Ivan Petrović[1], Marija Seder[1], Ivan Marković[1]

Autonomous Navigation of Mobile Robots in Complex Dynamic Environments

[1]University of Zagreb Faculty of Electrical Engineering and Computing (UNIZG-FER), Zagreb, Croatia

Abstract

Most of the future robots will be mobile, and the main challenge is to develop algorithms for their autonomous navigation as well as for human-robot interactions. The Laboratory for Autonomous Systems and Mobile Robotics (LAMOR) at the Faculty of Electrical Engineering and Computing of the University of Zagreb is involved in the research of such mobile robotic systems, and currently participates in a number of related international and national research projects. This paper addresses the issue of autonomous navigation of mobile robots in complex dynamic environments, providing state of the art of the domain and major LAMOR's contribution to it. At the end, we present an application example of the autonomous navigation technologies in flexible warehouses, which we have been developing within a Horizon 2020 project SafeLog.

***Keywords**: autonomous systems, mobile robotics, navigation, safe human-robot interaction*

1. Introduction

There is no doubt that robotics, as a disruptive technology, will change the life as we know it over the next 50 years, enriching and augmenting all the aspects of life. The main robotics challenges in the next few decades will be to develop autonomous robotic systems that can perform complex tasks in human environments and safely cooperate with humans in arbitrary settings. Robots with these capabilities will transform our everyday lives as well as industrial processes, like the Internet, cell phones and computers had in the past two decades. Most of the future robots will be mobile, and therefore the main challenge is to develop algorithms for their autonomous navigation.

The paper is organized as follows. Section 2 and 3 describe respective activities of the Laboratory for Autonomous Systems and Mobile Robotics and the state of the art in the area autonomous navigation, while Section 3 describes the SafeLog project that deals with human navigation and safe interaction with robots in large flexible warehouses.

2. Laboratory for Autonomous Systems and Mobile Robotics (LAMOR)

LAMOR (lamor.fer.hr) is a research laboratory at the Faculty of Electrical Engineering and Computing of the University of Zagreb (UNIZG-FER) that holds expertise in autonomous mobile robotics systems. LAMOR's research activities are focused on the following aspects:

- Autonomous navigation of mobile robots in complex dynamic environments with three major research axes: (i) motion planning and control, (ii) simultaneous localization and mapping and (iii) detection and tracking of moving objects.
- Safe human-robot interactions to enable cohabitation of autonomous mobile robots and humans in the same environment with two major research axes: human intention recognition and human aware motion planning.

LAMOR's methodology relies on a strong coupling between theoretical research, algorithm development, experimental evaluations and a healthy dose of serendipity. The Laboratory is equipped with the state-of- the-art ground and aerial robotic platforms, advanced perception sensors, and a motion capture covered arena.

LAMOR has large experience in conducting international and national research projects. For example, LAMOR is currently involved in the following projects:

- SafeLog – Safe human-robot interaction in logistic applications for highly flexible warehouses (H2020 RIA project)
- L4MS – Logistics for Manufacturing SMEs (H2020 IA project)
- DIH2 – A Network of Robotics DIHs for Agile Production (H2020 DT-ICT-02-2018 – Robotics – Digital Innovation Hubs project)
- RoboCom++ – Rethinking Robotics for the Robot Companion of the future (FLAG-ERA project)
- SafeTRAM – System for Increased driving safety in public urban rail traffic (ERDF project)
- DUV-NRKBE – Development of a remotely controlled vehicle for operation in extreme CBRNe conditions (ERDF project)

- MAS – Development of a multi-functional anti-terrorism system (ERDF project), and
- DATACROSS – Advanced methods and technologies for data science and cooperative systems (ERDF – Top-level researches in Centres of Excellence project).

3. State of the art in autonomous navigation of mobile robots

3.1. Motion Planning and Control

Current state-of-the-art motion planning methods focus on the trajectory optimization aspects and they play two important roles in robot motion planning. Firstly, they can be used to smooth and shorten trajectories computed by other planning methods such as sampling-based planners. Secondly, they can be used to compute locally optimal, collision-free trajectories from scratch starting from naive trajectory initializations that might be in collision with obstacles.

