

Heterogeneous Sensor Fusion via Confidence-rich 3D Grid Mapping: Application to Physical Robots

Eric Heiden[†], Daniel Pastor[†], Pradyumna Vyshnav, Ali-akbar Agha-mohammadi

Abstract Autonomous navigation of intelligent physical systems largely depend on the ability of the system to generate an accurate map of its environment. Confidence-rich grid mapping algorithm provides a novel representation of the map based on range data by storing richer information at each voxel, including an estimate of the variance of occupancy. Capabilities and limitations are attributes of any given sensor, and therefore a single sensor may not be effective in providing detailed assessment of dynamic terrains. By incorporating multiple sensory modalities in a robot and extracting fused sensor information from them leads to higher certainty, noise reduction, and improved failure tolerance when mapping in real-world scenarios. In this work we investigate and evaluate sensor fusion techniques using confidence-rich grid mapping through a series of experiments on physical robotic systems with measurements from heterogeneous ranging sensors.

1 Introduction

Biological systems integrate information from multiple sources (*e.g.* vision, touch, etc.) to create a coherent and rich representation of the environment which allows them to reliably act in it. Robotic systems, too, largely depend on a map of their surroundings generated through various sensing instruments on-board to enable robust autonomous navigation in obstacle-laden environments. Consider a robot operating in a subterranean environment, such as a cave (see Fig. 1). This robot has to take into consideration



Fig. 1 Hybrid aerial-ground robot developed at NASA-JPL operating in a cave.

[†] Equal contribution

Eric Heiden

Department of Computer Science, University of Southern California, 3650 McClintock Ave, Los Angeles, CA 90089, e-mail: heiden@usc.edu

Daniel Pastor

Graduate Aerospace Laboratories, California Institute of Technology, 1200 E California Blvd, Pasadena, CA 91125, e-mail: dpastorm@caltech.edu

Pradyumna Vyshnav, Aliakbar Aghamohammadi

NASA Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91106, e-mail: pradyumna.vyshnav, aliagha@jpl.nasa.gov

the various physical constraints such as elevation, illumination, visual occlusion, etc. Although, in some cases, perception operates independently in the context of exploration, due to the ever-changing nature of the real-world setting in question, it is imperative to make use of robotic systems with multiple sensor modalities. This provides the ability to utilize the intrinsic benefits of each sensor system based on operational requirements.

1.1 Multi-modal Mapping

In mobile robotics, occupancy grids are the most common frameworks for modeling obstacles. Compared to state-of-the-art grid-based mapping methods, *Confidence-rich Grid Mapping* (CRM) [1] not only provides an accurate representation of the map, but also estimates the consistency between the map error and the reported confidence value in each voxel. When dealing with robotic systems fusing sensor data, this form of map representation is particularly useful as it can capture noise introduced by imperfect sensors and models more accurately.

1.2 Heterogeneous Sensors and Sensor Fusion

Depth cameras provide rich and dense information about the scene based on stereo matching or structured light (*e.g.* Kinect v1). However, the usable range of such sensors is typically considerably limited as the measurement error increases significantly at longer distances [2]. On the other hand, although scanning lasers are very reliable and accurate, even at larger ranges, they only provide range information in the horizontal plane. As a consequence of this, significant amounts of data must be collected to obtain detailed topographical information about a region [3]. Instead of relying on singular sensor modalities, it is beneficial to leverage a variety of sensors that complement each other [4] in order to improve measurement accuracy, robustness, and spatial and temporal coverage of perception pipelines.

2 Related Work

Grid-based maps have been constructed using a variety of range sensors, including sonars [5, 6], depth cameras [7], and scanning lasers [8]. Sensor fusion in occupancy grids for mobile robots was first studied by Moravec [9]. Detailed discussions on advantages of sensor fusion techniques in occupancy grids including different sensor architectures was presented by Elmenreich [4]. In 1987, Matthies and Elfes [10] combined improved versions of the sonar and stereo vision-based algorithms into a single method that builds maps integrating data from both sensors. This approach was extended to achieve robust data fusion of data from a monocular camera and a rangefinder [11]. The main technique using forward sensor models for integrating point-clouds obtained from different 3D sensors, in particular, time-of-flight sensors (Swiss-ranger, scanning laser rangefinders), and stereo vision cameras was developed by Pathak *et al.* [12]. A factor graph-based optimization framework, designed using a modular sensor-fusion system that allows for efficient and accurate incorporation of any navigation sensor of different sampling rates was discussed by Geneva

et al. [13]. Patel *et al.* [14] propose a deep learning architecture for the sensor fusion problem that consists of two convolutional neural networks (CNNs), each consisting of a different input modality, which are fused with a gating mechanism. Deep neural network architectures that are able to fuse information generated by multiple sensors and are robust to sensor failures at runtime are demonstrated by Bohnez *et al.* [15].

