Survey Article

Towards Collaborative Simultaneous Localization and Mapping: a Survey of the Current Research Landscape

Pierre-Yves Lajoie[1] , **Benjamin Ramtoula[2]** , **Fang Wu[3]** and **Giovanni Beltrame[1]**
[1]Department of Computer and Software Engineering, Polytechnique Montréal, Canada
[2]Oxford Robotics Institute, Department of Engineering Science, University of Oxford, United Kingdom
[3]Desay SV Automotive, Singapore

Abstract: Motivated by the tremendous progress, we witnessed in recent years, this paper presents a survey of the scientific literature on the topic of Collaborative Simultaneous Localization and Mapping (C-SLAM), also known as multi-robot SLAM. With fleets of self-driving cars on the horizon and the rise of multi-robot systems in industrial applications, we believe that Collaborative SLAM will soon become a cornerstone of future robotic applications. In this survey, we introduce the basic concepts of C-SLAM and present a thorough literature review. We also outline the major challenges and limitations of C-SLAM in terms of robustness, communication, and resource management. We conclude by exploring the area's current trends and promising research avenues.

Keywords: SLAM, cooperative robots, localization, mapping, perception

1. Introduction

Collaborative Simultaneous Localization and Mapping (C-SLAM), also known as multi-robot SLAM, has been studied extensively with early techniques dating back as far as the early 2000s (e.g., (Jennings et al., 1999; Fox et al., 2000; Thrun, 2001; Williams et al., 2002; Fenwick et al., 2002)). These techniques were introduced only a short time after the inception of single-robot SLAM by researchers who were already envisioning collaborative perception of the environment. Although they were small-scale proofs of concept, they laid the foundations that still shape the field nowadays.

After years of confinement to laboratory settings, C-SLAM technologies are finally coming to fruition into industry applications, ranging from warehouse management to fleets of self-driving cars. Those long awaited success stories are a strong indicator that C-SLAM technologies are poised to permeate other fields such as marine exploration (Paull et al., 2014; Bonin-Font and Burguera, 2020), cooperative object transportation (Rioux et al., 2015), or search and rescue operations (Tian et al., 2020c; Lee et al., 2020).

SLAM is the current method of choice to enable autonomous navigation, especially in unknown and GPS-denied environments. SLAM provides an accurate representation of the robot surroundings which can in turn enable autonomous control and decision making. Similarly, in multi-robot systems, C-SLAM enables collaborative behaviors by building a collective representation of the environment and a shared situational awareness.

Moreover, many ambitious applications remain for multi-robot systems, such as the exploration of other planets (Vitug, 2021; Ebadi et al., 2021). To reach those moonshot goals, ongoing trends in the research community aim to push the boundaries of multi-robot systems towards increasingly larger teams, or swarms of robots (Beni, 2004; Kegeleirs et al., 2021), which potentially allow parallel operations that are more efficient and versatile. However, this is still largely uncharted territory since current multi-robot applications either involve very few robots or rely upon large amounts of centralized computation in server clusters. Current C-SLAM techniques are no exception. They are prone to deteriorated performance when the team size increases above a few robots, and could be infeasible when minimal or no prior information is available about the operating environment.

Even though C-SLAM-enabled swarms of robots are still far from reality, C-SLAM remains a useful tool when operating as few as two autonomous robots. In exploration and mapping applications, even small teams can yield a significant boost in performance compared to a single robot system (Simmons et al., 2000). Notably, autonomous mapping using C-SLAM has recently received a lot of attention due to the latest DARPA Subterranean Challenge (DARPA, 2020) and its potential applications in space technologies (Bezouska and Barnhart, 2019).

Thus this paper presents a survey of the relevant literature on the topic of C-SLAM, aiming to give a complete overview of the main concepts, current developments, open challenges, and new trends in the field. We hope it will help new as well as established researchers to evaluate the state-of-the-art and offer valuable insights to guide future design choices and research directions. Compared to previous reviews (Saeedi et al., 2016; Rone and Ben-Tzvi, 2013), this paper provides an update on the tremendous progress in the past five years. In particular, we delve into the major advances towards the deployment of complete C-SLAM systems outside closely monitored laboratory environments, and we address the specific challenges of the different submodules (i.e., front-end, back-end, etc.). We also focus on the emergent trends and new opportunities coming from adjacent fields of research (e.g., deep learning, edge computing, etc.). This paper aims for a broader overview of the field than surveys covering specific C-SLAM subproblems such as map merging (Lee et al., 2012), practical implementations (Kshirsagar et al., 2018), particle filter techniques (Gupta and Conrad, 2019), vision-based techniques (Zou et al., 2019), and search and rescue applications (Queralta et al., 2020).

1.1. Outline

The rest of this paper consists of seven sections covering the main C-SLAM subfields of research presented in Table 1: Section 2 presents an overview of the single robot SLAM problem; Section 3 explains the core differences with C-SLAM; Section 4 explores the different modules of the C-SLAM front-end and their challenges; Section 5 introduces the C-SLAM back-end and discusses the different inference techniques; Section 6 looks into important system-level challenges. Section 7 discusses the available benchmarking datasets; Section 8 presents open problems and ongoing trends in the fields; and Section 9 concludes the survey and discusses future research avenues.

2. What is SLAM?

At its core, SLAM is a joint estimation of a robot's state and a model of its surrounding environment, with the key assumption that a moving robot performs the data collection sequentially. On one hand, the robot's state comprises its pose (position and orientation) and possibly other quantities such as sensors' calibration parameters. On the other hand, the environment model (i.e., the *map*) consists of representations of landmarks, built with processed data from the robot's exteroceptive sensors such as cameras or lidars. This makes SLAM an essential part of many applications that require building

Table 1. Collaborative Simultaneous Localization and Mapping Subfields of Research

SLAM	Odometry Intra-Robot Loop Closures Pose Estimation
C-SLAM Front-End	Direct Inter-Robot Loop Closures Indirect Inter-Robot Loop Closures Heterogeneous Sensing
C-SLAM Back-End	Extended Kalman Filters Particle Filters Pose Graph Optimization Perceptual Aliasing Mitigation
System-Level Challenges	Map Representation Communication Constraints
Open Problems	Resilient Inter-Robot Communication Managing Limited Computation Resources Adapting to Dynamic Environments Active C-SLAM Semantic C-SLAM Augmented Reality

an accurate map of the surrounding environment, whether it be for collision-free navigation, scene understanding, or visual inspection by a remote human operator. Since dead-reckoning approaches (e.g., IMU, wheel or visual odometry) drift over time due to noise accumulation, the environment map in SLAM is also used internally to correct the robot trajectory when known areas are revisited. The recovered links between previously visited locations are called loop closures. SLAM is useful when neither an a priori map nor localization information are available, when a map needs to be built, or long-term accurate localization estimates are required. Common scenarios include robotics applications without external positioning systems, such as the exploration of unknown indoor environments, caves, mines, or other planets.

2.1. Single-Robot SLAM problem

Formally, the overall goal of SLAM is to maximize the posterior of the map and robot state. We can formulate this with the state variables X of both the landmarks (map) and the robot, and the set of measurements Z acquired by the moving robot (Thrun et al., 2005):

$$p(X|Z) \tag{1}$$

This estimation problem is solved by either updating the current state at each time step given the new observations (i.e., filtering) or optimizing over the whole trajectory and past observations (i.e., smoothing).

Although filtering in SLAM is still an active research topic, current state-of-the-art techniques are mostly based on smoothing (Cadena et al., 2016; Rosen et al., 2021). The common formulation for smoothing techniques is a *Maximum A Posteriori* (*MAP*) estimation problem that leverages the moving robot assumption by introducing a prior distribution (e.g., obtained by odometry) over the robot trajectory.

Thus the SLAM problem for a single robot, designated with the lower case letter α, can be expressed as finding X_α^*, the solution of the *MAP* problem:

$$X_\alpha^* \doteq \underset{X_\alpha}{\operatorname{argmax}}\, p(X_\alpha|Z_\alpha) = \underset{X_\alpha}{\operatorname{argmax}}\, p(Z_\alpha|X_\alpha)p(X_\alpha) \tag{2}$$
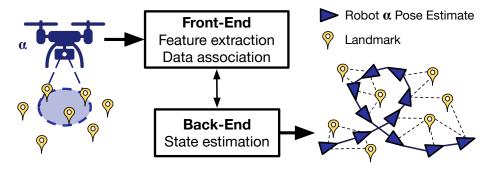
Figure 1. Single-robot SLAM Overview.

The decomposition of the posterior distribution is obtained with Bayes' theorem: $p(Z_\alpha|X_\alpha)$ is the likelihood of the measurements Z_α given a certain X_α, and $p(X_\alpha)$ is the prior distribution of the robot motion state. Intuitively, the SLAM problem finds the set of state variables (environment landmarks and robot poses) X_α^* that is most likely to produce the measurements Z_α given a prior estimation $p(X_\alpha)$.

It is important to also note that SLAM is closely related to the well-studied problem of bundle adjustment in *Structure from Motion* for which we refer the reader to (Özyeşil et al., 2017).

2.2. SLAM Systems Architecture

SLAM systems are commonly divided into a front-end and a back-end, each involving different fields of research. The front-end is in charge of perception-related tasks, such as feature extraction and data association which are both related to fields such as computer vision and signal processing. The back-end produces the final state estimates using the front-end's outputs. The back-end uses tools from the fields of optimization, probability theory and graph theory. In practice, the front-end processes the sensor data to generate ego-motion, loop closure, and landmark measurements, while the back-end performs the joint estimation of the map and the robot state. Figure 1 provides an overview of a common SLAM structure in which the robot trajectory is represented as a graph of poses at consecutive discrete times (i.e., a pose graph) and the map as a set of observed landmarks (Cadena et al., 2016). In a 3D pose graph, the nodes are the robot poses $[\mathbf{R}, \mathbf{t}] \in \mathrm{SE}(3)$ comprised of a rotation matrix $\mathbf{R} \in \mathrm{SO}(3)$ and a translation $\mathbf{t} \in \mathbb{R}^3$, and the edges represent the relative measurements between the poses (Barfoot, 2017).

Single-robot SLAM still faces many challenges that consequently apply to C-SLAM such as its long-term use, its robustness to perception failures and incorrect estimates, or its need for performance guarantees (Cadena et al., 2016). To circumvent those limitations in their specific settings, SLAM and C-SLAM developers often have to adapt the architecture and consider some trade-offs between the sensors capabilities, the onboard computing power, and available memory.

3. What is Collaborative SLAM?

Many tasks can be performed faster and more efficiently by using multiple robots instead of a single one. Whether SLAM is used to provide state estimation to support an application (e.g., estimate each robot's position to plan for actions), or whether it is at the core of the task (e.g., mapping an environment), it is beneficial and sometimes necessary to extend SLAM solutions into coordinated C-SLAM algorithms rather than performing single-robot SLAM on each robot.

C-SLAM algorithms aim to combine data collected on each individual robot into globally consistent estimates of a common map and of each robot's state. This coordination allows each robot to benefit from experience of the full team, leading to more accurate localization and mapping than multiple instances of single-robot SLAM. However, this coordination introduces many new features and challenges inherent to multi-robot systems.

3.1. Multi-robot systems

In multi-robot systems, data collection and state estimation are no longer entirely located on a single entity, so there is an inevitable need for communication between the agents (i.e., robots, base stations, etc.) which is the crux of the problem.

Moreover, multi-robot systems have additional properties to consider when designing C-SLAM systems, and taxonomies can be defined to classify approaches and highlight their benefits and tradeoffs. The taxonomy proposed in (Dudek et al., 1993) presents considerations that are well suited to the C-SLAM problem. It distinguishes approaches according to the following aspects:

Team size The number of robots in the system. Larger teams usually perform tasks more efficiently but may be harder to coordinate.

Communication range Direct communication between robots is limited by their spatial distribution and the communication medium. In some cases, robots might be unable to communicate for long periods of time, while in others they might always be in range of another robot.

Communication topology The communication network topology affects how robots communicate with one another. For example, they might be limited to either broadcast or one-to-one messages.

Communication bandwidth The bandwidth of the communication channel affects what information robots can afford to share.

System reconfigurability The robots will move and are likely to change spatial configuration over time. This can affect the communication topology and bandwidth.

Team unit processing ability Individual robot's computational capability can affect the computation cost of C-SLAM approaches and the distribution of computation tasks.

Team composition Robots can be homogeneous or heterogeneous over several aspects such as locomotion methods and available sensors.

The main differences between most C-SLAM techniques in the literature lie in the properties of the multi-robot system considered, especially their resource management strategy. One subclass of multi-robot systems particularly relevant to the future of C-SLAM are swarm robotics systems (Brambilla et al., 2013), which are inspired by social animals. Two main characteristics are required for swarm-compatibility in C-SLAM: robots' sensing and communication capabilities must be local, and robots can not have access to centralized control and/or to global knowledge. Such systems would present considerable benefits: they would have robustness to the loss of individual units, and they could scale well to large numbers of robots.

