

Distributed Optimization with Consensus Constraint for Multi-Robot Semantic Octree Mapping

Arash Asgharivaskasi

Nikolay Atanasov

Abstract—This work develops a distributed optimization algorithm for multi-robot 3-D semantic mapping using streaming range and visual observations and single-hop communication. Our approach relies on gradient-based optimization of the observation log-likelihood of each robot subject to a map consensus constraint to build a common multi-class map of the environment. This formulation leads to closed-form updates which resemble Bayes rule with one-hop prior averaging. To reduce the amount of information exchanged among the robots, we utilize an octree data structure that compresses the multi-class map distribution using adaptive-resolution.

I. INTRODUCTION

Dense semantically-rich environment representations, such as multi-class volumetric maps [1], can be built online by mobile robots thanks to the availability of on-board GPU-accelerated segmentation models. Distributed 3-D mapping requires designing optimization algorithms that take into account the communication and computation limits of multi-robot systems. Among the recent works in distributed mapping for robotics, Corah *et al.* [2] proposed a distributed Gaussian mixture model (GMM) in which a GMM representation is communicated over a network of robots and is utilized by each robot to build a local occupancy grid map for planning. The distributed estimation work by Paritosh *et al.* [3] uses mirror descent optimization of an information-theoretic cost function to obtain the probability density function (PDF) of unknown continuous variables by communicating local estimates among the agents. The metric-semantic mapping approach in Kimera-multi [4] builds a dense multi-class mesh representation from visual-inertial observations and local trajectory communication used to estimate a global reference frame in a decentralized manner.

In this work, we propose a multi-class mapping algorithm for a team of mobile robots. Our main **contribution** is an online distributed multi-robot semantic octree mapping method, where each node in the octree maintains a categorical distribution over various object classes, and is updated using local range and category observations. We approach the multi-robot mapping problem using distributed gradient optimization of the observation log-likelihood, and show that every iteration contains a consensus step where neighbors average their map distributions and a Bayesian update step where new local observations are incorporated into the map

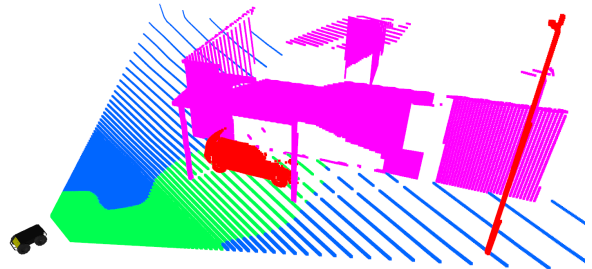


Fig. 1: Semantically annotated point cloud obtained from a range and category observation z_t^i . Each object category is shown with a unique color.

distribution by each robot. Through a simulated experiment, we show that our algorithm leads to globally consistent Bayesian multi-class mapping and the octree data structure helps to achieve memory and communication efficiency.

II. PROBLEM STATEMENT

Consider a network of robots represented by an undirected graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ where \mathcal{V} denotes the set of robots and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ encodes the existence of a communication link between each pair of robots, assumed to be constant throughout the task. The robots are navigating in an environment consisting of a collection of disjoint sets $\mathcal{S}_k \subset \mathbb{R}^3$, each associated with a semantic category $c \in \mathcal{C} := \{0, 1, \dots, C\}$. Let \mathcal{S}_0 denote free space, while each \mathcal{S}_c for $c > 0$ represents a different category, such as building, vegetation, terrain. Each robot i is equipped with a sensor that provides a stream of ray-based observations $z_t^i \in \mathbb{R} \times \mathcal{C}$ at time t . Each observation contains information about the distance to and semantic category of the observed object along the ray. See Fig. 1 for an example. Such information may be obtained by processing the measurements of an RGBD camera or a LiDAR with a semantic segmentation algorithm.