3 Problem Statement

The goal of this paper is to generate a confidence-aware representation of an unknown environment using measurements obtained from a set of heterogeneous sensors mounted on vehicles with multi-modal capabilities. In particular, we adopt CRM as a spatial representation of the environment to fuse sensor data from a multitude of robots and sensors. The contributions of this paper are *a*) the real-time implementation of distributed CRM, and *b*) testing and analyzing CRM as a consistent, accurate and convenient mapping algorithm for heterogeneous vehicles using different ranging sensors.

4 Technical Approach

To obtain more accurate and confidence-aware maps, CRM relaxes several incorrect assumptions made by traditional mapping algorithms. First, the dependency between the voxels inside the same measurement cone is taken into account for each map update. Second, the inverse sensor model is replaced by a novel “sensor cause model” which enables a principled approach to integrating forward sensor models into the map update algorithm. Finally, besides the mean of occupancy, each voxel in the CRM stores and estimates the variance as a confidence value of occupancy. We refer to [1] for more details on CRM and provide a short summary in the following.

4.1 Confidence-rich Representation

An occupancy map $m = [m^1, \dots, m^n]$ is defined as a set of values over a 2D or 3D grid of n voxels. The sensor measurement and the sensor configuration at the k -th time step are given by z_k and x_k , respectively. $b_k^m = p(m|z_{0:k}, x_{0:k})$ is the probability distribution (belief) of occupancy on the map m formulated in a Bayesian framework, by compressing the information obtained from past measurements $z_{0:k} = \{z_0, \dots, z_k\}$ and $x_{0:k} = \{x_0, \dots, x_k\}$. In [1], we show how the map belief can be updated iteratively via a linear term α^i and a constant term β^i for each voxel i as follows:

$$p(m^i|z_{0:k}, x_{0:k}) = (\alpha^i m^i + \beta^i) p(m^i|z_{0:k-1}, x_{0:k-1}) \quad (1)$$

In our C++ implementation, the map belief b^{m^i} for each voxel i is realized as an array of particles, *i.e.* floating-point 32-bit numbers, that are updated by the α^i and β^i coefficients. We have experimentally determined the number of particles to be

50 as a sufficient trade-off between mapping accuracy, computational and storage efficiency.

4.2 Sensor Cause Model

Unlike conventional occupancy-grid algorithms, CRM maintains the interdependence between voxels through the *Sensor Cause Model* (SCM). It is a probabilistic model that reasons about a voxel c_k being cause of the measurement z_k under sensor ray x_k and is proportional to the sensor forward model $p(z_k|c_k, x_k)$, as well as the probabilities of the measurement ray bouncing off that voxel c_k and traversing through all the voxels between c_k and the sensor:

$$p(c_k|z_{0:k}, x_{0:k}) = p(c_k|b_{k-1}^m, z_k, x_k) = \eta' p(z_k|c_k, x_k) \hat{m}_{k-1}^{c_k} \prod_{j=1}^{c_k^l - 1} (1 - \hat{m}_{k-1}^{g(j,x)}) \quad (2)$$

where c^l denotes the local index of voxel c along the ray x , i.e., $c^l = g^{-1}(c, x)$ and η' is the normalization constant.

4.3 Sensor Fusion in CRM

The Sensor Cause Model allows the CRM to embed forward models $p(z|c, x)$ for various sensors through the same update algorithm while considering the sensor's noise characteristics.

For example, for a stereo camera with focal length f and base line d_b , we can obtain the forward model $p(z|c, x) = \mathcal{N}(\|G^c - x_{cam}\|^{-1} f d_b, V)$ where G^c is the 3D point on the voxel grid estimated from the disparity matching between the two images recorded by the camera, x_{cam} is the camera's position and V is the variance of the Gaussian noise on the measured disparity.

The forward model provides us with its own uncertainty measurement that is embedded in the SCM such that the map update step (cf. Eq. 1) does not need to be hand-engineered for various sensors, i.e. it is *sensor-agnostic*, as opposed to conventional grid mapping algorithms. Measurements obtained from heterogeneous sensors contribute based on their accuracy to the map update, such that sensor fusion follows naturally from the Bayesian framework CRM is based on.