3.2. C-SLAM Problem definition

When all robots' initial states are known or can be estimated, the C-SLAM problem is a simple extension of the single-robot SLAM *MAP* problem that includes all the robots' states, measurements, and additional inter-robot measurements linking different robots' maps. In a setup with two robots (α, β), where X_α and X_β are the state variables from robot α and β to be estimated, Z_α and Z_β are the set of measurements gathered by robot α and β independently, $Z_{\alpha\beta}$ is the set of inter-robot measurements linking both robot maps containing relative pose estimates between one pose of robot α and one of robot β in their respective trajectories, and X_α^*, X_β^* are the solutions, the problem can be formulated as:

$$
\begin{aligned}
(X_\alpha^*, X_\beta^*) &\doteq \underset{X_\alpha, X_\beta}{\operatorname{argmax}} \, p(X_\alpha, X_\beta | Z_\alpha, Z_\beta, Z_{\alpha\beta}) \\
&= \underset{X_\alpha, X_\beta}{\operatorname{argmax}} \, p(Z_\alpha, Z_\beta, Z_{\alpha\beta} | X_\alpha, X_\beta) p(X_\alpha, X_\beta)
\end{aligned}
\tag{3}
$$

However, when the relative starting locations and orientations of the robots cannot be determined, the initial guess of the robots states $p(X_\alpha, X_\beta)$ is not available. In that case, there are infinite
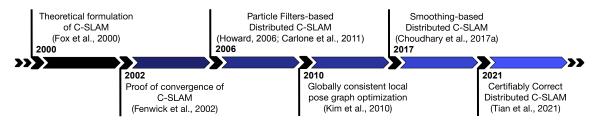
Figure 2. C-SLAM Problem Major Milestones.

possible initial alignments between the multiple robot trajectories. Therefore, in absence of a prior distribution, C-SLAM is reduced to the following *Maximum Likelihood Estimation* (MLE) problem.

$$(X_\alpha^*, X_\beta^*) \doteq \underset{X_\alpha, X_\beta}{\mathrm{argmax}} \, p(Z_\alpha, Z_\beta, Z_{\alpha\beta}|X_\alpha, X_\beta) \tag{4}$$

The C-SLAM problem formulation is still evolving to this day and progress still needs to be made to achieve an efficient decentralized, distributed and robust implementation. To give some perspective, Figure 2 presents some major milestones in the evolution of the C-SLAM problem over time. More details on these milestone works are provided in the following sections.

3.3. Centralized, Decentralized and Distributed Systems

An important distinction in C-SLAM, and in multi-robot systems in general, is the difference between the *global* and *local* perspectives. The local perspective is the default point of view in single-robot SLAM. Accordingly, the pose and map estimates are expressed in an internal reference frame which is usually the starting location of the robot's mission. However, in C-SLAM, one has to consider the global perspective of the system since the pose and map of each robot need to be expressed in a shared global reference frame. This means that every landmark can be expressed within the same coordinates system by every robot in the team. Otherwise, shared information (e.g., position of observed landmarks) would have no significance to the receiving robot due to the representation being in another unknown local reference frame. Establishing this global reference frame using C-SLAM allows the robots to collectively perceive the environment and benefit from each other's observations.

To achieve this global understanding, one could either solve C-SLAM in a centralized or decentralized manner. In a centralized solution, the estimator has a global view of the entire team of robots: it performs the estimation given perfect knowledge of the measurements of each robot. These measurements can be raw or preprocessed, and shared on demand depending on the communication limits.

Unfortunately, due to communication constraints, solving centralized C-SLAM quickly becomes intractable as the number of robots increases (Saeedi et al., 2016). Thus a better solution for scalability is to solve C-SLAM in a decentralized manner (Cieslewski et al., 2018). This means that each robot only has access to a local view comprised of its own data and partial information from its neighbors. Therefore, decentralized systems cannot solve the C-SLAM problem for all the robots at once, but aim instead for local solutions on each robot that are consistent with their neighbors. Then, iteratively and over time, with the robots gradually improving their estimates given their neighbors' latest data, decentralized techniques converge to local solutions that are mutually consistent across the team of robots. So, upon convergence, the individual robots reach a common understanding and their local maps are aligned with the common (global) reference frame. Figure 3 provides examples of the C-SLAM problem and output in both perspectives.

Aside from the centralized/decentralized classification, a system is distributed if the computation load is divided among the robots. The two notions are independent. Therefore, a system could be centralized and distributed at the same time, if, for example, each robot performs parts of the
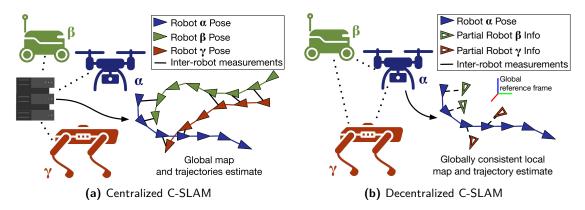
Figure 3. Illustration of centralized and decentralized approaches to solve the C-SLAM estimation problem. The decentralized illustration is from the local perspective of robot α.

computation, but a central node is required to merge the individual results from all the robots (Leung, 2012).

Centralized C-SLAM. Many seminal C-SLAM works are centralized and solve the estimation problem in eq. 4 from the global perspective. In those approaches, the robots are essentially reduced to mobile sensors whose data is collected and processed on a single computer. Examples of centralized C-SLAM techniques include (Andersson and Nygards, 2008; Kim et al., 2010) that gather all the robots' measurements at a central station to compute the global map. (Lázaro et al., 2013) improves this solution by marginalizing unnecessary nodes in the local pose graphs so only a few condensed measurements need to be shared to the central computer. Other centralized approaches (Forster et al., 2013; Schmuck and Chli, 2017; Schmuck and Chli, 2019) perform C-SLAM with monocular cameras, successfully solving the associated 3D estimation challenges, while (Loianno et al., 2015) focuses on micro-aerial vehicles constraints. (Deutsch et al., 2016) proposes a framework to reuse existing single robot SLAM solutions for C-SLAM. The same idea is explored in (Li et al., 2018), in which a popular single-robot SLAM technique (Mur-Artal and Tardós, 2017) is converted into C-SLAM. (Karrer et al., 2018; Karrer and Chli, 2018) integrate inertial measurements from IMUs in their centralized C-SLAM systems. (Jiménez et al., 2018) proposes that the central node spreads the resulting map across the robots to limit the memory usage.

Improving upon the pure centralized methods, some techniques do not rely on a single computer, but can use different robots or base stations for the computation. This way, the system can adapt itself to the failure of one node or communication link and complete the mission. A typical solution is to use replicated central servers among the robots (Bailey et al., 2011).

Decentralized C-SLAM. Solving the C-SLAM problem in a decentralized manner is radically different, but offers major benefits in terms of computation, communication and privacy (Choudhary et al., 2017a; Cieslewski et al., 2018). Such systems are usually distributed and solve the estimation problem from eq. 4 partially on each robot. As shown in Figure 3b, each robot computes its own local map and uses partial information from other robots as well as inter-robot measurements to achieve a local solution. Over several iterations with its neighbors, each robot's resulting local solution converges to a solution consistent with the global reference frame. These techniques mitigate communication and computation bottlenecks since the loads are spread across the robot team (Pfingsthorn et al., 2008). Alternatively, the full mapping data can be sent to every robot for redundancy and a subset of robots can be designated for computation (Saeedi et al., 2011a; Bresson et al., 2013; Saeedi et al., 2015).

As one would expect, decentralized and distributed techniques come with many additional challenges that need to be tackled such as complex bookkeeping, information double counting or synchronization issues.

3.4. Complete C-SLAM Systems

In C-SLAM, as well as in single-robot SLAM, the front-end handles perception-related tasks and the back-end generates state estimates using all measurements gathered. However, in C-SLAM, the front-end and back-end computations do not necessarily occur fully on a single robot anymore depending on the sensing, communication, and estimation strategies. For example, in a centralized system, all robots could send their sensor data directly to a single unit which would then perform the front-end and back-end steps for the whole team. While in a decentralized and distributed setup, a robot could perform feature extraction on its own and communicate with other robots for data association and state estimation. Every part of a C-SLAM system can be subject to distribution or decentralization.

In addition, the front-end needs to find links and relative measurements between the individual maps. Therefore, the front-end must also perform data association to detect and compute inter-robot loop closures, which will be detailed in Section 4. It follows that the back-end must generate estimates combining the individual and shared measurements as described in Section 5.

In the recent years, several complete C-SLAM systems have been developed and deployed in realistic scenarios. For example, some solutions deployed in large-scale environments during the DARPA Subterranean Challenge (Hudson et al., 2021; Agha et al., 2021) led to the developments of new C-SLAM systems, such as the robust lidar-based approach of (Ebadi et al., 2020). Alternatively, (Schmuck et al., 2021) proposes a vision-based centralized C-SLAM system incorporating inertial measurements, which has been tested with up to 12 robots in simulation. In another line of work, (Lajoie et al., 2020) presents a distributed and decentralized system robust to spurious measurements, along with online experiments on real robots, and a publicly available implementation. A subsequent approach, detailed in (Tian et al., 2021a), puts together a completed decentralized and distributed C-SLAM system including a novel robust distributed pose graph optimization back-end, and a front-end producing globally consistent metric-semantic 3D mesh models of the explored environment. Those works are some of the best starting points for researchers and engineers looking to implement, improve and deploy practical C-SLAM systems in challenging conditions.

4. C-SLAM Front-End

Although the division between the front-end and the back-end is sometimes blurry due to the presence of feedback loops between the two processes, a typical C-SLAM front-end is in charge of producing landmark estimates, odometry measurements, and both intra-robot and inter-robot loop closures.

Odometry measurements aim to capture the translation and rotation of a robot from one time step to the next. Common techniques measure wheel movements, integrate from an IMU, and/or perform geometric matching between consecutive images or laser-scans. Intra-robot loop closures are the measurements used by a SLAM system to relocalize itself and reduce its estimate error caused by odometry drift. Using place recognition, the system can detect previously visited locations and compute relative measurements between them. In other words, intra-robot loop closures are estimates relating non-consecutive poses in the robot trajectory that observed the same places. Since the computing of odometry and intra-robot loop closure measurements can be fully done locally on each robot, the approaches used are the same in both SLAM and C-SLAM. Thus we refer the reader to (Mohamed et al., 2019; Cadena et al., 2016; Lowry et al., 2016) for surveys of the current techniques.

Conversely, inter-robot loop closures relate poses of different robots trajectories. They are the seams that stitch together the estimates from multiple robots: they draw connections between the individual robots' local maps to build the global understanding of the environment. Generating inter-robot loop closures is the main focus of contributions to the front-end of C-SLAM systems, and key to ensure consistency of the estimates.

4.1. Direct vs Indirect Inter-Robot Loop Closures Measurements

Inter-robot loop closures can be classified as direct or indirect (Kim et al., 2010). Direct inter-robot loop closures occur when two robots meet, and they are able to estimate their current relative location with respect to each other through direct sensing. Indirect inter-robot loop closures are produced by looking back into maps to find partial overlaps for places that have been visited by both robots. Given these measurements, the robots can estimate the relative transformation between their maps. In general, indirect inter-robot loop closures detection produce more measurements and usually achieve a higher accuracy, but require more communication and processing. Indeed, the detection process is often the communication bottleneck of C-SLAM given the large amount of data required to compare landmarks between the individual local maps (Tardioli et al., 2015).

4.1.1. Direct Inter-Robot Loop Closures
The idea of direct inter-robot loop closures is to compute the relative pose between two or more robots when they physically meet in the same location. This is usually done through direct sensing of one another. For example, (Kim et al., 2010) operated a quadcopter and a ground robot and the latter was equipped with a checkerboard pattern that could be detected by the quadcopter's camera. (Zhou and Roumeliotis, 2006) used a combination of direct and indirect detection approaches, where colored cylinders were installed to be detected by omnidirectional cameras. In addition, (Gentner et al., 2018; Boroson et al., 2020; Cao and Beltrame, 2021) propose to replace visual loop closures by Ultra-Wide Band (UWB) measurements from beacons onboard the robots. Given a few distance measurements provided by the UWB sensors, the robots can estimate their current relative pose with respect to each other and establish a common reference frame.

4.1.2. Indirect Inter-Robot Loop Closures
Indirect inter-robot loop closure detection is the extension of single-robot loop closure detection to multiple maps. In fact, approaches to find indirect inter-robot loop closures often rely on the same core algorithms as intra-robot loop closures. The first challenge is the loop closures detection, which consists of detecting overlaps between the individual maps. This task is usually handled by a place recognition module which can efficiently compare new observations against previous sections of the robots' maps. Following place recognition matches, geometric estimation is performed to compute the relative pose between the two places.

In the case of visual sensors, the place recognition problem has been studied extensively (Lowry et al., 2016). The seminal work of visual bags of binary words (Galvez-López and Tardos, 2012) is still popular, but newer approaches based on deep learning, such as NetVLAD (Arandjelović et al., 2018), are more accurate and data-efficient. Loop closure relative pose measurements can be computed using visual features matching and multi-view geometry (Hartley and Zisserman, 2003).

Finding inter-robot overlaps is a harder task with 3D point clouds given the dense data that need to be shared and the lack of expressive features to perform place recognition. To that end, compact and robust global point cloud descriptors (Uy and Lee, 2018) can be relied upon to compare point clouds for place recognition. Other approaches extract features from the point cloud that can serve for place recognition while providing initial guesses for later geometric alignments (Ebadi et al., 2021), or even directly compute loop closure measurements (Dubé et al., 2017a). While the classical *Iterative Closest Point* method (Besl and McKay, 1992) is still commonly used in single robot SLAM to compute relative pose measurements between two matching point clouds, it is not well suited for multi-robot operation due to its reliance on a good initial guess, which is usually not available between the robots' local maps. To handle the initialization problem, early work from (Olson, 2009) presents a correlation-based algorithm that can be efficiently solved on a GPU for real-time scan matching. Another common solution is to use submaps matching for both stereo cameras (Schuster et al., 2015; Schulz et al., 2019; Dubois et al., 2020b) and lidars (Dubé et al., 2017b; Ebadi et al., 2021). During this process, multiple laser scans or 3D point clouds are clustered into submaps which can in turn be registered more efficiently.

4.2. Heterogeneous Sensing

In many applications, the teams of robots are composed of different platforms equipped with different onboard sensors. Heterogeneous sensing comes with the additional challenge of matching map data in different representation to perform relocalization and/or map fusion. To this end, a recent study evaluated the repeatability of existing keypoint detectors between data from stereo cameras and lidars For example, when matching data from both stereo cameras and lidars, one needs to chose repeatable 3D feature representations that are consistent despite the differences in density and field-of-view (Boroson and Ayanian, 2019). Another approach is to use an intermediate map representation that can be produced by different kinds of sensors (Koch and Lacroix, 2016). For example, (Käslin et al., 2016) proposes to compare elevation maps that are invariant to sensor choice: lidars or cameras.