The robots aim to construct a common multi-class map \mathbf{m} of the environment by utilizing their semantically-annotated local range measurements as well as communication with their teammates. We model the map \mathbf{m} as a grid of independent cells, where each individual cell m is labeled with a category in \mathcal{C} . To model measurement noise, we consider a PDF $p(z_t^i | \mathbf{m})$ for each observation. The goal is to compute the map estimate that maximizes the sum of expected log-likelihood of data, i.e. local observations $\{z_\tau^i\}_{\tau=1}^t$ for each robot $i \in \mathcal{V}$:

$$\max_{p \in \mathcal{P}} \sum_{i=1}^N \mathbb{E}_{\mathbf{m} \sim p} \left[\sum_{\tau=1}^t \log p(z_\tau^i | \mathbf{m}) \right],$$

The authors are with the Department of Electrical and Computer Engineering, University of California San Diego, CA 92093, USA {aasghari, natanasov}@eng.ucsd.edu.

We gratefully acknowledge support from ARL DCIST CRA W911NF-17-2-0181 and NSF FRR CAREER 2045945.

where $N = |\mathcal{V}|$ and \mathcal{P} is the space of all probability mass functions (PMF) over the set of possible maps. To remove the constraint $p \in \mathcal{P}$, we utilize a multi-class log-odds ratio representation of the categorical distribution [1]:

$$\mathbf{h} := \left[\log \frac{p(m=0)}{p(m=0)} \quad \dots \quad \log \frac{p(m=C)}{p(m=0)} \right]^\top \in \mathbb{R}^{C+1}.$$

A probability mass function and its log-odds representation have a one-to-one correspondence through the softmax function $\sigma : \mathbb{R}^{C+1} \mapsto \mathbb{R}^{C+1}$:

$$p(m=c) = \sigma_{c+1}(\mathbf{h}) := \frac{\mathbf{e}_{c+1}^\top \exp(\mathbf{h})}{\mathbf{1}^\top \exp(\mathbf{h})},$$

where \mathbf{e}_c is the standard basis vector with c -th element equal to 1 and 0 elsewhere, $\mathbf{1}$ is the vector with all elements equal to 1, and $\exp(\cdot)$ is applied element-wise to the vector \mathbf{h} . Applying Bayes rule to the observation model $p(z_\tau^i | \mathbf{m})$ and using the map cell independence assumption makes it possible to solve the optimization for individual cells instead of maintaining the joint distribution over all cells. Thus, the estimation objective can then be written as an unconstrained optimization for each map cell:

$$\max_{\mathbf{h} \in \mathbb{R}^{C+1}} \sum_{i=1}^N \sum_{m=0}^C \sigma_{m+1}(\mathbf{h}) \log \frac{q_i(m)}{\sigma_{m+1}(\mathbf{h})},$$

where $q_i(m) = \sqrt[t]{\prod_{\tau=1}^t p(m|z_\tau^i)}$ is the temporal geometric average of the inverse observation model. In order to enable distributed optimization of the objective, we introduce a constraint that requires the robots to agree on a common map estimate via only one-hop communication as follows.

Problem. Let $\mathcal{G}(\mathcal{V}, \mathcal{E})$ be a network of N robots, where each robot $i \in \mathcal{V}$ collects observations $\{z_\tau^i\}_{\tau=1}^t$ and maintains its own *local estimate* of the map cell distribution \mathbf{h}_i . Estimate the globally consistent common map such that:

$$\begin{aligned} \max_{\mathbf{h}_{1:N} \in \mathbb{R}^{(C+1) \times N}} & \sum_{i=1}^N \sum_{m=0}^C \sigma_{m+1}(\mathbf{h}_i) \log \frac{q_i(m)}{\sigma_{m+1}(\mathbf{h}_i)}, \\ \text{s.t.} & \sum_{\{i,j\} \in \mathcal{E}} A_{ij} \|\mathbf{h}_j - \mathbf{h}_i\|_2^2 = 0, \end{aligned} \quad (1)$$

where A_{ij} is the ij -th entry of the weighted adjacency matrix of \mathcal{G} with $\sum_j A_{ij} = 1$.