5 Experiments

In this section, we present experiments to analyze the performance of multiple sensors and fusion thereof on physical robotic systems. All the experiments were carried out at the Center for Autonomous Systems and Technologies (CAST³) located at California Institute of Technology (see Fig. 3 left). CAST has a $20 \times 12 \times 13$ m flying arena equipped with an *OptiTrack*⁴ motion capturing system instrumented with 48 high-definition video cameras. We manually control the robots and use OptiTrack pose estimation throughout all our experiments.

³ <http://cast.caltech.edu>

⁴ <http://optitrack.com>

5.1 Sensors and Vehicles

Throughout our experiments, we use the following two robotic mobility systems on which CRM was running onboard, implemented in C++ with ROS [16] integration. We provide a schematic data flow diagram of our hardware configuration in Fig. 2. While each robot independently computes the map onboard, a laptop serves as ground station to collect diagnostic information and data for visualization purposes.

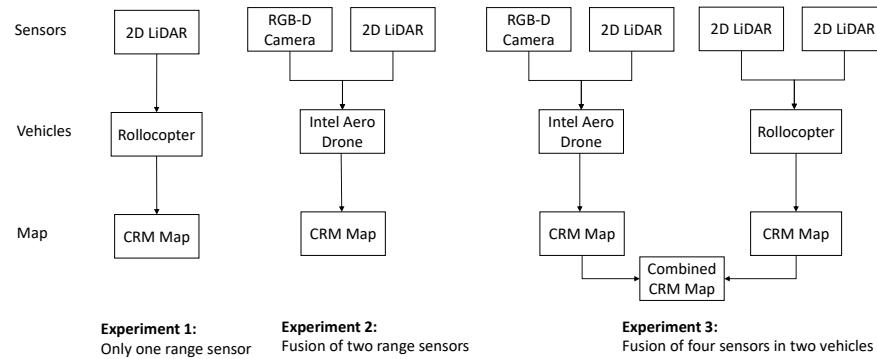


Fig. 2 Block diagram of the data flows in the three experiments that were conducted.

5.1.1 Rollocopter

Developed at NASA Jet Propulsion Laboratory (JPL), *Rollocopter* is a six degrees-of-freedom autonomous hybrid robot for both aerial and terrestrial modes (see Fig. 3 center). It is equipped with an NVIDIA Jetson TX2 microcomputer and two Hokuyo URG-04LX-UG01 LiDARs that measure ranges within a field of view of 240° and 0.36° angular resolution at a frequency of 10 Hz. Each Hokuyo sensor is rigidly mounted to the wheels which rotate as the vehicle moves, giving the planar depth sensors more coverage of the 3D space. The platform combines energy-efficient rolling on flat terrains with flying capabilities for more demanding scenarios.

5.1.2 Intel Aero Drone

The *Intel Aero* drone is a quadcopter powered by the Intel Aero Compute Board with an Intel Atom x7-Z8750⁵ CPU, equipped with an Intel RealSense R200 depth-sensing camera which records depth images at 20 Hz. In addition, we mounted a forward-facing Hokuyo URG-04-UG01 LiDAR (see Fig. 3 right) onto the vehicle.

5.2 Experiment 1

In our first experiment, we investigate the performance of CRM compared to log-odds mapping on Rollocopter with two LiDARs. We place boxes on the floor as obstacles to be registered in the map. This presents the first time that any mapping algorithm has been implemented on this novel robot platform.

⁵ <https://software.intel.com/en-us/aero/drone-kit>



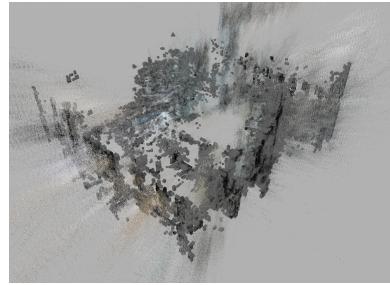
Fig. 3 Flying arena at CAST (left). Rollocopter (center) and Intel Aero Drone (right).

5.3 Experiment 2

We control the Intel Aero drone to take measurements with the RealSense camera and the Hokuyo LiDAR. Both sensors are aligned forward-facing in this experiment to compare how well CRM can leverage the combination of both sources of measurements in comparison to log-odds mapping. This experiment serves as our first sensor fusion application.

The depth measurements from both sensors are processed in real time by the mapping algorithms, at frequencies of 20 Hz and 10 Hz for RealSense and LiDAR, respectively.