4.3. Non-Conventional Sensing

While most C-SLAM techniques use the typical SLAM sensors (i.e., lidars and monocular, RGB-D, or stereo cameras), many recent research works have explored the use of non-conventional sensors: (Choi et al., 2014) uses omnidirectional (i.e., fish-eye) cameras, (Waniek et al., 2015) performs C-SLAM with event-based vision sensors, and (Morales and Kassas, 2018) integrates ambient radio signals (i.e., signals of opportunity) into their system. In a similar vein, (Liu et al., 2020) leverages existing WiFi access points in most indoor environments to perform loop closures based on their radio signal fingerprint. Alternatively, some approaches use only a few higher-level landmarks, such as objects, for tracking and place recognition (Cunningham et al., 2010; Choudhary et al., 2017b). This type of approach have regained popularity with the increasing performance of deep learning-based methods in semantic segmentation as discussed in Section 8.5.

5. C-SLAM Back-End

As mentioned before, the role of the C-SLAM back-end is to estimate the state of the robot and the map given the front-end measurements. The difference with single-robot SLAM is the presence of inter-robot measurements, the need to reach consensus, and the potential lack of an initial guess since the global reference frame and the starting positions of the robots are usually initially unknown. Nevertheless, similar to single-robot solvers, C-SLAM back-ends are roughly divided in two main categories of inference techniques: filtering-based and smoothing-based. Although filtering-based approaches were the most common among the early techniques (e.g., EKF (Rekleitis et al., 2003) and particle filters (Madhavan et al., 2004)), smoothing-based approaches quickly gained in popularity and are currently considered as superior in most applications (Strasdat et al., 2012). This section provides an overview of the different categories of estimation workhorses for C-SLAM and presents examples from the literature.

5.1. Filtering-Based Estimation

Filtering approaches are often characterized as online in the sense that only the current robot pose is estimated and all previous poses are marginalized out (Thrun et al., 2005) at each time step. Consequently, the estimation of the posterior in eq. 1 at time t only depends on the posterior at time $t-1$ and the new measurements.

The classical filtering technique for nonlinear problems (i.e., all problems in robotics except trivial ones) is the Extended Kalman Filter (EKF). It has been applied to C-SLAM in various ways among which the information filter method presented in (Thrun and Liu, 2005). In a nutshell, EKF are Gaussian filters that circumvent the linear assumptions of Kalman filters through linearization (i.e., local linear approximation); however, the linearization process potentially leads to inconsistencies when the noise is too large. A major advantage of EKF techniques (Thrun and Liu, 2005; Sasaoka

et al., 2016; Luft et al., 2016; Schuster et al., 2019) over smoothing techniques is that the covariance matrix is available without requiring additional computation, which can be useful for feature tracking or active exploration. For example, one could prioritize the exploration in the most uncertain directions. Yet, an explicit covariance matrix is rarely required, so alternative filtering techniques seek to avoid its computation, such as the smooth variable structure filters approach presented in (Demim et al., 2017).

Building on the EKF, Rao-Blackwellized Particle Filters (RBPF) (Doucet et al., 2000) are another popular filtering approach for the C-SLAM problem. Techniques, such as (Howard, 2006), use samples (particles) to represent the posterior distribution in eq. 1 and perform variable marginalization using an EKF which drastically reduce the size of the sampling space. (Carlone et al., 2011) extends on (Howard, 2006) and improves its consistency while making it fully distributed. (Gil et al., 2010) adapts RBPF to visual C-SLAM and (Dörr et al., 2016) showcases the potential of RBPF C-SLAM for industrial applications.

It is important to note that, according to theoretical analysis results (Mourikis and Roumeliotis, 2004), reducing the number of relative position measurements between the robots to a minimum, to limit communication and computation, only inflicts a small penalty on the localization performance. It was also shown that the presence of even only one robot equipped with an absolute positioning sensor is enough to bound the positioning uncertainty of the whole team. Additionally, analytical upper bounds can be computed to predict the positioning uncertainty as function of the size of the explored area, the number of landmarks, the number of robots, and the accuracy of the onboard sensors (Mourikis and Roumeliotis, 2006). Those theoretical results can be of great use in the design of a C-SLAM system.

5.2. Smoothing-Based Estimation

Besides the linearization error, another drawback of filtering techniques is that the marginalization of past pose variables leads to many new links among the remaining variables. Indeed, the elimination of each pose variable leads to interdependence between every landmark variables to which it was connected. As a result, the variables become increasingly coupled and this leads to more computation. This problem also affects smoothing approaches, but a clever ordering during variable elimination can significantly reduce its impact on performance (Dellaert and Kaess, 2006). Moreover, in smoothing, there is less marginalization required which means that the variables will stay sparsely connected. This sparsity is exploited by modern solvers to yield significant speed-ups (Strasdat et al., 2012). In addition, unlike filtering-based approaches, smoothing techniques improve their accuracy by revisiting past measurements instead of only working from the latest estimate. Hence, filtering techniques fell out of favor due to the better performance of smoothing both in terms of accuracy and efficiency. Moreover, in the context of C-SLAM, the sparsity reduces the amount of data to be exchanged during the estimation process (Paull et al., 2015).

In order to formalize the estimation problem solved by C-SLAM back-ends, we present a general smoothing formulation for pose-graph C-SLAM with two robots (α, β) in which the map landmarks are marginalized into odometry and loop closure measurements. The robots poses and measurements are elements of the special Euclidean manifold $\mathrm{SE}(d)$ where d is the dimension of the problem (i.e., 2 or 3) (Dellaert, 2021).

First, assuming that the measurements noises are uncorrelated, we can factorize eq. 4 as follows:

$$
\begin{aligned}
(X_\alpha^*, X_\beta^*) &\doteq \operatorname*{argmax}_{X_\alpha, X_\beta} p(Z_\alpha, Z_\beta, Z_{\alpha\beta} | X_\alpha, X_\beta) \\
&\doteq \operatorname*{argmax}_{X_\alpha, X_\beta} \left(\prod_{i=1}^{l} p(z_\alpha^i | X_\alpha^i) \prod_{j=1}^{m} p(z_\beta^j | X_\beta^j) \right. \\
&\qquad\qquad \left. \prod_{k=1}^{n} p(z_{\alpha\beta}^k | X_\alpha^k, X_\beta^k) \right)
\end{aligned}
\tag{5}
$$

where $p(z_\alpha^i|X_\alpha^i)$ is the likelihood of the ith measurement of robot α (i.e., z_α^i) given the subset of variables X_α^i on which it is dependent, $p(z_\beta^j|X_\beta^j)$ is the likelihood of the jth measurement of robot β (i.e., z_β^j) given the subset of variables X_β^j on which it is dependent, and $p(z_{\alpha\beta}^k|X_\alpha^k, X_\beta^k)$ is the likelihood of the kth inter-robot measurement (i.e., $z_{\alpha\beta}^k$) given the subset of variables X_α^k and X_β^k. There are l measurements related only to state variables from robot α, m measurements related only to state variables from robot β, and n measurements related to state variables from both robots.

Second, assuming that the measurements are disturbed by zero-mean Gaussian noise with information matrix Ω (i.e., inverse of the covariance), we can express the individual measurement likelihood as

$$p(z_\alpha^i|X_\alpha^i) \propto \exp\left(-\frac{1}{2}\left\|h_\alpha^i(X_\alpha^i) - z_\alpha^i\right\|_{\Omega_\alpha^i}^2\right) \tag{6}$$

where h_α^i is a function that maps the state variables to the measurements.

Third, since maximizing the likelihood is equivalent to minimizing the negative log-likelihood, we obtain the following nonlinear least squares formulation of problem 4:

$$
\begin{aligned}
(X_\alpha^*, X_\beta^*) &\doteq \underset{X_\alpha, X_\beta}{\operatorname{argmin}} -\log\Bigg(\prod_{i=1}^{l} p(z_\alpha^i|X_\alpha^i) \prod_{j=1}^{m} p(z_\beta^j|X_\beta^j) \\
&\qquad\qquad \prod_{k=1}^{n} p(z_{\alpha\beta}^k|X_\alpha^k, X_\beta^k)\Bigg) \\
&\doteq \underset{X_\alpha, X_\beta}{\operatorname{argmin}} \Bigg(\sum_{i=1}^{l}\left\|h_\alpha^i(X_\alpha^i) - z_\alpha^i\right\|_{\Omega_\alpha^i}^2 + \sum_{j=1}^{m}\left\|h_\beta^j(X_\beta^j) - z_\beta^j\right\|_{\Omega_\beta^j}^2 \\
&\qquad\qquad + \sum_{k=1}^{n}\left\|h_{\alpha\beta}^k(X_\alpha^k, X_\beta^k) - z_{\alpha\beta}^k\right\|_{\Omega_{\alpha\beta}^k}^2\Bigg)
\end{aligned}
\tag{7}
$$

This nonlinear least squares problem can be solved either on a single computer or in a distributed fashion. In the centralized case, one can simply use single-robot pose graph optimization solvers (Agarwal et al., ; Kümmerle et al., 2011; F. Dellaert et al., ; Rosen et al., 2019). Incremental single-robot solvers (Kaess et al., 2011) can also be adapted for the centralized C-SLAM problem to continuously update the global pose graph with the latest measurements from the robots (Dong et al., 2015). Recently, a client-server architecture has been proposed in which resource-limited clients (e.g., robots or mobile phones) only optimize small parts of the map while the server processes the rest (Zhang et al., 2021b). This centralized and distributed approach allows for accurate real-time state estimation even with limited computation and memory capacity onboard the clients.

Among the distributed solvers, many early techniques used Gaussian elimination (Cunningham et al., 2010; Cunningham et al., 2013; Cunningham et al., 2012). Although popular, those approaches require the exchange of dense marginals which means that the communication cost is quadratic on the number of inter-robot measurements. Furthermore, those approaches rely on linearization, so they require complex bookkeeping to ensure the consistency at the linearization point within the team of robots. To reduce the complexity, (Nerurkar et al., 2009) introduces a distributed marginalization scheme to limit the size of the optimization problem.

More recently, the approach in (Choudhary et al., 2017a) leverages the Distributed Gauss-Seidel algorithm introduced in (Bertsekas and Tsitsiklis, 1989) to solve eq. 7. This technique avoids complex bookkeeping and information double-counting in addition of satisfying privacy constraints by exchanging minimal information on the robot individual trajectories. In another line of work, (Zhang et al., 2021a) extends a single-robot incremental solver (Kaess et al., 2011) towards distributed multi-robot setups. This is useful in online missions as it can update the current estimate based on the latest observations without recomputing the whole problem.

Optimization on Riemannian manifolds (Boumal, 2020) has also been considered extensively to solve the C-SLAM problem (Knuth and Barooah, 2012; Knuth and Barooah, 2013). Approaches in (Tron and Vidal, 2009; Tron and Vidal, 2014; Tron et al., 2016) introduce a multi-stage distributed Riemannian consensus protocol with convergence guarantees to globally optimal solutions in noiseless scenarios. Expanding on those ideas, a recent technique (Tian et al., 2021b), based upon a sparse semidefinite relaxation, provides exactness guarantees even in the presence of moderate measurement noise. Moreover, this latter technique has been extended to consider asynchronous scenarios and parallel computation (Tian et al., 2020b), which are often critical to deal with communication delays inherent to multi-robot systems.

5.3. Other Estimation Techniques

Other estimation techniques have been proposed for C-SLAM. Among them, the distributed Jacobi approach has been shown to work for 2D poses (Aragues et al., 2011). (Franceschelli and Gasparri, 2010; Aragues et al., 2012) look into consensus-based algorithms and prove their convergence across teams of robots. Also, apart from the solver itself, researchers have studied which measurement and noise models are the best suited for C-SLAM (Indelman et al., 2012).

We observe that more exciting new directions are still being discovered, considering that recent approaches such as (Tian et al., 2021b) have been shown to outperform, both in accuracy and convergence rate, the well established Distributed Gauss-Seidel pose graph optimization method (Choudhary et al., 2017a) reused in many state-of-the-art C-SLAM systems such as (Cieslewski et al., 2018; Lajoie et al., 2020; Wang et al., 2019). Those promising approaches also include the majorization-minimization technique from (Fan and Murphey, 2020), the consensus-based 3D pose estimation technique inspired by distributed formation control from (Cristofalo et al., 2019; Cristofalo et al., 2020), and (Zhu et al., 2021) distributed estimator based on covariance intersection.

5.4. Perceptual Aliasing Mitigation

As it is the case in single robot SLAM, loop closure detection is vulnerable to spurious measurements, i.e., outliers, due to perceptual aliasing. This phenomenon occurs when two different places are conflated as the same during the place recognition process. This motivates the need for robust techniques that can detect and remove those outliers to avoid dramatic distortions in the C-SLAM estimates. A common approach is to adopt a robust objective function less sensitive to outliers (Sünderhauf and Protzel, 2012; Agarwal et al., 2013; Latif et al., 2013; Lajoie et al., 2019; Yang et al., 2020). Outliers mitigation might also help against adversarial attacks by rejecting spurious measurements injected by a nefarious agent.

The classic approach to remove outliers is to use the RANSAC algorithm (Fischler and Bolles, 1981) to find a set of mutually consistent measurements (Dong et al., 2015). While RANSAC works well in centralized settings, it is not adapted to distributed systems. Therefore, researchers recently explored other ways of detecting outliers such as leveraging extra information from the wireless communication channels during a rendezvous between two robots (Wang et al., 2019). Since such approaches work only for direct inter-robot loop closures, there is a need for general robust data association in the back-end. To that end, (Indelman et al., 2014) uses expectation maximization to infer which inter-robot measurements are inliers and which ones are outliers. One currently popular approach in C-SLAM is the use of pairwise consistency maximization to search for the maximal clique of pairwise consistent measurements among the inter-robot loop closures (Mangelson et al., 2018). (Lajoie et al., 2020) introduces a distributed implementation of this technique which does not require any additional communication when paired with distributed pose graph optimization, while (Chang et al., 2021) proposes an incremental version, and (Do et al., 2020) extends the pairwise consistency evaluation with a data similarity metric. Another recent work (Tian et al., 2021a) extends to distributed setups the Graduated Non-Convexity approach for outlier rejection previously applied to single-robot SLAM (Yang et al., 2020). It is important to note that those latest

approaches only apply to smoothing-based C-SLAM since, unlike filtering, it allows the removal of past measurements from the estimation.