III. DISTRIBUTED SEMANTIC MAPPING

The optimization objective in (1) has a specific structure as the sum of local objective functions over all robots, subject to a consensus constraint among all \mathbf{h}_i , $i \in \mathcal{V}$. We develop an iterative gradient-based method inspired by the general distributed optimization framework presented in [5]. The underlying idea is to interleave local gradient updates for the individual objective function and for the consensus constraint at each robot. Alg. 1 presents the update steps for each map cell in order to solve (1) in a distributed manner. We initialize the log-odds $\mathbf{h}_i^{(0)}$ with prior \mathbf{h}_0 for each robot i . We continuously apply the update steps until a maximum number of iterations is reached, or the update norm is

smaller than a threshold. The update step in line 6 guides the local log-odds towards satisfaction of the consensus constraint, which only requires single-hop communication between neighboring robots $j \in \mathcal{N}_i$. Line 11 incorporates the local observations $\log \mathbf{q}_i = [\log q_i(m)]_{m=0}^C$ via Δ_i , where \odot is the element-wise multiplication. This step is only local to each robot i and does not require communication. Note that both update steps 6 and 13 resemble the log-odds equivalent of the Bayes rule for updating multi-class probabilities (see (8) in [1]). Alg. 1 results in a globally consistent multi-class probabilistic grid map of the environment that is shared amongst all robots in \mathcal{V} . Proofs of convergence to consensus and optimality with respect to the objective function (1) will be shown in future work.

The distributed semantic mapping algorithm we developed assumes a regular grid representation of the environment. To reduce the storage and communication requirements, we may utilize a semantic octree data structure which provides a lossless compression of the original 3-D multi-class map. In this case, the update rules in Alg. 1 should be applied to all leaf nodes in the semantic octree map of each robot i . Please refer to Alg. 3 in [1] for the octree equivalents of the update steps in 6 and 13.

Algorithm 1 Distributed Semantic Mapping

Input: Local observations $\{z_\tau^i\}_{\tau=1}^t$ for each robot $i \in \mathcal{V}$

- 1: $\mathbf{h}_i^{(0)} = \mathbf{h}_0 \quad \forall i \in \mathcal{V} \quad \triangleright$ Initialize all log-odds to \mathbf{h}_0
- 2: $k = 0$
- 3: **loop**
- 4: **for** each robot $i \in \mathcal{V}$ **do**
- 5: \triangleright Enforce consensus with step size ϵ :
- 6: $\tilde{\mathbf{h}}_i^{(k+1)} = \mathbf{h}_i^{(k)} + 2\epsilon \sum_{j \in \mathcal{N}_i} A_{ij} (\mathbf{h}_j^{(k)} - \mathbf{h}_i^{(k)})$
- 7: \triangleright Gradient computation:
- 8: $\Delta_i = \tilde{\mathbf{h}}_i^{(k+1)} - \log \mathbf{q}_i$
- 9: $\alpha_i^{(k+1)} = (\exp(\tilde{\mathbf{h}}_i^{(k+1)})^\top \Delta_i) \mathbf{1}$
- 10: $\beta_i^{(k+1)} = (\exp(\tilde{\mathbf{h}}_i^{(k+1)})^\top \mathbf{1}) \Delta_i$
- 11: $\mathbf{g}_i^{(k+1)} = (\alpha_i^{(k+1)} - \beta_i^{(k+1)}) \odot \frac{\exp(\tilde{\mathbf{h}}_i^{(k+1)})}{(\exp(\tilde{\mathbf{h}}_i^{(k+1)})^\top \mathbf{1})^2}$
- 12: \triangleright Apply gradient with step size $\gamma^{(k+1)}$:
- 13: $\mathbf{h}_i^{(k+1)} = \tilde{\mathbf{h}}_i^{(k+1)} + \gamma^{(k+1)} \mathbf{g}_i^{(k+1)}$
- 14: $h_{i,1}^{(k+1)} = 0 \quad \triangleright h_{i,1}^{(k+1)} = \log \frac{p(m=0)}{p(m=0)} = 0$
- 15: $k = k + 1$
- 16: **return** $\mathbf{h}_i^{(k)} \quad \forall i \in \mathcal{V}$