The CHOMP algorithm introduced in [1] was one of the first successful attempts at using such methods in robotics. This method significantly outperformed naive RRT as well as grid search methods. STOMP introduced in [2] and ITOMP [3] further improved upon the CHOMP paradigm, exploring gradient-free optimization methods and dynamic obstacle avoidance, respectively. In [4] authors used sequential quadratic programming over a discrete trajectory representation, showing that it can be used to enforce both equality and inequality constraints. The underlying sparsity of the problem graph can be exploited by using exactly sparse Gaussian process (GP) regression [5]–[6]. GPs inherently provide a notion of trajectory optimality through a *prior*. Using this representation, a gradient-based optimization algorithm called GPMP (Gaussian Process Motion Planner) was proposed that can efficiently overcome the large computational costs of fine discretization while still maintaining smoothness of the result [7]–[9].

LAMOR's major contributions: In [10] we presented a framework for estimating intention of workers in a robotized warehouse. An active SLAM algorithm based on D* planning was introduced in [11], while a convergent navigation receding horizon control for differential drive robots was proposed in [12]. We also proposed for robot path planning a real-time approximation of clothoids with bounded error in [13].

3.2. Simultaneous Localization and Mapping

The simultaneous localization and mapping (SLAM) is a prerequisite for mobile robot's autonomy with applications in many areas, including modern logistics, autonomous driving, transportation, search and rescue missions, human assistance etc. The SLAM problem was introduced in the late 80's in the work of Smith et al. [14], but it came to focus at the beginning of 20th century. For a long period of time, the SLAM solutions were based on the filtering methods [14]-[17], but in the past years the focus has shifted to optimization methods that structures SLAM as an undirected graph in which nodes represent either the robot's pose or map's landmarks, and edges represent robot's observations. The approach uses *maximum a posteriori* method to find the relations of poses and landmarks that maximize the probability of consistent robot and landmark poses. Most prominent examples of this paradigm are square-root SAM [18], GraphSLAM [19], and *incremental* SAM (iSAM) [20]. In the last few years the implementation of graph based SLAMs has been made easier since there are open source libraries like g2o [21] and Ceres [22]. One of the most popular SLAM approaches in robotics are those using cameras, either in monocular or stereo setups. The most accurate and used approaches include semi-direct visual odometry method (SVO) [23], large-scale direct monocular SLAM (LSD-SLAM) [24], ORB-SLAM [25], and SOFT-SLAM [26].

LAMOR' major contributions: In [26] we introduced stereo visual odometry, dubbed SOFT, which ranked as the most accurate visual odometry on multiple popular datasets. In [27] we proposed LG-ESDSF, the exactly sparse delayed state filter on Lie groups which can solve SLAM accurately and efficiently. Combined with the SOFT odometry, it yielded SOFT-SLAM, a SLAM algorithm ranking first among visual SLAM approaches on several datasets. Finally, a theoretical foundation for LG-ESDSF lies in the extended information filter in Lie groups which we introduced in [28].

3.3. Detection and Tracking of Moving Objects

Tracking of a multiple moving object is another fundamental component of autonomous robotic systems. The objective of multi-target tracking (MTT) is to jointly estimate the number of objects as well as their dynamic states. In addition to the time varying number of targets, there are many other difficulties in MTT such as clutter detections (false alarms) and unknown association between detections and targets. In recent years there have been major breakthroughs in the MTT field resulting in diverse tracking algorithms, although most of them can be divided into these three paradigms: probabilistic data association (PDA) [29]-[30], multiple hypothesis tracking (MHT) [31]-[33] and random finite sets (RFS) [34]-[37].

LAMOR's major contributions: We introduced two interesting versions of the PHD filter: one on the unit circle with the von Mises distribution [38] and the other on Lie groups [39]. Furthermore, joint integrated PDA filter on Lie groups for MTT with the radar and stereo camera was introduced in [40], while a JIPDA using the von Mises-Fisher distribution was proposed in [41].

4. SafeLog project

4.1. About SafeLog

SafeLog – Safe Human-Robot Interaction for Highly Flexible Warehouses (safelog-project.eu) is a four-year H2020 research project (1/2016 – 12/2019, grant No 688117). It is coordinated by Prof. Björn Hein from the Karlsruhe Institute of Technology (KIT). It has six partners in total and, besides KIT, also includes the industrial partner Swisslog, whose automated warehouse system is used as a case-study in the project (Fig. 1).