6. System-Level Challenges

Along with the front-end and back-end specific challenges, some issues and design choices affect the whole C-SLAM system. As in single-robot SLAM, the map representation has strong repercussions on motion tracking, place recognition and state estimation. On top of this, multi-robot systems (described in 3.1) present unique challenges to C-SLAM in terms of communication and coordination.

6.1. Map Representation

When designing large multi-robot systems, the choices of map representation could affect computation load, memory usage, and communication bandwidth. First, it is important to note that an interpretable map is not always required. For example, when the sole objective is collaborative localization, a feature map can be sufficient. In those cases, each robot locally tracks landmarks, or features, and searches for correspondences in other robots' feature maps to obtain indirect inter-robot loop closure measurements. The local feature maps can be merged frequently so that the robots can navigate and track features in a global map, or they can be shared on demand upon place recognition events. This way, the robots can operate in the same reference frame without the computation and communication burden of building an interpretable map model.

When required, the chosen map representation depends on the mission objective and environment. For example, in the case of ground robots in flat indoor environments, a 2D map might be sufficient (Caccavale and Schwager, 2018). In those scenarios, occupancy grid maps have been shown to be a compact and more accurate solution (Martin and Emami, 2010; Saeedi et al., 2011a) than feature-based maps (Benedettelli et al., 2010). However, 3D representations are sometimes required (e.g., for rough terrain navigation) at the cost of more computation, storage, and communication, which can be difficult to handle when resources are limited on the robots. Given the communication constraints in C-SLAM systems, compact or sparse representations, such as topological maps (H. Jacky Chang et al., 2007; Saeedi et al., 2014), are often preferred. In the same vein, some works aim for semantic-based representations in the form of sparse maps of labelled regions (Choudhary et al., 2017b). Map representations can also affect long-term operations due to the increasing size of the map in memory (Zhang et al., 2018a), which is also a challenge in single-robot SLAM.

6.2. Efficient and Robust Communication

One of the core implementation differences between SLAM and C-SLAM is the need for communication and coordination within the robotic team. For efficiency, the required bandwidth needs to be minimal, and the communication network needs to be robust to robot failures and varying topologies.

The communication bottleneck of a C-SLAM system is generally caused by the exchanges of sensor data or representations used to compute inter-robot loop closures (Tardioli et al., 2015). Robots need to share enough data to detect whether other robots have visited the same area, and then estimate a map alignment using any overlapping sections of their maps. Hence, contributions to the front-end of C-SLAM systems often consist of mechanisms to perform an efficient search for loop closure candidates over a team, considering communication constraints. Conversely, the back-end generally relies on a pose-graph which can be shared without the need for large bandwidth.

6.2.1. Efficient Data Sharing
While some early techniques simply share all the data from one robot to another, new sensors produce increasingly rich and dense data. The days of raw sensor data transmission are over and most current techniques in literature opt for some sort of compression or reduction. Even among the early

techniques (Nettleton et al., 2006), the idea of a communication budget has been explored. More recently, the topic has gathered more attention with new techniques carefully coordinating the exchange of data when two robots meet, accounting for the available communication and computation resources (Giamou et al., 2018; Tian et al., 2018a; Tian et al., 2018b; Tian et al., 2020a). One idea is to compress the generated maps using self-organizing maps obtained through unsupervised learning (Saeedi et al., 2011b; Best and Hollinger, 2020). The use of compact representations has also been explored with high-level semantic features: (Choudhary et al., 2017b) relies on objects as landmarks, needing to communicate only object labels and poses to other robots, and (Ramtoula et al., 2020) presents a compact object-based descriptor relying on the configuration of objects in a scene to perform place recognition. In addition to making representations compact, it is useful to ensure that only helpful information is shared. Hence, (Kepler and Stilwell, 2020) introduces a novelty metric so that only sufficiently novel measurements compared to the existing map are transmitted.

The problem has been extensively studied specifically for visual C-SLAM: (Tardioli et al., 2015) proposes to share visual vocabulary indexes instead of feature descriptors to reduce the required bandwidth. Other approaches focus on scalable team-wide place recognition by assigning each robot with a predetermined range of words from a pretrained visual bag of words (Cieslewski and Scaramuzza, 2017b), or regions of full-image descriptors (Cieslewski and Scaramuzza, 2017a). (Dymczyk et al., 2015; Contreras and Mayol-Cuevas, 2017) remove landmarks that are not necessary for localization, (Opdenbosch and Steinbach, 2019) introduces a new coding to efficiently compress features, and (Dubois et al., 2019) proposes data sharing algorithms specialized for visual inertial C-SLAM.

In some extreme cases, communication is severely limited due to the properties of the transmission medium or the large distance between the robots: (Paull et al., 2014; Paull et al., 2015) explore the special case of underwater operations with low bandwidth acoustic communication, and (Schulz et al., 2019) considers long distance radio modules with very limited bandwidth to build the collaborative map through small incremental updates.

6.2.2. Network Topology

Another important aspect to consider is the network topology. Current techniques either assume full connectivity, multi-hop connectivity or are rendezvous-based. Full connectivity means that each robot can directly communicate with all other robots at any time such as in (Cieslewski and Scaramuzza, 2017a; Cieslewski and Scaramuzza, 2017b). Multi-hop connectivity implies that robots can only share information with their neighbors and it might require multiple neighbor-to-neighbor transmissions to reach all robots (Aragüés et al., 2010; Montijano et al., 2013). Rendezvous-based communication means that the robots share data only when they meet and, therefore, do not require any connectivity maintenance. Rendezvous-based C-SLAM also offers the opportunity to perform direct inter-robot loop closure detection during the encounters (Zhou and Roumeliotis, 2006).

The impact of the network topology is especially important during the inference step because disconnections or multi-hop paths can lead to inconsistencies or synchronization issues. Thus (Leung et al., 2011b; Leung et al., 2012) examine the conditions that allow distributed inference to reach the same result as a centralized equivalent approach. Another approach (Quraishi et al., 2016) leverages the progress in the field of distributed computing to improve the robustness to connectivity losses, while (Tuna et al., 2015) evaluates the use of *Wireless Sensor Network*-based communication which is less reliable and predictable, but offers a flexible architecture with self-organization capabilities.

7. Benchmarking C-SLAM

Despite the tremendous progress in the field during the last decade, C-SLAM techniques face tough challenges in terms of reproducibility and benchmarking. C-SLAM systems involve multiple software modules and lots of different hardware components, making it hard to replicate perfectly. While standardized benchmarking approaches have been emerging for single-robot SLAM (Bujanca et al., 2019), such systematic evaluation techniques are still lacking for C-SLAM.

Moreover, only a few datasets dedicated to C-SLAM exist. (Leung et al., 2011a) consists of 9 monocular camera subdatasets and (Dubois et al., 2020a) is dedicated to stereo-inertial C-SLAM. Therefore, the common approach to evaluate C-SLAM solutions is to split single robot SLAM datasets into multiple parts and to associate each one to a robot. When splitting the dataset, careful attention has to be given to ensure the presence of overlaps between the parts for loop closing. In addition, one should avoid overlaps near the cutting points, where the viewpoint and lighting conditions are exactly the same since they depict the same place viewed by the robot at the same point in time: this kind of overlaps is highly unrealistic in multi-robot operations. One of the most used dataset in the literature is the KITTI self-driving car dataset comprised of lidar and stereo camera data (Geiger et al., 2012). New datasets of interest include KITTI360 (Xie et al., 2016) which adds fish-eye cameras, the very large Pit30M lidar and monocular camera dataset that contains over 30 million frames (Martinez et al., 2020), and the DARPA SubT datasets which come with standardized evaluation tools for SLAM (Rogers et al., 2020; Rogers et al., 2021).

8. Open Problems and Ongoing Trends

This section presents open problems and trending ideas in the research community to improve C-SLAM. These new trends push the boundaries of what C-SLAM can do and offer an exciting view of the field's future.

8.1. Resilient Inter-Robot Communication

Although the limitations of inter-robot communication have been a major concern since the inception of C-SLAM, it is still one of the main open problems in the field. In particular, there is a need for resilient communication strategies, aiming beyond robustness to endure unexpected disruptions and ensure swift recovery (Prorok et al., 2021). Delays and dropouts are inevitable in realistic systems, and their effects are amplified when multiple robots operating simultaneously are flooding the network. Delays and out-of-sequence messages can have dramatic effects on real-time robot control which heavily relies on accurate and up-to-date state estimates from C-SLAM (Bresson et al., 2017), and yet they still have not been thoroughly addressed by the research community. Instead, current approaches focus primarily on minimizing communication, which can be achieved, for example, by posing distributed loop closure detection as an optimization problem subject to a budget constraint on total data transmission (Tian et al., 2018a).

Another open problem inherent to C-SLAM and inter-robot communication is the risk of adversarial attacks. In a future in which robots, such as autonomous cars, collaborate on a large scale, security and data integrity will be one of the major concerns of consumers. In addition to the usual risks of infection and hijacking, byzantine data manipulation could lead to map merging poisoning and intentionally erroneous C-SLAM estimates (Deng et al., 2021). Thus further investigation and more efforts have to be deployed on system security.

An interesting, yet still underdeveloped, trend is to leverage the communication medium for inter-robot measurements. This has been successfully done with UWB (Gentner et al., 2018; Boroson et al., 2020; Cao and Beltrame, 2021) or WiFi (Liu et al., 2020), and could be a promising avenue using multipath analysis with channel estimators in 5G networks (Ge et al., 2021). Future techniques might even sidestep inter-robot data transfer completely by communicating via sensor observations of each other and predetermined cues such as visual tags or behavioral patterns (Kim et al., 2010).

8.2. Managing Limited Computing Resources

Aside from communication, computational constraints are an essential consideration in robotics since robots are usually equipped with limited onboard processing devices. It is particularly important in C-SLAM where multiple sub-processes from sensory analysis to inter-robot communication need to be run simultaneously. Thus, to support the current expansion of C-SLAM capabilities, there is

a constant need for efficiency gains. In fact, computation improvements are often at the forefront of new trends in C-SLAM such as the rise of semantic methods, discussed in Section 8.5, which were enabled by GPU-based deep learning (Krizhevsky et al., 2012). Moreover, as discussed in Section 8.6, many new applications of C-SLAM are designed for even smaller platforms such as mobile phones.

Centralized techniques are a natural solution to limited onboard computation capabilities, and, in that regard, recent research suggest that C-SLAM could efficiently leverage the progress in cloud computing. The connection between the two fields is somewhat intuitive: why perform all the processing on robots with limited resources when we could use powerful remote clusters of servers instead? For example, (Riazuelo et al., 2014) offloads the expensive map optimization and storage to a server in the cloud. (Yun et al., 2017) proposes a cloud robotics framework for C-SLAM based on available commercial platforms. Using a similar approach, (Zhang et al., 2018b) manage to perform C-SLAM with up to 256 robots. This is orders of magnitude more than the current techniques based on onboard computation can achieve.

However, while cloud techniques solve the problem of limited computing power onboard the robots, they still face the issue of limited communication bandwidth which is exacerbated when many robots transmit their data through a single communication link. Hence, instead of using remote servers, other strategies need to be explored. For example, a subset of a team of robots could act as a computing cluster to free other robots from the heavy computation burden (Gouveia et al., 2015). Such moving clusters performing computing closer to the sources of data are in accordance with the edge computing paradigm (Satyanarayanan, 2017) to save bandwidth and reduce response time (Huang et al., 2021).

8.3. Adapting to Dynamic Environments

Another inherent problem in multi-robot system is the presence of moving objects in the environment (e.g., people or vehicles). In this regard, the other moving robots in the team are especially problematic. This is a substantial issue since SLAM techniques rely on the tracking of static landmarks. Attempting to solve this problem, (Lee and Lee, 2009) proposes the simple idea of pointing the cameras towards the ceiling when operating indoors with ground robots so that they cannot see each other. (Zou and Tan, 2013) proposes instead to classify dynamic points using the reprojection error and to keep only the static points for estimation. In a different vein, (Moratuwage et al., 2013; Moratuwage et al., 2014; Battistelli et al., 2017) and more recently (Gao et al., 2020) extend upon the Rao-Blackwellized particle filters framework to track moving features, potentially neighboring robots, and remove them from the estimation process. Those works use *Random-Finite-Sets* which were originally developed for multi-target tracking. This way, they manage to incorporate data association, landmark appearance and disappearance, missed detections, and false alarms in the filtering process. Nevertheless, handling dynamic landmarks remains an open topic given that most current works still rely on static environment assumptions.

8.4. Active C-SLAM

The concept of active C-SLAM comes from the powerful idea that while C-SLAM naturally improves path planning and control, path planning and control can also improve C-SLAM. Interestingly, some of the earliest works in collaborative localization were already leveraging coordination between robots to improve accuracy. Instead of mapping the environment, they relied on subsets of robots, in alternance, to serve as landmarks for the others (Kurazume et al., 1994; Trawny and Barfoot, 2004). In an interesting turn of events, the recent progress in C-SLAM has brought back this active sensing trend to the forefront of research.

In C-SLAM, gains can be made by leveraging the coordination between the mapping robots. Having feedback loops to the C-SLAM algorithm allows path planning optimization for faster coverage and mapping of the environment (Bryson and Sukkarieh, 2007; Bryson and Sukkarieh,

2009). To achieve those goals, (Mahdoui et al., 2020) aims to minimize the global exploration time and the average travelled distance among the robots. Other examples of the coupling between path planning and SLAM include (Trujillo et al., 2018), which shows the advantages of UAVs flying in formation for monocular C-SLAM, and (Pei et al., 2020) which uses deep Q-learning to decide whether a robot should localize the others or continue exploring on its own.