Fig. 4 Confidence-rich grid map and overlaid point cloud generated using the RealSense stereo sensor and a 2D LiDAR on the Intel Aero drone from experiment 2.



5.4 Experiment 3

In our final experiment, we combine the two vehicles, Rollocopter and Intel Aero, to exchange map updates between each other and fuse measurements from their particular sensor suites. We manually control the robots in parallel to follow paths which provide a good coverage of the obstacle course (see Fig. 6 right). The RealSense on the Intel drone is attached downward-facing to obtain depth images of the floor which overlap significantly with the measurements taken by Rollocopter moving below the drone. This enables the mapping algorithm to fuse updates from both robots to obtain a more accurate model of the space that is traversable for Rollocopter. A Velodyne VLP-16 3D LiDAR is used to capture the ground-truth of the environment, as shown in Fig. 5.

With this experiment, we investigate how CRM can leverage heterogeneous sensors *and* heterogeneous robots in a multi-agent setting. We implement a procedure that allows for the exchange of map updates across different sensing agents: by broadcasting the α^i, β^i coefficients from Eq. 1 for each voxel i , the change in the

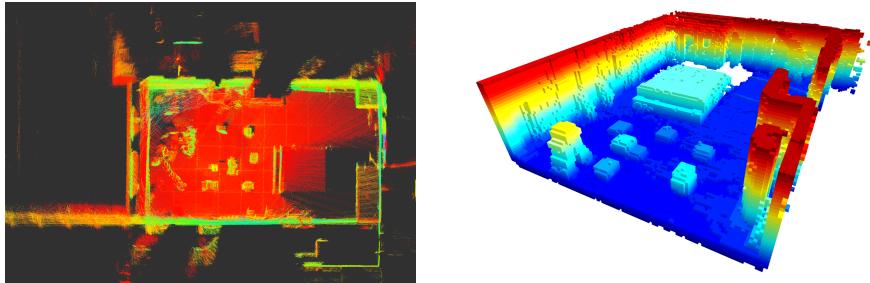


Fig. 5 *Left:* Orthographic top-down projection of point cloud of CAST captured with a Velodyne VLP-16 laser scanner. *Right:* Ground-truth map from Velodyne laser scans.

CRM map can be communicated efficiently without sending the 50 particle values for each voxel. Every 10 depth measurements, the CRM instance running on each robot broadcasts these map updates in the form of ROS messages, resulting in a frequency of ca. 2 Hz (see Fig. 6 left). The receiving CRM instance listens for such update messages and applies them in accordance to the Bayesian update Eq. 1 as soon as they arrive.



Fig. 6 *Left:* Architecture diagram for Experiment 3 with measurement rates for the different sensors (brown) and map update rates (green) between the two CRM instances (blue) running on Rollocopter and the Intel Aero Drone. *Right:* Experimental setup at CAST during Experiment 3.

6 Results

In this section, we demonstrate the performance of the heterogeneous sensors when applied to multi-modal physical robotic systems. For the first experiment, using a single robot and a single range sensor, Fig. 8 shows the map mean absolute error (MAE) over time for different sensor noise std deviations. We compare CRM against log-odds-based grid mapping in the *OctoMap* framework [17]. The MAE is averaged over all voxels that were updated throughout the mapping process to best show the improvement of the affected parts of the map. As an alternative measure of consistency, we compute the *Pearson correlation coefficient* between the true error and the estimated std. deviation (see Fig. 8), which indicates that the std. deviation estimated by CRM is highly correlated with the error, hence, it reliably captures the mapping confidence.

For the second experiment, using a single robot and multiple range sensors, we com-

pare the MAE for CRM and log-odds maps using both modalities combined and separated. Our results show that CRM yields higher accuracy than log-odds (Fig. 7) and they indicate that sensor fusion in occupancy grids benefit from our Bayesian formulation.

For the third experiment, using two robots and multiple range sensors on each robot, we show qualitative results of the resulting global map using CRM.