Fig. 1. SafeLog partner Swisslog's automated CarryPick system

In the sequel we continue with the LAMOR research activities carried out within the project, and end with concluding remarks.

4.2. LAMOR Research Results

Moving objects detection using a wearable stereo camera

The aim is to detect moving objects from a stereo camera mounted on a human, as part of a Safety Vest. The stereo camera was modeled as two pinhole cameras which project the three-dimensional space on two-dimensional image plane. With a stereo camera it is possible to reconstruct the three-dimensional space back from its two two-dimensional projections. Apart from the Euclidean representation, which is the usual way of describing three-dimensional space, there are other representations. One such representation is the disparity space where x and y coordinates remain unchanged, but the third coordinate is the inverse of the depth z, and this inverse is called disparity. The disparity image is a way of showing the scene's depth by using pixel intensities.

We implemented the *Semi-global matching* method that efficiently computes consistent disparity images. Our problem was oriented to the reconstruction of depth from a video sequence. Previously computed disparities are used to improve the disparity computation, where the ego-motion estimation is used to transform the disparity from the previous step into the next step. Under the assumption of a static world, the transformed disparity is equivalent to the newly computed disparity. In reality, this is not true because of the noise (dynamic objects, discretization noise, errors in disparity computation etc.) and the transformed disparity will not always match the new one. Nevertheless, the disparity prediction from the previous step is still used. For each pixel we deterministically compute the displacement based on ego-motion and stochastically track the value of its disparity while updating its uncertainty through time with Kalman filtering. The disparity of each pixel is estimated by combining the newly matched (measured) disparity map and predicted disparity map. This way we managed to reduce the complexity of the algorithm.

The described framework is focused on stable, precise and fast spatio-temporal reconstruction, thus constraining the use case of the proposed method to static scenes. Although this can be seen as a limitation, in fact, this approach forms the base for dense stereo detection of dynamic objects by detecting discrepancies between static and dynamic flow. The assumption of the static world is not valid for the moving objects and they will cause discrepancies between the predicted and measured disparity images. By grouping such areas in the image, we manage to find the parts of the image with moving objects. In order to avoid the need to introduce any other sensors, we obtain ego-motion using the visual odometry algorithm.

The implemented algorithm is evaluated on real-world data from the KITTI dataset (Fig. 2) and the results are compared with an open source implementation of semi-global matching method in OpenCV library. The results show that our implementation is faster and more accurate than the implementation from OpenCV.

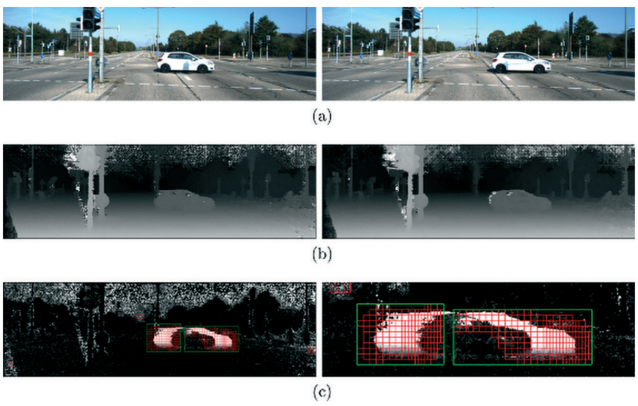

Fig. 2. Process of moving object detection. Part (a) shows the first and second scene in the sequence. Part (b) shows the predicted disparity map and the matched one. Lastly, part (c) shows the difference between the disparity maps and the final result of detection.

Multiple Moving Objects Tracking

This task is mainly concerned with the stochastic estimation of the state of multiple moving objects. We leverage moving object detections from the previous section and use them as inputs into the tracking algorithm. We present the Gaussian mixture PHD (GM-PHD) filter tested on various simulated and real-world scenarios. In Fig. 3 we show tracking of a rather complex scenario with 13 moving objects, from which we can see that even with a high clutter rate, the algorithm is capable of estimating the position of multiple moving objects on the scene.