IV. EVALUATION

We tested our distributed semantic octree mapping algorithm in a 3-D photo-realistic Unity simulation environment, resembling an outdoor village area with various object classes, such as grass, dirt road, building, car, etc. Six ClearPath Husky robots each equipped with an RGB-D camera explore the unknown environment using frontier-based exploration [6]. The RGB sensor measurements are used to perform pixel classification, which combined with the depth input, are provided as local observations to build semantic octree maps. We assume known robot poses and

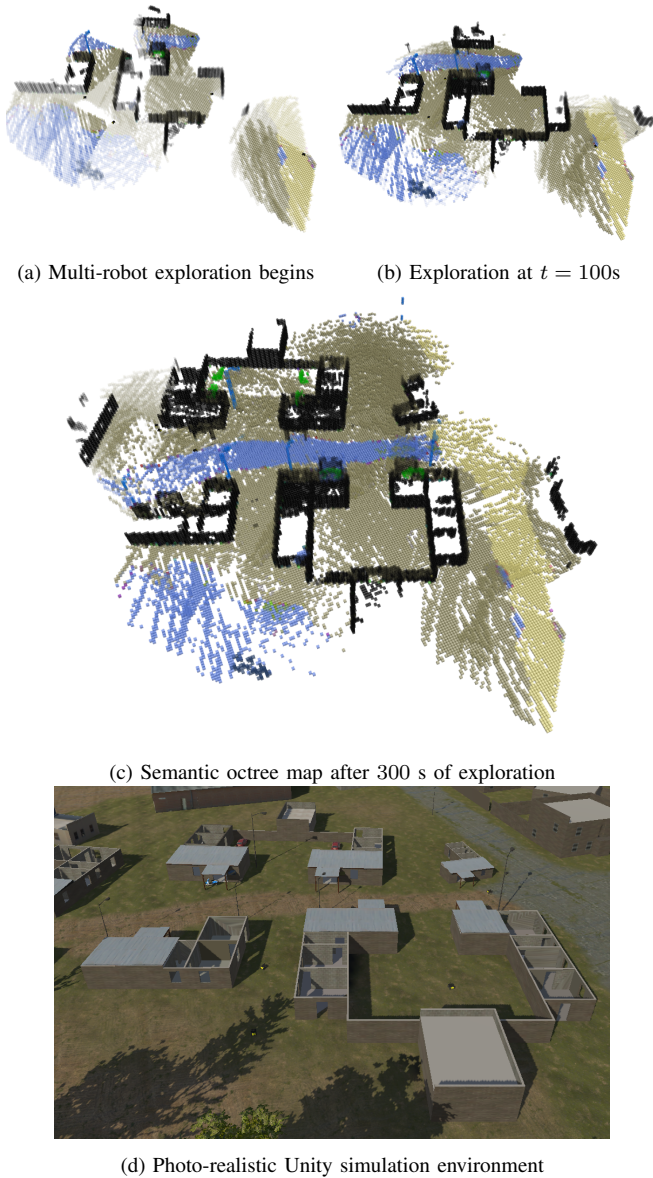


Fig. 2: Time lapse of multi-robot autonomous exploration and semantic octree mapping in the environment shown in (d). Different colors represent different semantic categories (building, dirt road, grass, etc.). Local semantic maps are overlaid with each other such that the transparency of each cell directly relates to the deviation between the local map estimates.