Active C-SLAM can also increase the estimation accuracy. To that end, (Dinnissen et al., 2012) uses reinforcement learning to determine the best moment to merge the local maps, and (Kontitsis et al., 2013) leverages instead the covariance matrix computed by the EKF-based inference engine to select trajectories that reduce the map uncertainty. Similarly, (Atanasov et al., 2015) develop a theoretical approach to design a sensor control policy which minimizes the entropy of the estimation task, while (Chen et al., 2020) proposes to broadcast the weakest nodes in the C-SLAM pose graph topology to actively increase the estimation accuracy.

Those works are most likely the mere beginning of active C-SLAM research given that C-SLAM systems are now being integrated on actual industrial, scientific or consumer robots, opening many possibilities of interaction between C-SLAM and other robotics subsystems.

8.5. Semantic C-SLAM

With the rise of deep learning and its impressive semantic inference capabilities, a lot of interest have been directed towards semantic mapping in which the environment is interpreted using class labels (i.e., person, car, chair, etc.). Representing maps as a collection of objects or semantic classes usually leads to much more compressed representations of the environment (Salas-Moreno et al., 2013), and this can be especially beneficial for C-SLAM. Indeed, fewer landmarks and smaller maps are better suited to tight communication constraints since they reduce the amount of data sharing between the robots.

Semantic segmentation was first applied to C-SLAM in (Wu et al., 2009) which detects blobs of colors as salient landmarks in the robots maps. (Choudhary et al., 2017b) later leverages deep learning-based object detection to perform object-based C-SLAM. However, such object-based techniques rely heavily upon the presence of many objects of the known classes in the environment (i.e., classes in the training data). Thus they do not generalize well to arbitrary settings.

The other current preferred approach for semantic C-SLAM is to annotate maps of the environment with class labels. For example, (Frey et al., 2019; Ramtoula et al., 2020) use constellations of landmarks each comprised of a 3D point cloud, a class label and an appearance descriptor. The relatively small number of semantic landmarks reduces the required inter-robot communication significantly. (Tchuiev and Indelman, 2020) considers the joint estimation of object labels and poses in addition to the robots poses in order to improve both estimates. (Chang et al., 2021) build instead globally consistent local metric maps that are enhanced with local semantic labelling, hence preserving the accuracy of pure geometric C-SLAM approaches while incorporating useful high-level information in the robots individual maps.

The tremendous progress still occurring in the field of deep learning strongly suggests that there is more to come in terms of integration with C-SLAM and enhanced collaborative understanding of the environment.

8.6. Augmented Reality

Apart from the well known UAV or self-driving cars applications, Augmented Reality (AR) is probably one of the biggest field of application of SLAM. Indeed, SLAM makes markerless AR applications possible by building a map of the surrounding environment which is essential to overlay digital interactive augmentations. In other words, SLAM is required to make AR work in environments without motion capture, localization beacons or predetermined markers. In the foreseeable future, AR applications and games will push for multi-agent collaboration and this is where C-SLAM comes into play (Egodagamage and Tuceryan, 2017; Egodagamage and Tuceryan,

2018). To that end, (Morrison et al., 2016) proposes a centralized approach in which virtual elements are shared by all agents, and (Sartipi et al., 2019) introduces a decentralized AR technique with smartphones, making use of the visual and inertial sensors already present in those devices. In a similar vein, (Guo et al., 2018) presents a resource-aware technique capable of trading off accuracy to adjust the computational cost to the available resources on mobile devices.

Some other techniques also look at the tremendous potential of collaborative AR for intuitive human-robot interfaces which is especially complex when the number of agents (i.e., humans or robots) and viewpoints increases. For example, to improve supervised mapping tasks, (Sidaoui et al., 2019) equips a human operator with an AR system to edit and correct the map produced by a robot during a mission. Interestingly, (Yu et al., 2020) goes in the opposite direction: humans equipped with smartphones map an environment and get feedback from a central server to indicate which unscanned areas still need to be explored.

Augmented Reality might soon become the main application of C-SLAM in our daily lives, but there is still a lot of research work ahead to efficiently satisfy its inherent constraints and achieve robust large-scale deployments.

9. Conclusions

In this paper, we presented the core ideas behind Collaborative Simultaneous Localization and Mapping and provided a survey of existing techniques. First, we introduced the basic concepts of a C-SLAM system. We provided explanations and bits of historical context to better understand the astonishing progress recently made in the field. Then, we presented the building blocks of a typical C-SLAM system and the associated techniques in the literature. We also touched upon the difficulties of reproducibility and benchmarking. Afterwards, we explored new trends and challenges in the field that will certainly receive a lot more interests in the future. In summary, we focused on providing a complete overview of the C-SLAM research landscape.

We have shown, through numerous examples, how C-SLAM systems are varied and need to match closely the application requirements: sparse or dense maps, precise or topological localization, the number of robots involved, the networking limitations, etc. We wish for this survey to be a useful tool for C-SLAM practitioners looking for adequate solutions to their specific problems.

Nevertheless, despite the current growing interest for C-SLAM applications, it is still a young topic of research and many fundamental problems have to be resolved before the advance of C-SLAM-based commercial products. In particular, we believe that current systems scale poorly and are often limited to very few robots. So, a lot of work is still required to achieve large teams of robots building maps and localizing themselves collaboratively. We also note the growing interest for semantic C-SLAM to make robotic maps more interpretable and more actionable. Scene understanding techniques in the computer vision field could bring more compact and expressive environment representations into the SLAM system, which potentially increase the map readability while reducing the inter-robot communication burden. Furthermore, the rise of AR, in conjunction with C-SLAM and semantics, will offer incredible opportunities of innovation in the fields of collaborative robotics, mobile sensing, and entertainment.

ORCID

Pierre-Yves Lajoie ⬤ https://orcid.org/0000-0002-7495-8567
Benjamin Ramtoula ⬤ https://orcid.org/0000-0002-9185-7995

Fang Wu ⬤ https://orcid.org/0000-0003-2365-6283
Giovanni Beltrame ⬤ https://orcid.org/0000-0001-9755-8630

References

Agarwal, P., Tipaldi, G. D., Spinello, L., Stachniss, C., and Burgard, W. (2013). Robust map optimization using dynamic covariance scaling. In *2013 IEEE International Conference on Robotics and Automation*, pages 62–69.

Agarwal, S., Mierle, K., and Others. Ceres Solver — A Large Scale Non-linear Optimization Library. http://ceres-solver.org/.

Agha, A., Otsu, K., Morrell, B., Fan, D. D., Thakker, R., Santamaria-Navarro, A., Kim, S.-K., Bouman, A., Lei, X., Edlund, J., Ginting, M. F., Ebadi, K., Anderson, M., Pailevanian, T., Terry, E., Wolf, M., Tagliabue, A., Vaquero, T. S., Palieri, M., Tepsuporn, S., Chang, Y., Kalantari, A., Chavez, F., Lopez, B., Funabiki, N., Miles, G., Touma, T., Buscicchio, A., Tordesillas, J., Alatur, N., Nash, J., Walsh, W., Jung, S., Lee, H., Kanellakis, C., Mayo, J., Harper, S., Kaufmann, M., Dixit, A., Correa, G., Lee, C., Gao, J., Merewether, G., Maldonado-Contreras, J., Salhotra, G., Da Silva, M. S., Ramtoula, B., Kubo, Y., Fakoorian, S., Hatteland, A., Kim, T., Bartlett, T., Stephens, A., Kim, L., Bergh, C., Heiden, E., Lew, T., Cauligi, A., Heywood, T., Kramer, A., Leopold, H. A., Choi, C., Daftry, S., Toupet, O., Wee, I., Thakur, A., Feras, M., Beltrame, G., Nikolakopoulos, G., Shim, D., Carlone, L., and Burdick, J. (2021). NeBula: Quest for Robotic Autonomy in Challenging Environments; TEAM CoSTAR at the DARPA Subterranean Challenge. *Accepted for publication in the Journal of Field Robotics, 2021*.

Andersson, L. A. A. and Nygards, J. (2008). C-SAM: Multi-Robot SLAM using square root information smoothing. In *2008 IEEE International Conference on Robotics and Automation*, pages 2798–2805.

Aragues, R., Carlone, L., Calafiore, G., and Sagues, C. (2011). Multi-agent localization from noisy relative pose measurements. In *2011 IEEE International Conference on Robotics and Automation*, pages 364–369.

Aragues, R., Cortes, J., and Sagues, C. (2012). Distributed Consensus on Robot Networks for Dynamically Merging Feature-Based Maps. *IEEE Transactions on Robotics*, 28(4):840–854.

Aragüés, R., Montijano, E., and Sagüés, C. (2010). Consistent data association in multi-robot systems with limited communications. In *In Robotics: Science and Systems*, pages 97–104.

Arandjelović, R., Gronat, P., Torii, A., Pajdla, T., and Sivic, J. (2018). NetVLAD: CNN Architecture for Weakly Supervised Place Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(6):1437–1451.

Atanasov, N., Ny, J. L., Daniilidis, K., and Pappas, G. J. (2015). Decentralized active information acquisition: Theory and application to multi-robot SLAM. In *2015 IEEE International Conference on Robotics and Automation (ICRA)*, pages 4775–4782.

Bailey, T., Bryson, M., Mu, H., Vial, J., McCalman, L., and Durrant-Whyte, H. (2011). Decentralised cooperative localisation for heterogeneous teams of mobile robots. In *2011 IEEE International Conference on Robotics and Automation*, pages 2859–2865.

Barfoot, T. D. (2017). *State Estimation for Robotics*. Cambridge University Press, Cambridge.

Battistelli, G., Chisci, L., and Laurenzi, A. (2017). Random Set Approach to Distributed Multivehicle SLAM. *IFAC-PapersOnLine*, 50(1):2457–2464.

Benedettelli, D., Garulli, A., and Giannitrapani, A. (2010). Multi-robot SLAM using M-Space feature representation. In *49th IEEE Conference on Decision and Control (CDC)*, pages 3826–3831.

Beni, G. (2004). From swarm intelligence to swarm robotics. In *Proceedings of the 2004 International Conference on Swarm Robotics*, SAB'04, pages 1–9, Berlin, Heidelberg. Springer-Verlag.

Bertsekas, D. and Tsitsiklis, J. (1989). *Parallel and Distributed Computation*. Englewood Cliffs, NJ: Prentice-Hall.

Besl, P. J. and McKay, N. D. (1992). A method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):239–256.

Best, G. and Hollinger, G. (2020). Decentralised Self-Organising Maps for Multi-Robot Information Gathering. *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, page 8.

Bezouska, W. and Barnhart, D. (2019). Decentralized Cooperative Localization with Relative Pose Estimation for a Spacecraft Swarm. In *2019 IEEE Aerospace Conference*, pages 1–13.

Bonin-Font, F. and Burguera, A. (2020). Towards Multi-Robot Visual Graph-SLAM for Autonomous Marine Vehicles. *Journal of Marine Science and Engineering*, 8(6):437.

Boroson, E., Hewitt, R., and Ayanian, N. (2020). Inter-Robot Range Measurements in Pose Graph Optimization. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, page 8.

Boroson, E. R. and Ayanian, N. (2019). 3D Keypoint Repeatability for Heterogeneous Multi-Robot SLAM. In *2019 International Conference on Robotics and Automation (ICRA)*, pages 6337–6343.

Boumal, N. (2020). An introduction to optimization on smooth manifolds. Available online.

Brambilla, M., Ferrante, E., Birattari, M., and Dorigo, M. (2013). Swarm robotics: a review from the swarm engineering perspective. *Swarm Intelligence*, 7(1):1–41.

Bresson, G., Alsayed, Z., Yu, L., and Glaser, S. (2017). Simultaneous Localization and Mapping: A Survey of Current Trends in Autonomous Driving. *IEEE Transactions on Intelligent Vehicles*, 2(3):194–220.

Bresson, G., Aufrère, R., and Chapuis, R. (2013). Consistent multi-robot decentralized SLAM with unknown initial positions. In *Proceedings of the 16th International Conference on Information Fusion*, pages 372–379.

Bryson, M. and Sukkarieh, S. (2007). Co-operative Localisation and Mapping for Multiple UAVs in Unknown Environments. In *2007 IEEE Aerospace Conference*, pages 1–12.

Bryson, M. and Sukkarieh, S. (2009). Architectures for Cooperative Airborne Simultaneous Localisation and Mapping. *Journal of Intelligent and Robotic Systems*, 55(4):267–297.

Bujanca, M., Gafton, P., Saeedi, S., Nisbet, A., Bodin, B., O'Boyle, M. F. P., Davison, A. J., Kelly, P. H. J., Riley, G., Lennox, B., Luján, M., and Furber, S. (2019). SLAMBench 3.0: Systematic Automated Reproducible Evaluation of SLAM Systems for Robot Vision Challenges and Scene Understanding. In *2019 International Conference on Robotics and Automation (ICRA)*, pages 6351–6358.

Caccavale, A. and Schwager, M. (2018). Wireframe mapping for resource-constrained robots. In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 1–9.

Cadena, C., Carlone, L., Carrillo, H., Latif, Y., Scaramuzza, D., Neira, J., Reid, I., and Leonard, J. J. (2016). Past, Present, and Future of Simultaneous Localization and Mapping: Toward the Robust-Perception Age. *IEEE Transactions on Robotics*, 32(6):1309–1332.

Cao, Y. and Beltrame, G. (2021). VIR-SLAM: Visual, inertial, and ranging SLAM for single and multi-robot systems. *Autonomous Robots*, 45(6):905–917.

Carlone, L., Kaouk Ng, M., Du, J., Bona, B., and Indri, M. (2011). Simultaneous Localization and Mapping Using Rao-Blackwellized Particle Filters in Multi Robot Systems. *Journal of Intelligent & Robotic Systems*, 63(2):283–307.