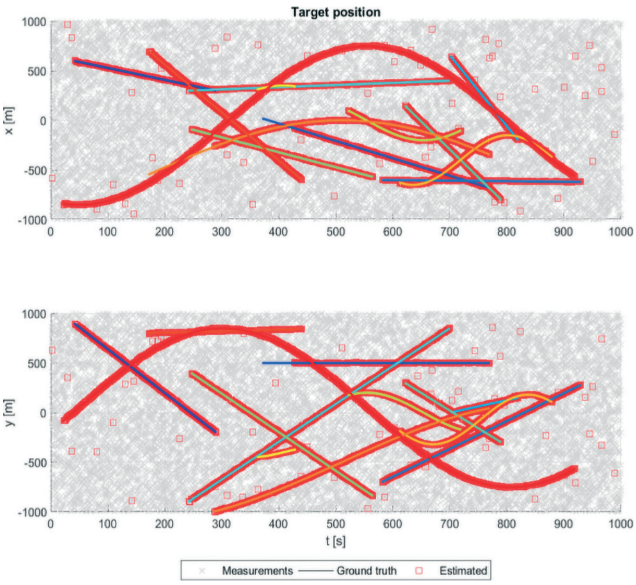

Fig. 3. GM-PHD tracking 13 moving objects in a high clutter rate.

For real-world experiments we used a dataset recorded with a hand-held PerceptIn Ironsides stereo camera at the premises of UNIZG-FER. Inputs into the tracking algorithm were detections produced from the previous section. In Fig. 4 we can see some examples for the Ironsides dataset.

Human Worker Localization with the Safety Vest

The aim is to develop a localization concept that will provide a stable and consistent location of the worker in the warehouse. The first part of this research dealt with estimating ego-motion of the worker from a stereo camera.

The stereo camera odometry can be seen as a sequence of several constituent blocks. Firstly, after a new stereo-pair acquisition, high-quality features are detected. Feature management starts with extraction and matching of corner-like features in both left and right images of the stereo pair. For this purpose, we utilize blob and corner masks on the gradient image and apply the non-maximum suppression, thus obtaining a set of available features. The features are then used in the matching

Fig. 4. Tracking example for the Ironsides dataset.

process, where the correspondences are determined by calculating the sum of absolute differences (SAD) over a pattern of pixels around the detected maxima. The inertial measurement unit (IMU) is used to predict the relative displacement of the worker in order to define a search radius for feature matching. Features are then weighted based on the distance to the predicted coordinates. The output of this step is a sparse feature set that can be further used within a RANSAC optimization procedure in order to yield final displacement information.

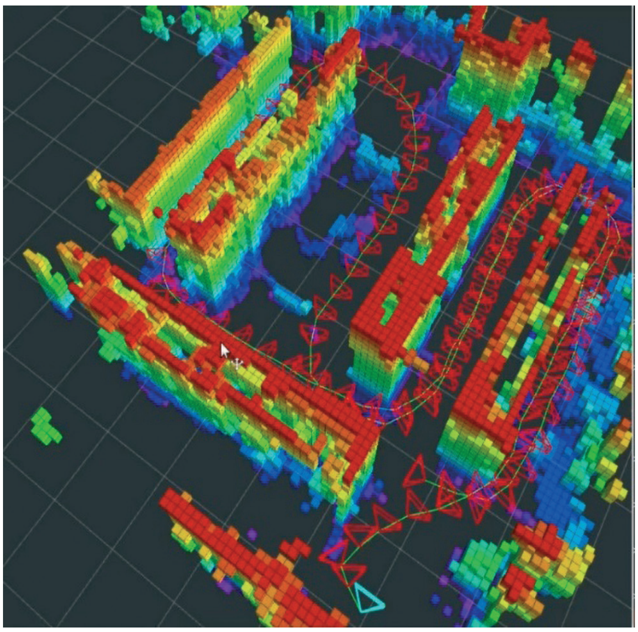

Fig. 5. An example of a 3D map built with the stereo visual odometry. Besides the map, the estimated trajectory (green line) is shown together with key frames (red pyramids).

Fig. 6. An example of an image of the stereo camera during the experiment in the library. The numbers represent tracked feature statistics that are used to decide whether to keep or discard the feature.