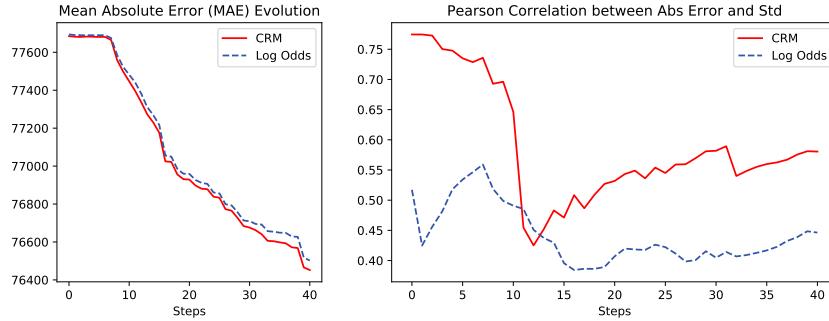


Fig. 8 Evolution of mean absolute error (left) and Pearson correlation coefficient between absolute error and estimated mapping std (right) for the LiDAR dataset obtained with Rollocopter (Experiment 1).

7 Experimental Insights

Range sensors have particular noise characteristics which we observed in the measurements obtained from the LiDAR sensors in our experiments.

Spurious Measurements: We observed that scanning lasers produce a variety of erroneous points in the vicinity of edges. These *phantom points* were usually found behind the edges close to the laser ray. This effect became more apparent at points on laser scans which were affected by a multitude of far-distanced objects. LiDARs, e.g. Velodyne VLP-16, that capture multiple returns for each single laser scan, are more robust to spurious measurements compared to the light-weight Hokuo LiDARs used on the robots in our experiments.

Surface Reflectance: As scanning lasers rely on the beam reflected from the object surface to the receiving unit, the strength of the returning signal is affected by the reflective properties of the surface. Such specular reflection is influenced by the

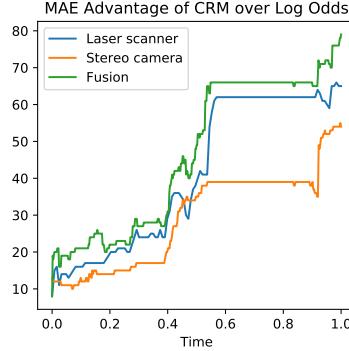


Fig. 7 Mean Absolute Error advantage of CRM over Log Odds under different sensor modalities (forward-facing LiDAR, forward-facing stereo camera, fusion of both) on the Intel Drone (Experiment 2). CRM shows an improvement in accuracy that increases with time regardless of the sensor.

distance to the surface, operating frequency and the incidence angle, among others. As shown in Fig. 9, accurate depth measurements from objects with shiny surfaces are particularly difficult to obtain and result in missing or erroneous data. Filtering approaches have been shown to effectively remove echo measurements resulting from reflective and transparent objects [18].

8 Conclusion and Future Work

Our experiments have demonstrated that CRM can be used in a distributed setting where a network of robots equipped with heterogeneous sensors collaboratively construct a consistent representation of the environment in real time. While log-odds mapping achieves a similar level of accuracy as the CRM map given measurements from two equal laser scanners, the consistency in CRM is significantly higher. Furthermore, CRM has been shown to leverage the advantages of heterogeneous sensors more effectively, leading to significant improvements in accuracy.

Future research is directed towards integrating CRM with planning. Preliminary results in a simulated environment [19] have shown promising perspectives for a trajectory optimization procedure that leverages CRM’s estimated confidence, besides the mean of occupancy. Such confidence-aware planners could improve the ability of mobile robots to autonomously navigate in unknown, unstructured environments. To enable true autonomy, CRM must be combined with a localization pipeline, *e.g.* visual-inertial odometry, such as Stereo MSCKF-VIO [20], or LiDAR-based odometry [21].

In some applications the communication link between the different sensor agents may be restricted or temporally blocked. Although we only send CRM updates instead of the complete map to reduce bandwidth, the issue of compressing the CRM map in an unstructured environment remains open for future research.

Acknowledgements

This research was carried out at the Jet Propulsion Laboratory under a contract with the National Aeronautics and Space Administration. U.S. Government sponsorship acknowledged.

References

1. Ali-Akbar Agha-Mohammadi, Eric Heiden, Karol Hausman, and Gaurav Sukhatme. Confidence-rich grid mapping. In *Proc. Int. Symp. Robot. Res.*, pages 1–19, 2017.
2. Kourosh Khoshelham. Accuracy analysis of kinect depth data. In *ISPRS workshop laser scanning*, volume 38, page W12, 2011.

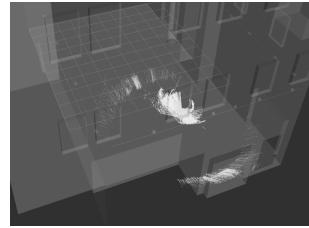


Fig. 9 While taking laser scans from reflective surfaces, such as the floor, a distinctive group of wrong measurements appears as a halo around the non-reflected measurements. This scan was taken by rotating the LiDAR around its axis. For clarity, a simplified CAD model of the testing area is displayed.