Chang, Y., Tian, Y., How, J. P., and Carlone, L. (2021). Kimera-Multi: A System for Distributed Multi-Robot Metric-Semantic Simultaneous Localization and Mapping. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pages 11210–11218.

Chen, Y., Zhao, L., Lee, K. M. B., Yoo, C., Huang, S., and Fitch, R. (2020). Broadcast Your Weaknesses: Cooperative Active Pose-Graph SLAM for Multiple Robots. *IEEE Robotics and Automation Letters*, 5(2):2200–2207.

Choi, Y.-W., Kwon, K.-K., Lee, S.-I., Choi, J.-W., and Lee, S.-G. (2014). Multi-robot Mapping Using Omnidirectional-Vision SLAM Based on Fisheye Images. *ETRI Journal*, 36(6):913–923.

Choudhary, S., Carlone, L., Nieto, C., Rogers, J., Christensen, H. I., and Dellaert, F. (2017a). Distributed mapping with privacy and communication constraints: Lightweight algorithms and object-based models. *The International Journal of Robotics Research*, 36(12):1286–1311.

Choudhary, S., Carlone, L., Nieto, C., Rogers, J., Liu, Z., Christensen, H. I., and Dellaert, F. (2017b). Multi Robot Object-Based SLAM. In Kulić, D., Nakamura, Y., Khatib, O., and Venture, G., editors, *2016 International Symposium on Experimental Robotics*, Springer Proceedings in Advanced Robotics, pages 729–741, Cham. Springer International Publishing.

Cieslewski, T., Choudhary, S., and Scaramuzza, D. (2018). Data-Efficient Decentralized Visual SLAM. In *2018 IEEE International Conference on Robotics and Automation (ICRA)*, pages 2466–2473.

Cieslewski, T. and Scaramuzza, D. (2017a). Efficient decentralized visual place recognition from full-image descriptors. In *2017 International Symposium on Multi-Robot and Multi-Agent Systems (MRS)*, pages 78–82.

Cieslewski, T. and Scaramuzza, D. (2017b). Efficient Decentralized Visual Place Recognition Using a Distributed Inverted Index. *IEEE Robotics and Automation Letters*, 2(2):640–647.

Contreras, L. and Mayol-Cuevas, W. (2017). O-POCO: Online point cloud compression mapping for visual odometry and SLAM. In *2017 IEEE International Conference on Robotics and Automation (ICRA)*, pages 4509–4514.

Cristofalo, E., Montijano, E., and Schwager, M. (2019). Consensus-based Distributed 3D Pose Estimation with Noisy Relative Measurements. In *2019 IEEE 58th Conference on Decision and Control (CDC)*, pages 2646–2653, Nice, France. IEEE.

Cristofalo, E., Montijano, E., and Schwager, M. (2020). GeoD: Consensus-based Geodesic Distributed Pose Graph Optimization. *arXiv:2010.00156 [cs, eess]*.

Cunningham, A., Indelman, V., and Dellaert, F. (2013). DDF-SAM 2.0: Consistent distributed smoothing and mapping. In *2013 IEEE International Conference on Robotics and Automation*, pages 5220–5227.

Cunningham, A., Paluri, M., and Dellaert, F. (2010). DDF-SAM: Fully distributed SLAM using Constrained Factor Graphs. In *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 3025–3030.

Cunningham, A., Wurm, K. M., Burgard, W., and Dellaert, F. (2012). Fully distributed scalable smoothing and mapping with robust multi-robot data association. In *2012 IEEE International Conference on Robotics and Automation*, pages 1093–1100, St Paul, MN, USA. IEEE.

DARPA (2020). DARPA Subterranean Challenge. https://www.subtchallenge.com/.

Dellaert, F. (2021). Factor Graphs: Exploiting Structure in Robotics. *Annual Review of Control, Robotics, and Autonomous Systems*, 4(1):141–166.

Dellaert, F. and Kaess, M. (2006). Square Root SAM: Simultaneous Localization and Mapping via Square Root Information Smoothing. *The International Journal of Robotics Research*, 25(12):1181–1203.

Demim, F., Nemra, A., Louadj, K., Hamerlain, M., and Bazoula, A. (2017). Cooperative SLAM for multiple UGVs navigation using SVSF filter. *Automatika*, 58(1):119–129.

Deng, G., Zhou, Y., Xu, Y., Zhang, T., and Liu, Y. (2021). An Investigation of Byzantine Threats in Multi-Robot Systems. In *24th International Symposium on Research in Attacks, Intrusions and Defenses*, pages 17–32, San Sebastian Spain. ACM.

Deutsch, I., Liu, M., and Siegwart, R. (2016). A framework for multi-robot pose graph SLAM. In *2016 IEEE International Conference on Real-time Computing and Robotics (RCAR)*, pages 567–572.

Dinnissen, P., Givigi, S. N., and Schwartz, H. M. (2012). Map merging of Multi-Robot SLAM using Reinforcement Learning. In *2012 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 53–60.

Do, H., Hong, S., and Kim, J. (2020). Robust Loop Closure Method for Multi-Robot Map Fusion by Integration of Consistency and Data Similarity. *IEEE Robotics and Automation Letters*, 5(4):5701–5708.

Dong, J., Nelson, E., Indelman, V., Michael, N., and Dellaert, F. (2015). Distributed real-time cooperative localization and mapping using an uncertainty-aware expectation maximization approach. In *2015 IEEE International Conference on Robotics and Automation (ICRA)*, pages 5807–5814.

Dörr, S., Barsch, P., Gruhler, M., and Lopez, F. G. (2016). Cooperative longterm SLAM for navigating mobile robots in industrial applications. In *2016 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*, pages 297–303.

Doucet, A., de Freitas, N., Murphy, K. P., and Russell, S. J. (2000). Rao-Blackwellised Particle Filtering for Dynamic Bayesian Networks. In *Proceedings of the 16th Conference on Uncertainty in Artificial Intelligence*, UAI '00, pages 176–183, San Francisco, CA, USA. Morgan Kaufmann Publishers Inc.

Dubé, R., Dugas, D., Stumm, E., Nieto, J., Siegwart, R., and Cadena, C. (2017a). Segmatch: Segment based place recognition in 3d point clouds. In *2017 IEEE International Conference on Robotics and Automation (ICRA)*, pages 5266–5272. IEEE.

Dubé, R., Gawel, A., Sommer, H., Nieto, J., Siegwart, R., and Cadena, C. (2017b). An online multi-robot SLAM system for 3D LiDARs. In *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 1004–1011.

Dubois, R., Eudes, A., and Frémont, V. (2019). On Data Sharing Strategy for Decentralized Collaborative Visual-Inertial Simultaneous Localization And Mapping. In *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 2123–2130.

Dubois, R., Eudes, A., and Frémont, V. (2020a). AirMuseum: A heterogeneous multi-robot dataset for stereo-visual and inertial Simultaneous Localization And Mapping. In *2020 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*, pages 166–172.

Dubois, R., Eudes, A., Moras, J., and Fremont, V. (2020b). Dense Decentralized Multi-Robot SLAM Based on Locally Consistent TSDF Submaps. In *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, page 8.

Dudek, G., Jenkin, M., Milios, E., and Wilkes, D. (1993). A taxonomy for swarm robots. In *Proceedings of 1993 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'93)*, volume 1, pages 441–447. IEEE.

Dymczyk, M., Lynen, S., Cieslewski, T., Bosse, M., Siegwart, R., and Furgale, P. (2015). The gist of maps - summarizing experience for lifelong localization. In *2015 IEEE International Conference on Robotics and Automation (ICRA)*, pages 2767–2773.

Ebadi, K., Chang, Y., Palieri, M., Stephens, A., Hatteland, A., Heiden, E., Thakur, A., Funabiki, N., Morrell, B., Wood, S., Carlone, L., and Agha-mohammadi, A.-a. (2020). LAMP: Large-Scale Autonomous Mapping and Positioning for Exploration of Perceptually-Degraded Subterranean Environments. *arXiv:2003.01744 [cs, eess]*.

Ebadi, K., Palieri, M., Wood, S., Padgett, C., and Agha-mohammadi, A.-a. (2021). DARE-SLAM: Degeneracy-Aware and Resilient Loop Closing in Perceptually-Degraded Environments. *Journal of Intelligent & Robotic Systems*, 102(1):2.

Egodagamage, R. and Tuceryan, M. (2017). A Collaborative Augmented Reality Framework Based on Distributed Visual Slam. In *2017 International Conference on Cyberworlds (CW)*, pages 25–32.

Egodagamage, R. and Tuceryan, M. (2018). Distributed monocular visual SLAM as a basis for a collaborative augmented reality framework. *Computers & Graphics*, 71:113–123.

F. Dellaert et al. Georgia Tech Smoothing And Mapping (GTSAM). http://gtsam.org/.

Fan, T. and Murphey, T. (2020). Majorization Minimization Methods for Distributed Pose Graph Optimization with Convergence Guarantees. In *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 5058–5065, Las Vegas, NV, USA. IEEE.

Fenwick, J. W., Newman, P. M., and Leonard, J. J. (2002). Cooperative concurrent mapping and localization. In *Proceedings 2002 IEEE International Conference on Robotics and Automation (Cat. No.02CH37292)*, volume 2, pages 1810–1817 vol.2.

Fischler, M. A. and Bolles, R. C. (1981). Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6): 381–395.

Forster, C., Lynen, S., Kneip, L., and Scaramuzza, D. (2013). Collaborative monocular SLAM with multiple Micro Aerial Vehicles. In *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 3962–3970.

Fox, D., Burgard, W., Kruppa, H., and Thrun, S. (2000). A Probabilistic Approach to Collaborative Multi-Robot Localization. *Autonomous Robots*, 8(3):325–344.

Franceschelli, M. and Gasparri, A. (2010). On agreement problems with gossip algorithms in absence of common reference frames. In *2010 IEEE International Conference on Robotics and Automation*, pages 4481–4486.

Frey, K. M., Steiner, T. J., and How, J. P. (2019). Efficient Constellation-Based Map-Merging for Semantic SLAM. In *2019 International Conference on Robotics and Automation (ICRA)*, pages 1302–1308.

Galvez-López, D. and Tardos, J. D. (2012). Bags of Binary Words for Fast Place Recognition in Image Sequences. *IEEE Transactions on Robotics*, 28(5):1188–1197.

Gao, L., Battistelli, G., and Chisci, L. (2020). Random-Finite-Set-Based Distributed Multirobot SLAM. *IEEE Transactions on Robotics*, 36(6):1758–1777.

Ge, Y., Jiang, F., Zhu, M., Wen, F., Svensson, L., and Wymeersch, H. (2021). 5G SLAM with Low-complexity Channel Estimation. In *2021 15th European Conference on Antennas and Propagation (EuCAP)*, pages 1–5.

Geiger, A., Lenz, P., and Urtasun, R. (2012). Are we ready for autonomous driving? The KITTI vision benchmark suite. In *2012 IEEE Conference on Computer Vision and Pattern Recognition*, pages 3354–3361, Providence, RI. IEEE.

Gentner, C., Ulmschneider, M., and Jost, T. (2018). Cooperative simultaneous localization and mapping for pedestrians using low-cost ultra-wideband system and gyroscope. In *2018 IEEE/ION Position, Location and Navigation Symposium (PLANS)*, pages 1197–1205.

Giamou, M., Khosoussi, K., and How, J. P. (2018). Talk Resource-Efficiently to Me: Optimal Communication Planning for Distributed Loop Closure Detection. In *2018 IEEE International Conference on Robotics and Automation (ICRA)*, pages 3841–3848.

Gil, A., Reinoso, Ó., Ballesta, M., and Juliá, M. (2010). Multi-robot visual SLAM using a Rao-Blackwellized particle filter. *Robotics and Autonomous Systems*, 58(1):68–80.

Gouveia, B. D., Portugal, D., Silva, D. C., and Marques, L. (2015). Computation Sharing in Distributed Robotic Systems: A Case Study on SLAM. *IEEE Transactions on Automation Science and Engineering*, 12(2):410–422.

Guo, C. X., Sartipi, K., DuToit, R. C., Georgiou, G. A., Li, R., O'Leary, J., Nerurkar, E. D., Hesch, J. A., and Roumeliotis, S. I. (2018). Resource-Aware Large-Scale Cooperative Three-Dimensional Mapping Using Multiple Mobile Devices. *IEEE Transactions on Robotics*, 34(5):1349–1369.

Gupta, R. U. and Conrad, J. M. (2019). A Survey on Multi-robot Particle Filter SLAM. In *2019 SoutheastCon*, pages 1–5.

H. Jacky Chang, C. S. George Lee, Y. Charlie Hu, and Yung-Hsiang Lu (2007). Multi-robot SLAM with topological/metric maps. In *2007 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1467–1472.

Hartley, R. and Zisserman, A. (2003). *Multiple View Geometry in Computer Vision.* Cambridge University Press, USA, second edition.

Howard, A. (2006). Multi-robot Simultaneous Localization and Mapping using Particle Filters. *The International Journal of Robotics Research*, 25(12):1243–1256.

Huang, P., Zeng, L., Chen, X., Luo, K., Zhou, Z., and Yu, S. (2021). Edge Robotics: Edge-Computing-Accelerated Multi-Robot Simultaneous Localization and Mapping. *arXiv:2112.13222 [cs]*.

Hudson, N., Talbot, F., Cox, M., Williams, J., Hines, T., Pitt, A., Wood, B., Frousheger, D., Surdo, K. L., Molnar, T., Steindl, R., Wildie, M., Sa, I., Kottege, N., Stepanas, K., Hernandez, E., Catt, G., Docherty, W., Tidd, B., Tam, B., Murrell, S., Bessell, M., Hanson, L., Tychsen-Smith, L., Suzuki, H., Overs, L., Kendoul, F., Wagner, G., Palmer, D., Milani, P., O'Brien, M., Jiang, S., Chen, S., and Arkin, R. C. (2021). Heterogeneous Ground and Air Platforms, Homogeneous Sensing: Team CSIRO Data61's Approach to the DARPA Subterranean Challenge. *Submitted to the DRC Finals Special Issue of the Journal of Field Robotics.*

Indelman, V., Gurfil, P., Rivlin, E., and Rotstein, H. (2012). Graph-based distributed cooperative navigation for a general multi-robot measurement model. *The International Journal of Robotics Research*, 31(9):1057–1080.

