environments. *The International Journal of Robotics Research*, 25: 449–483.

Leonard, J., How, J., Teller, S., Berger, M., Campbell, S., Fiore, G., et al. (2007–2008). *MIT DARPA Urban Grand Challenge team Public Data*. http://grandchallenge.mit.edu/wiki/index.php/PublicData.

Matteucci et al. (2009). *The Rawseeds Project*. Politecnico Di Milano, Italy http://www. rawseeds.org.

Nebot E. (2000). *Victoria Park Dataset*. ACFR, The University of Sydney, http://www-personal.acfr.usyd.edu.au/nebot/dataset.htm.

Peynot, T., Terho S., and Scheding S. (2009a). *Sensor Data Integrity: Multi-sensor Perception for Unmanned Ground Vehicles*. Technical Report ACFR-TR-2009-002, ACFR, The University of Sydney.

Peynot T., Underwood J., and Scheding, S. (2009b). Towards reliable perception for unmanned ground vehicles in challenging conditions. *IEEE/RSJ International Conference on Intelligent Robots and Systems*, St Louis, MO, USA.

Pless R. and Zhang Q. (2003). *Extrinsic Calibration of a Camera and Laser Range Finder*. Technical report, Washington University, St Louis, MO.

Smith, M., Baldwin, I., Churchill W., Paul, R. and Newman P. (2009). The new college vision and laser data set. *The International Journal of Robotics Research*, 28(5): 595–599.

Underwood, J., Hill, A., Peynot T. and Scheding S. (2010). Error modeling and calibration of exteroceptive sensors for accurate mapping applications. *Journal of Field Robotics*, 27(1): 2–20.

Urmson, C., Anhalt, J., Bagnell, D., Baker, C., Bittner R., Clark, MN. et al. (2008). Autonomous driving in urban environments: Boss and the urban challenge. *Journal of Field Robotics*, 25(8): 425–466.