Active SLAM using 3D Submap Saliency for Underwater Volumetric Exploration

Sudharshan Suresh, Paloma Sodhi, Joshua G. Mangelson, David Wettergreen, and Michael Kaess

Abstract— In this paper, we present an active SLAM framework for volumetric exploration of 3D underwater environments with multibeam sonar. Recent work in integrated SLAM and planning performs localization while maintaining volumetric free-space information. However, an absence of informative loop closures can lead to imperfect maps, and therefore unsafe behavior. To solve this, we propose a navigation policy that reduces vehicle pose uncertainty by balancing between volumetric exploration and revisitation. To identify locations to revisit, we build a 3D visual dictionary from real-world sonar data and compute a metric of submap saliency. Revisit actions are chosen based on propagated pose uncertainty and sensor information gain. Loop closures are integrated as constraints in our pose-graph SLAM formulation and these deform the global occupancy grid map. We evaluate our performance in simulation and real-world experiments, and highlight the advantages over an uncertainty-agnostic framework.

I. INTRODUCTION

The deployment of autonomous underwater vehicles (AUVs) in unstructured 3D environments requires integration of simultaneous localization and mapping (SLAM) with path-planning. Robust autonomy enables inspection of areas of interest—ship-hulls, bridges, reefs, dams, reactor pressure vessels, nuclear pools, and ballast tanks—with minimal human intervention. This is challenging due to the lack of global positioning underwater and failure of visual sensing.

The goal of active SLAM is to generate navigation strategies that complement mapping and localization. If SLAM and path-planning are considered to be mutually agnostic, their errors will affect each other. Uninformative trajectories can deteriorate map accuracy, and imperfect maps can generate unsafe trajectories.

Many state-of-the-art planning algorithms for exploration utilize a grid map representation and guide the robot towards informative viewpoints [1–3]. These methods, however, assume that the robot state estimate is accurate during operation and cannot accommodate map drift. In contrast, SLAM methods that use 3D point clouds localize and correct a global map, but do not preserve free-space information required for planning. Some of these formulations use *submaps*—a collection of smaller, locally consistent 3D point clouds [4–9]. Recently, Ho et al. [9] presented an integrated SLAM approach using a globally deformable grid map, amenable to both localization and planning. Even so, this method is
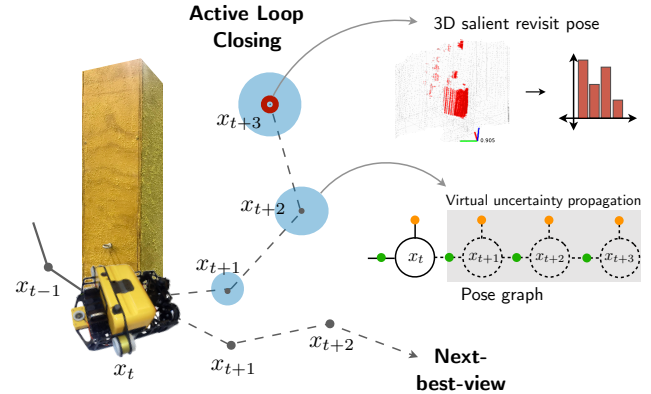
Fig. 1: Proposed active SLAM framework for underwater exploration with sonar sensing. It maintains candidate revisit locations for loop closures using a 3D saliency metric. Our method keeps track of maximum allowable uncertainty, and chooses between next-best-view exploration and revisiting. The vehicle state estimate and free-space is corrected by optimizing an underlying pose-graph.

not considered as active SLAM, which jointly considers the competing objectives of exploration and localization.

Active SLAM was demonstrated in marine environments to construct bathymetric sonar maps while re-localizing drifting vehicle pose [5, 10]. This approach uses particles rather than a pose-graph, and is thus prone to particle depletion. Subsequently, visual saliency was used to choose revisit locations for AUV ship-hull inspection [11, 12]. However, it uses a pre-computed trajectory focused on area coverage of the locally planar portion of the ship hull, a 2D manifold of the 3D exploration problem. In addition, while hull imagery at close standoff distance is feature-rich, vision in general marine environments is infeasible due to poor visibility and a lack of features. These issues are resolved with the use of multibeam acoustic sonar [13, 14], which is unaffected by high turbidity.