Indelman, V., Nelson, E., Michael, N., and Dellaert, F. (2014). Multi-robot pose graph localization and data association from unknown initial relative poses via expectation maximization. In *2014 IEEE International Conference on Robotics and Automation (ICRA)*, pages 593–600.

Jennings, C., Murray, D., and Little, J. J. (1999). Cooperative robot localization with vision-based mapping. In *Proceedings 1999 IEEE International Conference on Robotics and Automation (Cat. No.99CH36288C)*, volume 4, pages 2659–2665 vol.4.

Jiménez, A. C., García-Díaz, V., González-Crespo, R., and Bolaños, S. (2018). Decentralized Online Simultaneous Localization and Mapping for Multi-Agent Systems. *Sensors*, 18(8):2612.

Kaess, M., Johannsson, H., Roberts, R., Ila, V., Leonard, J., and Dellaert, F. (2011). iSAM2: Incremental smoothing and mapping with fluid relinearization and incremental variable reordering. In *2011 IEEE International Conference on Robotics and Automation*, pages 3281–3288.

Karrer, M. and Chli, M. (2018). Towards Globally Consistent Visual-Inertial Collaborative SLAM. In *2018 IEEE International Conference on Robotics and Automation (ICRA)*, pages 3685–3692.

Karrer, M., Schmuck, P., and Chli, M. (2018). CVI-SLAM—Collaborative Visual-Inertial SLAM. *IEEE Robotics and Automation Letters*, 3(4):2762–2769.

Käslin, R., Fankhauser, P., Stumm, E., Taylor, Z., Mueggler, E., Delmerico, J., Scaramuzza, D., Siegwart, R., and Hutter, M. (2016). Collaborative localization of aerial and ground robots through elevation maps. In *2016 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR)*, pages 284–290.

Kegeleirs, M., Grisetti, G., and Birattari, M. (2021). Swarm SLAM: Challenges and Perspectives. *Frontiers in Robotics and AI*, 8.

Kepler, M. and Stilwell, D. (2020). An Approach to Reduce Communication for Multi-Agent Mapping Applications. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*.

Kim, B., Kaess, M., Fletcher, L., Leonard, J., Bachrach, A., Roy, N., and Teller, S. (2010). Multiple relative pose graphs for robust cooperative mapping. In *2010 IEEE International Conference on Robotics and Automation*, pages 3185–3192.

Knuth, J. and Barooah, P. (2012). Collaborative 3D localization of robots from relative pose measurements using gradient descent on manifolds. In *2012 IEEE International Conference on Robotics and Automation*, pages 1101–1106.

Knuth, J. and Barooah, P. (2013). Collaborative localization with heterogeneous inter-robot measurements by Riemannian optimization. In *2013 IEEE International Conference on Robotics and Automation*, pages 1534–1539.

Koch, P. and Lacroix, S. (2016). Managing environment models in multi-robot teams. In *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 5722–5728.

Kontitsis, M., Theodorou, E. A., and Todorov, E. (2013). Multi-robot active SLAM with relative entropy optimization. In *2013 American Control Conference*, pages 2757–2764.

Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. In *Advances in Neural Information Processing Systems*, volume 25. Curran Associates, Inc.

Kshirsagar, J., Shue, S., and Conrad, J. M. (2018). A Survey of Implementation of Multi-Robot Simultaneous Localization and Mapping. In *SoutheastCon 2018*, pages 1–7.

Kümmerle, R., Grisetti, G., Strasdat, H., Konolige, K., and Burgard, W. (2011). G2o: A general framework for graph optimization. In *2011 IEEE International Conference on Robotics and Automation*, pages 3607–3613.

Kurazume, R., Nagata, S., and Hirose, S. (1994). Cooperative positioning with multiple robots. In *Proceedings of the 1994 IEEE International Conference on Robotics and Automation*, pages 1250–1257 vol.2.

Lajoie, P.-Y., Hu, S., Beltrame, G., and Carlone, L. (2019). Modeling Perceptual Aliasing in SLAM via Discrete–Continuous Graphical Models. *IEEE Robotics and Automation Letters*, 4(2):1232–1239.

Lajoie, P.-Y., Ramtoula, B., Chang, Y., Carlone, L., and Beltrame, G. (2020). DOOR-SLAM: Distributed, Online, and Outlier Resilient SLAM for Robotic Teams. *IEEE Robotics and Automation Letters*, 5(2):1656–1663.

Latif, Y., Cadena, C., and Neira, J. (2013). Robust loop closing over time for pose graph SLAM. *The International Journal of Robotics Research*, 32(14):1611–1626.

Lázaro, M. T., Paz, L. M., Piniés, P., Castellanos, J. A., and Grisetti, G. (2013). Multi-robot SLAM using condensed measurements. In *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1069–1076.

Lee, H., Seung-Hwan Lee, Tae-Seok Lee, Doo-Jin Kim, and Lee, B. (2012). A survey of map merging techniques for cooperative-SLAM. In *2012 9th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI)*, pages 285–287.

Lee, H. S. and Lee, K. M. (2009). Multi-robot SLAM using ceiling vision. In *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 912–917.

Lee, S., Kim, H., and Lee, B. (2020). An Efficient Rescue System with Online Multi-Agent SLAM Framework. *Sensors*, 20(1):235.

Leung, K. Y., Halpern, Y., Barfoot, T. D., and Liu, H. H. (2011a). The UTIAS multi-robot cooperative localization and mapping dataset. *The International Journal of Robotics Research*, 30(8):969–974.

Leung, K. Y. K. (2012). *Cooperative Localization and Mapping in Sparsely-Communicating Robot Networks*. PhD thesis, University of Toronto, Toronto, Ontario, Canada.

Leung, K. Y. K., Barfoot, T. D., and Liu, H. H. T. (2011b). Distributed and decentralized cooperative simultaneous localization and mapping for dynamic and sparse robot networks. In *2011 IEEE International Conference on Robotics and Automation*, pages 3841–3847.

Leung, K. Y. K., Barfoot, T. D., and Liu, H. H. T. (2012). Decentralized Cooperative SLAM for Sparsely-Communicating Robot Networks: A Centralized-Equivalent Approach. *Journal of Intelligent & Robotic Systems*, 66(3):321–342.

Li, F., Yang, S., Yi, X., and Yang, X. (2018). CORB-SLAM: A Collaborative Visual SLAM System for Multiple Robots. In Romdhani, I., Shu, L., Takahiro, H., Zhou, Z., Gordon, T., and Zeng, D., editors, *Collaborative Computing: Networking, Applications and Worksharing*, Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, pages 480–490, Cham. Springer International Publishing.

Liu, R., Marakkalage, S. H., Padmal, M., Shaganan, T., Yuen, C., Guan, Y. L., and Tan, U. (2020). Collaborative SLAM Based on WiFi Fingerprint Similarity and Motion Information. *IEEE Internet of Things Journal*, 7(3):1826–1840.

Loianno, G., Thomas, J., and Kumar, V. (2015). Cooperative localization and mapping of MAVs using RGB-D sensors. In *2015 IEEE International Conference on Robotics and Automation (ICRA)*, pages 4021–4028.

Lowry, S., Sünderhauf, N., Newman, P., Leonard, J. J., Cox, D., Corke, P., and Milford, M. J. (2016). Visual Place Recognition: A Survey. *IEEE Transactions on Robotics*, 32(1):1–19.

Luft, L., Schubert, T., I. Roumeliotis, S., and Burgard, W. (2016). Recursive Decentralized Collaborative Localization for Sparsely Communicating Robots. In *Robotics: Science and Systems XII*. Robotics: Science and Systems Foundation.

Madhavan, R., Fregene, K., and Parker, L. E. (2004). Distributed Cooperative Outdoor Multirobot Localization and Mapping. *Autonomous Robots*, 17(1):23–39.

Mahdoui, N., Frémont, V., and Natalizio, E. (2020). Communicating Multi-UAV System for Cooperative SLAM-based Exploration. *Journal of Intelligent & Robotic Systems*, 98(2):325–343.

Mangelson, J. G., Dominic, D., Eustice, R. M., and Vasudevan, R. (2018). Pairwise Consistent Measurement Set Maximization for Robust Multi-Robot Map Merging. In *2018 IEEE International Conference on Robotics and Automation (ICRA)*, pages 2916–2923.

Martin, A. and Emami, M. R. (2010). Just-in-time cooperative simultaneous localization and mapping. In *2010 11th International Conference on Control Automation Robotics Vision*, pages 479–484.

Martinez, J., Doubov, S., Fan, J., and Bã, I. A. (2020). Pit30M: A Benchmark for Global Localization in the Age of Self-Driving Cars. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, page 8.

Mohamed, S. A. S., Haghbayan, M., Westerlund, T., Heikkonen, J., Tenhunen, H., and Plosila, J. (2019). A Survey on Odometry for Autonomous Navigation Systems. *IEEE Access*, 7:97466–97486.

Montijano, E., Aragues, R., and Sagüés, C. (2013). Distributed Data Association in Robotic Networks With Cameras and Limited Communications. *IEEE Transactions on Robotics*, 29(6):1408–1423.

Morales, J. and Kassas, Z. M. (2018). Information fusion strategies for collaborative radio SLAM. In *2018 IEEE/ION Position, Location and Navigation Symposium (PLANS)*, pages 1445–1454.

Moratuwage, D., Vo, B., and Wang, D. (2013). Collaborative Multi-vehicle SLAM with moving object tracking. In *2013 IEEE International Conference on Robotics and Automation*, pages 5702–5708.

Moratuwage, D., Wang, D., Rao, A., Senarathne, N., and Wang, H. (2014). RFS Collaborative Multivehicle SLAM: SLAM in Dynamic High-Clutter Environments. *IEEE Robotics Automation Magazine*, 21(2): 53–59.

Morrison, J. G., Gálvez-López, D., and Sibley, G. (2016). MOARSLAM: Multiple Operator Augmented RSLAM. *Distributed Autonomous Robotic Systems*, pages 119–132.

Mourikis, A. and Roumeliotis, S. (2004). Analysis of positioning uncertainty in reconfigurable networks of heterogeneous mobile robots. In *IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA '04. 2004*, volume 1, pages 572–579 Vol.1.

Mourikis, A. I. and Roumeliotis, S. I. (2006). Predicting the Performance of Cooperative Simultaneous Localization and Mapping (C-SLAM). *The International Journal of Robotics Research*, 25(12):1273–1286.

Mur-Artal, R. and Tardós, J. D. (2017). ORB-SLAM2: an open-source SLAM system for monocular, stereo and RGB-D cameras. *IEEE Transactions on Robotics*, 33(5):1255–1262.

Nerurkar, E. D., Roumeliotis, S. I., and Martinelli, A. (2009). Distributed maximum a posteriori estimation for multi-robot cooperative localization. In *2009 IEEE International Conference on Robotics and Automation*, pages 1402–1409.

Nettleton, E., Thrun, S., Durrant-Whyte, H., and Sukkarieh, S. (2006). Decentralised SLAM with Low-Bandwidth Communication for Teams of Vehicles. In Yuta, S., Asama, H., Prassler, E., Tsubouchi, T., and Thrun, S., editors, *Field and Service Robotics: Recent Advances in Reserch and Applications*, Springer Tracts in Advanced Robotics, pages 179–188. Springer, Berlin, Heidelberg.

Olson, E. B. (2009). Real-time correlative scan matching. In *2009 IEEE International Conference on Robotics and Automation*, pages 4387–4393.

Opdenbosch, D. V. and Steinbach, E. (2019). Collaborative Visual SLAM Using Compressed Feature Exchange. *IEEE Robotics and Automation Letters*, 4(1):57–64.

Özyeşil, O., Voroninski, V., Basri, R., and Singer, A. (2017). A survey of structure from motion. *Acta Numerica*, 26:305–364.

Paull, L., Huang, G., Seto, M., and Leonard, J. J. (2015). Communication-constrained multi-AUV cooperative SLAM. In *2015 IEEE International Conference on Robotics and Automation (ICRA)*, pages 509–516.

Paull, L., Seto, M., and Leonard, J. J. (2014). Decentralized cooperative trajectory estimation for autonomous underwater vehicles. In *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 184–191.

Pei, Z., Piao, S., Quan, M., Qadir, M. Z., and Li, G. (2020). Active collaboration in relative observation for multi-agent visual simultaneous localization and mapping based on Deep Q Network. *International Journal of Advanced Robotic Systems*, 17(2):1729881420920216.

Pfingsthorn, M., Slamet, B., and Visser, A. (2008). A Scalable Hybrid Multi-robot SLAM Method for Highly Detailed Maps. In Visser, U., Ribeiro, F., Ohashi, T., and Dellaert, F., editors, *RoboCup 2007: Robot Soccer World Cup XI*, Lecture Notes in Computer Science, pages 457–464, Berlin, Heidelberg. Springer.

Prorok, A., Malencia, M., Carlone, L., Sukhatme, G. S., Sadler, B. M., and Kumar, V. (2021). Beyond Robustness: A Taxonomy of Approaches towards Resilient Multi-Robot Systems. *arXiv:2109.12343 [cs, eess]*.

Queralta, J. P., Taipalmaa, J., Pullinen, B. C., Sarker, V. K., Gia, T. N., Tenhunen, H., Gabbouj, M., Raitoharju, J., and Westerlund, T. (2020). Collaborative Multi-Robot Search and Rescue: Planning, Coordination, Perception, and Active Vision. *IEEE Access*, 8:191617–191643.

Quraishi, A., Cieslewski, T., Lynen, S., and Siegwart, R. (2016). Robustness to connectivity loss for collaborative mapping. In *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 4580–4585.