In order to test the odometry in a relevant scenario we made an experiment at the Library of UNIZG-FER that consists of bookshelf rows quite similar to the robotized warehouses of Swisslog. The main feature is that the lighting of the Library was also artificial, thus making the experiment even more relevant. In Fig. 5 we can see an example of a 3D map that was built in the library experiment, while in Fig. 6 we can see an example of the scene in the Library.

Human Worker Intention Estimation

The Fleet Management System of the automated warehouse needs to be able to estimate the worker's intentions correctly and control the robots accordingly, so that the warehouse operation efficiency is ensured.

Only the actions with the greatest influence on intention perception should be considered. For example, in the warehouse domain, the worker's orientation and motion have a large effect on the goal intention recognition and in this section, we track them using augmented reality glasses localization algorithm worn by the human worker. The human intention recognition algorithm is developed based on worker's movement validation which is used as observation for hidden Markov model (HMM) framework. Worker's movement is validated with respect to potential goal locations (i.e. warehouse racks and picking stations) using graph search algorithm on Generalized Voronoi Diagram's (GVD) nodes generated on the preexisting warehouse layout. If the mobile robot is located on a GVD edge, we cut that edge from the graph. The goal locations can be added or removed during the experiment.

The proposed HMM framework has one hidden state for every potential goal, one state indicating that the worker's intentions are not certain, and a state that declares the worker irrational meaning worker is not following path towards any. The irrational worker state includes the cases of the worker not following any proposed goal location or the worker's desire to go to an unknown goal. Every time the worker moves or turns significantly, we estimate the worker intention using Viterbi algorithm. The Viterbi algorithm outputs the most probable HMM's hidden states sequence and their probabilities which we consider intention estimates.

We carried out multiple experiments both in a real-world industrial setup in a test warehouse using augmented reality glasses and in a virtual reality generated warehouses in order to demonstrate the scalability of the algorithm (Fig. 7). Results corroborate that the proposed framework estimates warehouse worker's desires precisely and within reasonable expectations.

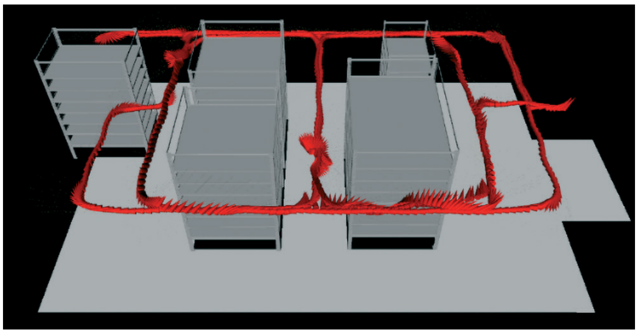

Fig. 7. Conducted experiment showcasing qualitatively the precision of the Hololens' localization. We have created s warehouse model in RViz – Robot Operating System's 3D visualization tool.

3.3. Ongoing research activities

As the SafeLog project is nearing its completion, we are working on the final integration and testing in a realistic real-world working automated warehouse. The final goal is to have a fully functional Safety Vest that will guarantee worker safety and localize the worker accurately in real-time relying just on the onboard sensors and onboard computing power. The data provided by the Safety Vest can then be utilized by other algorithms of higher safety levels, providing human intentions to the fleet management system, thus ensuring high efficiency of the warehouse and increasing worker comfort. We are optimistic that SafeLog results will also find its place in other industries and be exploited beyond the activities of the project itself.

References

[1] M. Zucker *et al.*, "Chomp: Covariant hamiltonian optimization for motion planning," *Int. J. Rob. Res.*, vol. 32, no. 9–10, pp. 1164–1193, 2013.

[2] M. Kalakrishnan, S. Chitta, E. Theodorou, P. Pastor, and S. Schaal, "STOMP: Stochastic Trajectory Optimization for Motion Planning," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA'11)*, 2011, pp. 4569–4574.

[3] C. Park, J. Pan, and D. Manocha, "ITOMP: Incremental Trajectory Optimization for Real-Time Replanning in Dynamic Environments," in *Proceedings of the International Conference on Automated Planning and Scheduling (ICAPS'12)*, 2012, pp. 207–215.