3. Juan Castorena, Charles D Creusere, and David Voelz. Modeling lidar scene sparsity using compressive sensing. In *Geoscience and Remote Sensing Symposium (IGARSS), 2010 IEEE International*, pages 2186–2189. IEEE, 2010.
4. Wilfried Elmenreich. An introduction to sensor fusion. *Vienna University of Technology, Austria*, 2002.
5. Hans P Moravec and Alberto Elfes. High resolution maps from wide angle sonar. In *Proceedings of the IEEE Conference on Robotics and Automation*, pages 19–24, 1985.
6. Brian Yamauchi. A frontier-based approach for autonomous exploration. In *IEEE International Symposium on Computational Intelligence in Robotics and Automation*, pages 146–151, 1997.
7. Paskin M and Thrun S. Robotic mapping with polygonal random fields. In *Conference on Uncertainty in Artificial Intelligence*, page 450458, 2005.
8. Sebastian Thrun. Learning occupancy grid maps with forward sensor models. *Autonomous robots*, 15(2):111–127, 2003.
9. Hans P Moravec. Sensor fusion in certainty grids for mobile robots. *AI magazine*, 9(2):61, 1988.
10. Alberto Elfes and Larry Matthies. Sensor integration for robot navigation: combining sonar and stereo range data in a grid-based representation. In *Decision and Control, 1987. 26th IEEE Conference on*, volume 26, pages 1802–1807. IEEE, 1987.
11. Petr Stepan, Miroslav Kulich, and Libor Preucil. Robust data fusion with occupancy grid. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 35(1):106–115, 2005.
12. Kaustubh Pathak, Andreas Birk, Jann Poppinga, and Soren Schwertfeger. 3d forward sensor modeling and application to occupancy grid based sensor fusion. In *Intelligent Robots and Systems, 2007. IROS 2007. IEEE/RSJ International Conference on*, pages 2059–2064. IEEE, 2007.
13. Patrick Geneva, Kevin Eckenhoff, and Guoquan Huang. Asynchronous multi-sensor fusion for 3d mapping and localization. Technical report, Tech. rep. RPNG-2017-002. Available: http://udel.edu/guang/papers/tr_async.pdf. University of Delaware, 2017.
14. Naman Patel, Anna Choromanska, Prashanth Krishnamurthy, and Farshad Khorrami. Sensor modality fusion with cnns for ugv autonomous driving in indoor environments. In *International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2017.
15. Steven Bohez, Tim Verbelen, Elias De Coninck, Bert Vankeirsbilck, Pieter Simoens, and Bart Dhoedt. Sensor fusion for robot control through deep reinforcement learning. In *Intelligent Robots and Systems (IROS), 2017 IEEE/RSJ International Conference on*, pages 2365–2370. Ieee, 2017.
16. Morgan Quigley, Ken Conley, Brian P Gerkey, Josh Faust, Tully Foote, Jeremy Leibs, Rob Wheeler, and Andrew Y Ng. ROS: an open-source Robot Operating System. In *ICRA Workshop on Open Source Software*, 2009.
17. Armin Hornung, Kai M. Wurm, Maren Bennewitz, Cyrill Stachniss, and Wolfram Burgard. Octomap: An efficient probabilistic 3d mapping framework based on octrees. *Autonomous Robots*, 34(3):189–206, April 2013.
18. R. Koch, S. May, and A. Nüchter. Detection and Purging of Specular Reflective and Transparent Object Influences in 3d Range Measurements. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pages 377–384, February 2017.
19. Eric Heiden, Karol Hausman, Gaurav S Sukhatme, and Ali-akbar Agha-mohammadi. Planning high-speed safe trajectories in confidence-rich maps. In *Intelligent Robots and Systems (IROS), 2017 IEEE/RSJ International Conference on*, pages 2880–2886. IEEE, 2017.
20. Ke Sun, Kartik Mohta, Bernd Pfommer, Michael Watterson, Sikang Liu, Yash Mulgaonkar, Camillo J Taylor, and Vijay Kumar. Robust stereo visual inertial odometry for fast autonomous flight. *IEEE Robotics and Automation Letters*, 3(2):965–972, 2018.
21. Ji Zhang and Sanjiv Singh. Loam: Lidar odometry and mapping in real-time. In *Robotics: Science and Systems Conference*, July 2014.