In this paper, we present an active SLAM framework for exploring 3D underwater environments with sonar submaps. An overview of our method is shown in Fig. 1. We consider a hovering autonomous underwater vehicle (HAUV) in an environment with proprioceptive sensing and multibeam sonar. Its navigation policy explores free-space with an information gain heuristic, and revisits loop closure candidates based on 3D submap saliency. The results show lower pose uncertainty, accurate dense maps, and reliable free-space information. Specifically, our main contributions are:

(1) A *global submap saliency* (GloSSy) metric to identify good loop closure candidates, coupled with a novel 3D visual dictionary generated from real-world sonar data,

(2) A revisit policy for these candidates based on propagated vehicle uncertainty and sensor information gain, and

(3) An evaluation of the active SLAM framework in simulation and real-world settings.

II. RELATED WORK

A. SLAM and Planning

SLAM is typically handled as a *maximum a posteriori* (MAP) estimate over a factor graph [15]. This is a bipartite graph comprised of variables to be optimized and factors that constrain the system. The variables represent the state we wish to estimate and the factors are the sensor measurements. In submap-based pose graph SLAM systems [6–8] scan-matching is used to constrain the pose of the vehicle.

Bircher et al. [1] used a rapidly-exploring random tree (RRT) planner [16] with the OctoMap representation [17] for a *next-best-view* algorithm for volumetric coverage. Ho et al. [9] combines the aforementioned SLAM and planning modules for a submap pose-graph framework with globally deformable free-space information.

B. Active SLAM

The area of active SLAM dates back to seminal work on active perception by Bajcsy et al. [18]. Sim and Roy [19] proposed that the performance of the SLAM solution is dependent on the choice of the robot trajectory. Underwater applications have focused on revisiting known loop closure candidates to reduce uncertainty in surveying tasks. These methods utilized visual saliency [11, 12] or bathymetric map saliency [20] as a heuristic to prune candidates. However, as detailed in Section I, these sensing modalities do not generalize to many underwater environments. Saliency detection from unorganized 3D point clouds have been presented [21, 22] and applied to LiDAR-based SLAM [23]. We adopt a similar method for sonar submap data, which is a collection of point clouds in local coordinate frames.

III. SUBMAP SALIENCY

A submap is an accumulation of sequential sonar scans over a defined time period, based on vehicle egomotion. It is assumed that within a submap, there is limited local odometry drift. The first scan pose is designated as the base pose, and all subsequent scans are registered to that coordinate frame. We create a 3D scene dictionary offline, and use it to measure the submap saliency of existing base poses. This gives us unique 3D structures in a scene—and empirically—good revisit candidates for scan-matching loop closures. Kim et al. [24] uses this strategy to identify visual candidates for loop closure camera registrations.

The concept of perceptual saliency stems from research on visual attention by Itti and Koch [25]. A bag-of-words (BoW) representation [26] creates a scene dictionary, which can be used for online loop closure detection [27]. While commonly used for visual features, it has also been applied on 3D descriptors for object recognition [28].

A. Vocabulary Generation

We use a 3D BoW representation of an underwater scene, that we generate from real sonar data. We illustrate the process of building this dictionary in Fig. 2. The input submaps are obtained from three datasets collected from teleoperating the HAUV in our underwater tank (Fig. 12). Alternatively, we can use sonar data collected from a pre-planned maneuver in the unknown environment.

For each submap, we obtain its Harris3D keypoints [29]. We extract 32-dimensional local 3D descriptors, known as the signature of histograms of orientations (SHOT) [30]. This jointly encodes the point normals and locations within the local support, and is invariant to rotation and translation. Finally, the SHOT descriptors are clustered into 3D visual words that make up our scene dictionary of size 50.

B. The GloSSy Metric

Global saliency was used for images by Kim et al. [24] as a measure of feature uniqueness across images. It was applied to locate revisit locations in their ship-hull survey trajectory. In place of their SURF descriptors, we use SHOT descriptors to compute our *global submap saliency* (GloSSy) metric. Similar to [24], we define an inter-submap rarity term as a summation of inverse document frequency (idf):

$$\mathcal{G}_s(t) = \sum_{j \in \mathcal{W}_s} \log_2 \frac{N(t)}{n_{w_j}(t)} \tag{1}$$

where $\mathcal{W}_s \subset \mathcal{W}$ is the subset of words found in submap s, $n_{w_j}(t)$ is the number of submaps that contain word w_j, and $N(t)$ is the total number of submaps so far. This is recomputed for all $N(t)$ submaps each time a new submap is received. The idf update has linear complexity in $N(t)$, but can also be implemented as an inverted index update scheme [24]. It is normalized to a GloSSy score $S_{\mathcal{G}_s}(t) \in [0, 1]$:

$$S_{\mathcal{G}_s}(t) = \frac{\mathcal{G}_s(t)}{\mathcal{G}_{\max}}, \quad \text{where} \quad \mathcal{G}_{\max} = \max_{i \in N(t)} \mathcal{G}_i(t) \tag{2}$$