[4] J. Schulman *et al.*, "Motion planning with sequential convex optimization and convex collision checking," *Int. J. Rob. Res.*, vol. 33, no. 9, pp. 1251–1270, 2014.

[5] M. Mukadam, X. Yan, and B. Boots, "Gaussian process motion planning," in *Robotics and Automation (ICRA), 2016 IEEE International Conference on*, 2016, pp. 9–15.

[6] J. Dong, M. Mukadam, F. Dellaert, and B. Boots, "Motion Planning as Probabilistic Inference using Gaussian Processes and Factor Graphs.," in *Proceedings of Robotics: Science and Systems (RSS'16)*, 2016.

[7] J. Dong, B. Boots, and F. Dellaert, "Sparse Gaussian Processes for Continuous-Time Trajectory Estimation on Matrix Lie Groups," *Arxiv*, vol. abs/1705.0, 2017.

[8] E. Huang, M. Mukadam, Z. Liu, and B. Boots, "Motion planning with graph-based trajectories and Gaussian process inference," in *Robotics and Automation (ICRA), 2017 IEEE International Conference on*, 2017, pp. 5591–5598.

[9] M. Mukadam, J. Dong, X. Yan, F. Dellaert, and B. Boots, "Continuous-time Gaussian process motion planning via probabilistic inference," *arXiv Prepr. arXiv1707.07383*, 2017.

[10] T. Petković, D. Puljiz, I. Marković, and B. Hein, "Human Intention Estimation based on Hidden Markov Model Motion Validation for Safe Flexible Robotized Warehouses," *Robot. Comput. Integr. Manuf.*, vol. 57, pp. 182–196, 2018.

[11] I. Maurović, M. Seder, K. Lenac, and I. Petrović, "Path Planning for Active SLAM Based on the D* Algorithm With Negative Edge Weights," *IEEE Trans. Syst. Man Cybern. – Syst.*, 2017.

[12] M. Seder, M. Baotić, and I. Petrović, "Receding Horizon Control for Convergent Navigation of a Differential Drive Mobile Robot," *IEEE Trans. Control Syst. Technol.*, vol. 25, no. 2, pp. 653–660, 2017.

[13] M. Brezak and I. Petrović, "Real-time Approximation of Clothoids with Bounded Error for Path Planning Applications," *IEEE Trans. Robot.*, vol. 30, no. 2, pp. 507–515, 2014.

[14] R. Smith, M. Self, and P. Cheeseman, "Estimating Uncertain Spatial Relationships in Robotics," *Mach. Intell. Pattern Recognit.*, vol. 5, no. C, pp. 435–461, 1988.

[15] S. Thrun, Y. Liu, D. Koller, A. Y. Ng, Z. Ghahramani, and H. Durrant-Whyte, "Simultaneous localization and mapping with sparse extended information filters," *Int. J. Rob. Res.*, vol. 23, no. 7–8, pp. 693–716, 2004.

[16] M. Montemerlo, S. Thrun, D. Koller, and B. Wegbreit, "FastSLAM: A factored solution to the simultaneous localization and mapping problem," *Proc. 8th Natl. Conf. Artif. Intell. Conf. Innov. Appl. Artif. Intell.*, vol. 68, no. 2, pp. 593–598, 2002.

[17] G. Grisetti, C. Stachniss, and W. Burgard, "Improving Grid Based SLAM with Rao Blackwellized Particle Filters by Adaptive Proposals and Selective Resampling," *Int. Conf. Robot. Autom.*, no. April, pp. 2443–2448, 2005.

[18] F. Dellaert and M. Kaess, "Square root SAM: Simultaneous localization and mapping via square root information smoothing," *Int. J. Rob. Res.*, vol. 25, no. 12, pp. 1181–1203, 2006.

[19] S. Thrun and M. Montemerlo, "The graph SLAM algorithm with applications to large-scale mapping of urban structures," *Int. J. Rob. Res.*, vol. 25, no. 5–6, pp. 403–429, 2006.

[20] M. Kaess, A. Ranganathan, and F. Dellaert, "{iSAM}: Incremental Smoothing and Mapping," *IEEE Trans. Robot.*, vol. 24, no. 6, pp. 1365–1378, 2008.