fully connected communication graph \mathcal{G} throughout the experiment. Each robot publishes its local map every 1s, and upon arrival of a map, one iteration of Alg. 1 is executed. Fig. 2 shows several snapshots of the multi-robot semantic octree mapping, as well as the simulation environment. Fig. 3 shows the total distance in local map estimates, i.e. $\sum_{\{i,j\} \in \mathcal{E}} \|\mathbf{h}_j - \mathbf{h}_i\|_2^2$, summed over all semantic octree map leaf nodes. Note that we did not account for the communication strength coefficients $A_{ij} \leq 1$ since doing so would underestimate the actual divergence in the local map estimates. As the figure shows, our distributed mapping algorithm leads to reduction of the total distance in local map estimates. Regarding communication requirements of

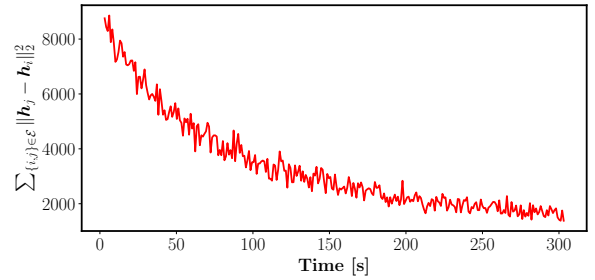


Fig. 3: Time evolution of the total distance in local map estimates. Iteration of Alg. 1 is executed in each robot every one second, upon arrival of a local map estimate from a neighboring robot.

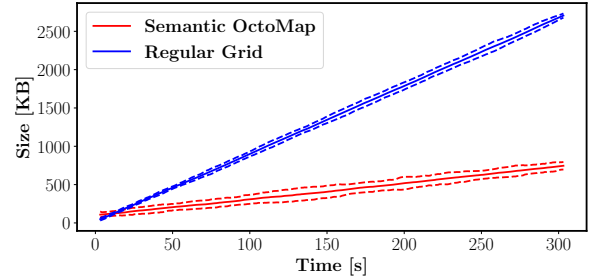


Fig. 4: Packet size used for local map communication for uniform resolution grid map (blue) and semantic octree map (red). The solid lines and the dashed lines show the average over all robots and one standard deviation from the average, respectively.

our distributed semantic octree mapping, Fig. 4 illustrates the size of the communicated local semantic octree maps, with comparison to running the same distributed mapping experiment using uniform resolution grid maps. Our semantic octree map shows more than four times saving in communication bandwidth, empirically validating the benefit of our map representation for distributed multi-robot applications.

In this work, we presented distributed construction of semantic octree maps of unknown environments using a team of robots. The global map can be accessed via query from any of the robots, eliminating the need for a central estimation unit. Such common map can be utilized by robots for decentralized planning. In a future work, we use our multi-robot mapping algorithm alongside a distributed planning method in order to achieve multi-robot exploration.

REFERENCES

- [1] A. Asgharivaskasi and N. Atanasov, “Semantic octree mapping and Shannon mutual information computation for robot exploration,” *IEEE Transactions on Robotics (TRO)*, 2023.
- [2] M. Corah, C. O’Meadhra, K. Goel, and N. Michael, “Communication-efficient planning and mapping for multi-robot exploration in large environments,” *IEEE Robotics and Automation Letters*, vol. 4, no. 2, pp. 1715–1721, 2019.
- [3] P. Paritosh, N. Atanasov, and S. Martinez, “Distributed bayesian estimation of continuous variables over time-varying directed networks,” *IEEE Control Systems Letters*, vol. 6, pp. 2545–2550, 2022.
- [4] Y. Tian, Y. Chang, F. Herrera Arias, C. Nieto-Granda, J. P. How, and L. Carlone, “Kimera-multi: Robust, distributed, dense metric-semantic slam for multi-robot systems,” *IEEE Transactions on Robotics*, vol. 38, no. 4, pp. 2022–2038, 2022.
- [5] S. M. Shah, “Distributed optimization on Riemannian manifolds for multi-agent networks,” *arXiv preprint arXiv:1711.11196*, 2017.
- [6] B. Yamauchi, “A frontier-based approach for autonomous exploration,” in *IEEE International Symposium on Computational Intelligence in Robotics and Automation*, 1997, pp. 146–151.