Ramtoula, B., de Azambuja, R., and Beltrame, G. (2020). CAPRICORN: Communication Aware Place Recognition using Interpretable Constellations of Objects in Robot Networks. In *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pages 8761–8768.

Rekleitis, I., Dudek, G., and Milios, E. (2003). Probabilistic cooperative localization and mapping in practice. In *2003 IEEE International Conference on Robotics and Automation (Cat. No.03CH37422)*, volume 2, pages 1907–1912 vol.2.

Riazuelo, L., Civera, J., and Montiel, J. M. M. (2014). C2TAM: A Cloud framework for cooperative tracking and mapping. *Robotics and Autonomous Systems*, 62(4):401–413.

Rioux, A., Esteves, C., Hayet, J., and Suleiman, W. (2015). Cooperative SLAM-based object transportation by two humanoid robots in a cluttered environment. In *2015 IEEE-RAS 15th International Conference on Humanoid Robots (Humanoids)*, pages 331–337.

Rogers, J. G., Gregory, J. M., Fink, J., and Stump, E. (2020). Test Your SLAM! The SubT-Tunnel dataset and metric for mapping. In *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pages 955–961.

Rogers, J. G., Schang, A., Nieto-Granda, C., Ware, J., Carter, J., Fink, J., and Stump, E. (2021). The DARPA SubT Urban Circuit Mapping Dataset and Evaluation Metric. In Siciliano, B., Laschi, C., and Khatib, O., editors, *Experimental Robotics*, pages 391–401, Cham. Springer International Publishing.

Rone, W. and Ben-Tzvi, P. (2013). Mapping, localization and motion planning in mobile multi-robotic systems. *Robotica*, 31(1):1–23.

Rosen, D. M., Carlone, L., Bandeira, A. S., and Leonard, J. J. (2019). SE-Sync: A certifiably correct algorithm for synchronization over the special Euclidean group. *The International Journal of Robotics Research*, 38(2-3):95–125.

Rosen, D. M., Doherty, K. J., Terán Espinoza, A., and Leonard, J. J. (2021). Advances in Inference and Representation for Simultaneous Localization and Mapping. *Annual Review of Control, Robotics, and Autonomous Systems*, 4(1):215–242.

Saeedi, S., Paull, L., Trentini, M., and Li, H. (2011a). Multiple robot simultaneous localization and mapping. In *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 853–858.

Saeedi, S., Paull, L., Trentini, M., and Li, H. (2011b). Neural Network-Based Multiple Robot Simultaneous Localization and Mapping. *IEEE Transactions on Neural Networks*, 22(12):2376–2387.

Saeedi, S., Paull, L., Trentini, M., and Li, H. (2015). Occupancy grid map merging for multiple robot simultaneous localization and mapping. *International Journal of Robotics and Automation 2015*, 30(6).

Saeedi, S., Paull, L., Trentini, M., Seto, M., and Li, H. (2014). Group Mapping: A Topological Approach to Map Merging for Multiple Robots. *IEEE Robotics Automation Magazine*, 21(2):60–72.

Saeedi, S., Trentini, M., Seto, M., and Li, H. (2016). Multiple-Robot Simultaneous Localization and Mapping: A Review. *Journal of Field Robotics*, 33(1):3–46.

Salas-Moreno, R. F., Newcombe, R. A., Strasdat, H., Kelly, P. H., and Davison, A. J. (2013). SLAM++: Simultaneous Localisation and Mapping at the Level of Objects. In *2013 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1352–1359, Portland, OR, USA. IEEE.

Sartipi, K., DuToit, R. C., Cobar, C. B., and Roumeliotis, S. I. (2019). Decentralized Visual-Inertial Localization and Mapping on Mobile Devices for Augmented Reality. In *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 2145–2152.

Sasaoka, T., Kimoto, I., Kishimoto, Y., Takaba, K., and Nakashima, H. (2016). Multi-robot SLAM via Information Fusion Extended Kalman Filters. *IFAC-PapersOnLine*, 49(22):303–308.

Satyanarayanan, M. (2017). The Emergence of Edge Computing. *Computer*, 50(1):30–39.

Schmuck, P. and Chli, M. (2017). Multi-UAV collaborative monocular SLAM. In *2017 IEEE International Conference on Robotics and Automation (ICRA)*, pages 3863–3870.

Schmuck, P. and Chli, M. (2019). CCM-SLAM: Robust and efficient centralized collaborative monocular simultaneous localization and mapping for robotic teams. *Journal of Field Robotics*, 36(4):763–781.

Schmuck, P., Ziegler, T., Karrer, M., Perraudin, J., and Chli, M. (2021). COVINS: Visual-Inertial SLAM for Centralized Collaboration. In *2021 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pages 171–176.

Schulz, C., Hanten, R., Reisenauer, M., and Zell, A. (2019). Simultaneous Collaborative Mapping Based on Low-Bandwidth Communication. In *2019 Third IEEE International Conference on Robotic Computing (IRC)*, pages 413–414.

Schuster, M. J., Brand, C., Hirschmüller, H., Suppa, M., and Beetz, M. (2015). Multi-robot 6D graph SLAM connecting decoupled local reference filters. In *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 5093–5100.

Schuster, M. J., Schmid, K., Brand, C., and Beetz, M. (2019). Distributed stereo vision-based 6D localization and mapping for multi-robot teams. *Journal of Field Robotics*, 36(2):305–332.

Sidaoui, A., Elhajj, I. H., and Asmar, D. (2019). Collaborative Human Augmented SLAM. In *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 2131–2138.

Simmons, R. G., Apfelbaum, D., Burgard, W., Fox, D., Moors, M., Thrun, S., and Younes, H. L. S. (2000). Coordination for Multi-Robot Exploration and Mapping. In *Proceedings of the Seventeenth National Conference on Artificial Intelligence and Twelfth Conference on Innovative Applications of Artificial Intelligence*, pages 852–858. AAAI Press.

Strasdat, H., Montiel, J. M. M., and Davison, A. (2012). Visual SLAM: Why filter? *Image Vis. Comput.*

Sünderhauf, N. and Protzel, P. (2012). Switchable constraints for robust pose graph SLAM. In *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1879–1884.

Tardioli, D., Montijano, E., and Mosteo, A. R. (2015). Visual data association in narrow-bandwidth networks. In *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 2572–2577.

Tchuiev, V. and Indelman, V. (2020). Distributed Consistent Multi-Robot Semantic Localization and Mapping. *IEEE Robotics and Automation Letters*, 5(3):4649–4656.

Thrun, S. (2001). A Probabilistic On-Line Mapping Algorithm for Teams of Mobile Robots. *The International Journal of Robotics Research*, 20(5):335–363.

Thrun, S., Burgard, W., and Fox, D. (2005). *Probabilistic Robotic | The MIT Press*. The MIT Press.

Thrun, S. and Liu, Y. (2005). Multi-robot SLAM with Sparse Extended Information Filers. In Dario, P. and Chatila, R., editors, *Robotics Research. The Eleventh International Symposium*, Springer Tracts in Advanced Robotics, pages 254–266, Berlin, Heidelberg. Springer.

Tian, Y., Chang, Y., Arias, F. H., Nieto-Granda, C., How, J. P., and Carlone, L. (2021a). Kimera-Multi: Robust, Distributed, Dense Metric-Semantic SLAM for Multi-Robot Systems. *arXiv:2106.14386 [cs]*.

Tian, Y., Khosoussi, K., Giamou, M., How, J., and Kelly, J. (2018a). Near-Optimal Budgeted Data Exchange for Distributed Loop Closure Detection. In *Robotics: Science and Systems XIV*. Robotics: Science and Systems Foundation.

Tian, Y., Khosoussi, K., and How, J. P. (2018b). Resource-Aware Algorithms for Distributed Loop Closure Detection with Provable Performance Guarantees. In *Algorithmic Foundations of Robotics XIII*, pages 422–438. Springer, Cham.

Tian, Y., Khosoussi, K., and How, J. P. (2020a). A resource-aware approach to collaborative loop-closure detection with provable performance guarantees. *The International Journal of Robotics Research*, page 0278364920948594.

Tian, Y., Khosoussi, K., Rosen, D. M., and How, J. P. (2021b). Distributed Certifiably Correct Pose-Graph Optimization. *IEEE Transactions on Robotics*, pages 1–20.

Tian, Y., Koppel, A., Bedi, A. S., and How, J. P. (2020b). Asynchronous and Parallel Distributed Pose Graph Optimization. *IEEE Robotics and Automation Letters*, 5(4):5819–5826.

Tian, Y., Liu, K., Ok, K., Tran, L., Allen, D., Roy, N., and How, J. P. (2020c). Search and rescue under the forest canopy using multiple UAVs. *The International Journal of Robotics Research*, 39(10-11):1201–1221.

Trawny, N. and Barfoot, T. (2004). Optimized motion strategies for cooperative localization of mobile robots. In *IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA '04. 2004*, volume 1, pages 1027–1032 Vol.1.

Tron, R., Thomas, J., Loianno, G., Daniilidis, K., and Kumar, V. (2016). A Distributed Optimization Framework for Localization and Formation Control: Applications to Vision-Based Measurements. *IEEE Control Systems Magazine*, 36(4):22–44.

Tron, R. and Vidal, R. (2009). Distributed image-based 3-D localization of camera sensor networks. In *Proceedings of the 48h IEEE Conference on Decision and Control (CDC) Held Jointly with 2009 28th Chinese Control Conference*, pages 901–908.

Tron, R. and Vidal, R. (2014). Distributed 3-D Localization of Camera Sensor Networks From 2-D Image Measurements. *IEEE Transactions on Automatic Control*, 59(12):3325–3340.

Trujillo, J.-C., Munguia, R., Guerra, E., and Grau, A. (2018). Cooperative Monocular-Based SLAM for Multi-UAV Systems in GPS-Denied Environments. *Sensors (Basel, Switzerland)*, 18(5).

Tuna, G., Güngör, V. Ç., and Potirakis, S. M. (2015). Wireless sensor network-based communication for cooperative simultaneous localization and mapping. *Computers & Electrical Engineering*, 41:407–425.

Uy, M. A. and Lee, G. H. (2018). Pointnetvlad: Deep point cloud based retrieval for large-scale place recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 4470–4479.

Vitug, E. (2021). Cooperative Autonomous Distributed Robotic Exploration (CADRE). http://www.nasa.gov/directorates/spacetech/game_changing_development/projects/CADRE.

Wang, W., Jadhav, N., Vohs, P., Hughes, N., Mazumder, M., and Gil, S. (2019). Active Rendezvous for Multi-Robot Pose Graph Optimization using Sensing over Wi-Fi. In *International Symposium on Robotics Research (ISRR)*, Hanoi.

Waniek, N., Biedermann, J., and Conradt, J. (2015). Cooperative SLAM on small mobile robots. In *2015 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, pages 1810–1815.

Williams, S. B., Dissanayake, G., and Durrant-Whyte, H. (2002). Towards multi-vehicle simultaneous localisation and mapping. In *Proceedings 2002 IEEE International Conference on Robotics and Automation (Cat. No.02CH37292)*, volume 3, pages 2743–2748 vol.3.

Wu, M., Huang, F., Wang, L., and Sun, J. (2009). Cooperative Multi-Robot Monocular-SLAM Using Salient Landmarks. In *2009 International Asia Conference on Informatics in Control, Automation and Robotics*, pages 151–155.

Xie, J., Kiefel, M., Sun, M., and Geiger, A. (2016). Semantic Instance Annotation of Street Scenes by 3D to 2D Label Transfer. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3688–3697.

Yang, H., Antonante, P., Tzoumas, V., and Carlone, L. (2020). Graduated Non-Convexity for Robust Spatial Perception: From Non-Minimal Solvers to Global Outlier Rejection. *IEEE Robotics and Automation Letters*, 5(2):1127–1134.

Yu, K., Ahn, J., Lee, J., Kim, M., and Han, J. (2020). Collaborative SLAM and AR-guided navigation for floor layout inspection. *The Visual Computer*, 36(10):2051–2063.

Yun, P., Jiao, J., and Liu, M. (2017). Towards a Cloud Robotics Platform for Distributed Visual SLAM. In Liu, M., Chen, H., and Vincze, M., editors, *Computer Vision Systems*, Lecture Notes in Computer Science, pages 3–15, Cham. Springer International Publishing.

Zhang, H., Chen, X., Lu, H., and Xiao, J. (2018a). Distributed and collaborative monocular simultaneous localization and mapping for multi-robot systems in large-scale environments. *International Journal of Advanced Robotic Systems*, 15(3):1729881418780178.

Zhang, P., Wang, H., Ding, B., and Shang, S. (2018b). Cloud-Based Framework for Scalable and Real-Time Multi-Robot SLAM. In *2018 IEEE International Conference on Web Services (ICWS)*, pages 147–154.

Zhang, Y., Hsiao, M., Dong, J., Engel, J., and Dellaert, F. (2021a). MR-iSAM2: Incremental Smoothing and Mapping with Multi-Root Bayes Tree for Multi-Robot SLAM. In *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 8671–8678.

Zhang, Y., Hsiao, M., Zhao, Y., Dong, J., and Engel, J. J. (2021b). Distributed Client-Server Optimization for SLAM with Limited On-Device Resources. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pages 5336–5342.

Zhou, X. S. and Roumeliotis, S. I. (2006). Multi-robot SLAM with Unknown Initial Correspondence: The Robot Rendezvous Case. In *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1785–1792.

Zhu, P., Geneva, P., Ren, W., and Huang, G. (2021). Distributed Visual-Inertial Cooperative Localization. *arXiv:2103.12770 [cs]*.

Zou, D. and Tan, P. (2013). CoSLAM: Collaborative Visual SLAM in Dynamic Environments. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(2):354–366.

Zou, D., Tan, P., and Yu, W. (2019). Collaborative visual SLAM for multiple agents: A brief survey. *Virtual Real. Intell. Hardw.*