[21] R. Kummerle, G. Grisetti, H. Strasdat, K. Konolige, and W. Burgard, "g2o: A General Framework for Graph Optimization," in *ICRA*, 2011.

[22] S. Agarwal, K. Mierle, and Others, "Ceres Solver." .

[23] C. Forster, M. Pizzoli, and D. Scaramuzza, "SVO: Fast semi-direct monocular visual odometry," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2014, pp. 15–22.

[24] J. Engel, T. Sch, and D. Cremers, "Direct Monocular SLAM," pp. 834–849, 2014.

[25] R. Mur-Artal and J. D. Tardós, "Orb-slam2: An open-source slam system for monocular, stereo, and rgb-d cameras," *IEEE Trans. Robot.*, vol. 33, no. 5, pp. 1255–1262, 2017.

[26] I. Cvišić, J. Ćesić, I. Marković, and I. Petrović, "SOFT-SLAM: Computationally Efficient Stereo Visual SLAM for Autonomous UAVs," *J. F. Robot.*, 2017.

[27] K. Lenac, J. Ćesić, I. Marković, and I. Petrović, "Exactly Sparse Delayed State Filter on Lie groups for Long-term Pose Graph SLAM," *Int. J. Rob. Res.*, vol. 37, no. 6, pp. 585–610, 2018.

[28] J. Ćesić, I. Marković, M. Bukal, and I. Petrović, "Extended information filter on matrix Lie groups," *Automatica*, vol. 82, 2017.

[29] Y. Bar-Shalom and E. Tse, "Tracking in a cluttered environment with probabilistic data association," *Automatica*, vol. 11, pp. 451–460, 1975.

[30] D. Mušicki and R. Evans, "Joint Integrated Probabilistic Data Association -- JIPDA," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 40, no. 3, pp. 1093–1099, 2004.

[31] D. Reid, "An algorithm for tracking multiple targets," *IEEE Trans. Automat. Contr.*, vol. 24, no. 6, pp. 843–854, 1979.

[32] S. He, H.-S. Shin, and A. Tsourdos, "Track-Oriented Multiple Hypothesis Tracking Based on Tabu Search and Gibbs Sampling," *IEEE Sens. J.*, vol. 18, no. 1, pp. 328–339, 2018.

[33] S. Coraluppi and C. Carthel, "Track Management in Multiple-Hypothesis Tracking.pdf," in *IEEE 10th Sensor Array and Multichannel Signal Processing Workshop (SAM)*, 2018, pp. 11–15.

[34] R. Mahler, "PHD filters of higher order in target number," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 43, no. 4, pp. 1523–1543, 2007.

[35] B.-N. Vo and W.-K. Ma, "The Gaussian Mixture Probability Hypothesis Density Filter," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4091–4104, 2006.

[36] B.-N. Vo, B.-T. Vo, and D. Phung, "Labeled Random Finite Sets and the Bayes Multi-Target Tracking Filter," vol. 62, no. 24, pp. 1–13, 2013.

[37] S. Reuter, B. N. T. Vo, B. N. T. Vo, and K. Dietmayer, "The labeled multi-Bernoulli filter," *IEEE Trans. Signal Process.*, vol. 62, no. 12, pp. 3246–3260, 2014.

[38] I. Marković, J. Ćesić, and I. Petrović, "Von Mises mixture PHD filter," *IEEE Signal Process. Lett.*, vol. 22, no. 12, pp. 2229–2233, 2015.

[39] J. Cesic, I. Markovic, and I. Petrovic, "Mixture Reduction on Matrix Lie Groups," *IEEE Signal Process. Lett.*, vol. 24, no. 11, pp. 1719–1723, 2017.

[40] J. Ćesić, I. Marković, I. Cvišić, and I. Petrović, "Radar and stereo vision fusion for multitarget tracking on the special Euclidean group," *Rob. Auton. Syst.*, 2016.

[41] I. Marković, M. Bukal, J. Ćesić, and I. Petrović, "Multitarget tracking with the von Mises-Fisher filter and probabilistic data association," *J. Adv. Inf. Fusion*, 2016.