C. Revisit Candidates

Candidate poses with a high GloSSy score are expected to be ideal candidates for loop closures. We empirically verify this by visualizing the highest and lowest GloSSy scores from the training data (Fig. 3). We observe the
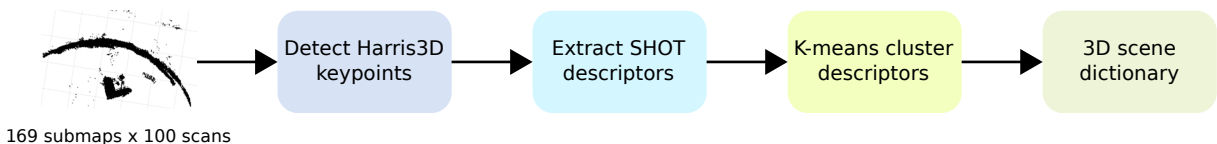

169 submaps x 100 scans

Fig. 2: Process of building a 3D submap scene dictionary offline from real-world sonar data.
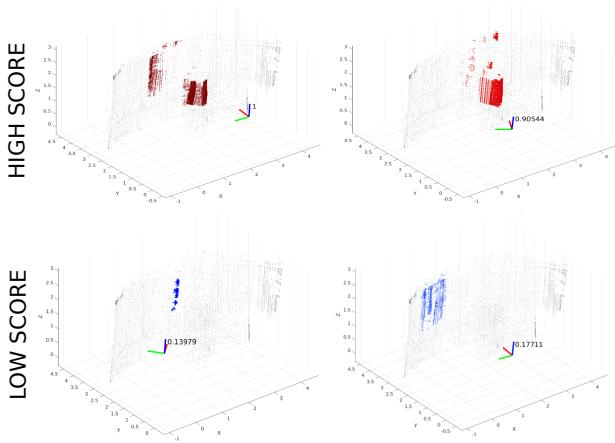
Fig. 3: The revisit poses with highest/lowest GloSSy scores from the collected datalogs. The gray point cloud represents the global map, and the colored sections show the submap that belongs to the revisit pose. They are colored based on the magnitude of the scores. The 6-DoF pose of the vehicle is visualized, along with the GloSSy score. We see a correlation between GloSSy scores and good loop closure candidates.

highest scores are those poses oriented facing a submerged object. These submaps are spatially unique, and can be easily aligned in scan-matching. The low scores comprise of the poses imaging the tank wall's curvature, or the flat side of the piling. These submaps are not spatially distinct in this environment, and can lead to incorrect loop closures. We see a correlation between the GloSSy metric and *good poses to revisit* similar to the visual analogue [24]. We sort them by their $S_{\mathcal{G}_s}(t)$ values and consider the top $N = 3$ candidates.

IV. ACTIVE SLAM FORMULATION

A. Factor Graph Representation

Our pose-graph is similar to that of [8, 9] (Fig. 4). The variable nodes $\mathcal{X} = \{\mathbf{x}_1, \dots \mathbf{x}_N\}$ are the 6-DoF submap base poses, each attributed to a local occupancy map. Upon loop closure, the local maps are merged into a global map with the occupancy probabilities updated according to [17, 31].

AUVs typically have a pressure sensor that directly observes depth (Z), along with an inertial measurement unit (IMU) that observes absolute pitch and roll. The X, Y and yaw directions are obtained through dead reckoning, with drift accumulating over time. Thus, while the local pose estimates are accurate, they drift in the plane over long dives. This divides the vehicle odometry into two factors: **(i)** 3-DoF relative odometry constraint (XYH), and **(ii)** 3-DoF absolute unary constraint (ZPR) [8, 32]. Loop closure constraints are obtained through iterative closest point (ICP) scan-matching of submap point clouds. Finally, we add a pose prior at the first node to bind the trajectory to a global coordinate frame.

Assuming Gaussian noise model, the MAP inference reduces to a nonlinear least-squares optimization problem [15]:

$$
\mathcal{X}^* = \underset{\mathcal{X}}{\operatorname{argmin}} \Big(\sum_{i=1}^{N} (\underbrace{||\mathcal{U}(x_{i-1}, x_i) - u_i||_{\Psi_i}^2}_{\text{XYH factor}} + \underbrace{||\mathcal{V}(x_i) - v_i||_{\Phi_i}^2}_{\text{ZPR factor}})
$$
$$
+ \sum_{(i,k)\in\mathbf{LC}} (\underbrace{||\mathcal{L}(x_i, x_k) - l_{ik}||_{\Gamma_{i,k}}^2}_{\text{loop closure factor}}) + \underbrace{||\mathbf{p}_0 \ominus \mathbf{x}_0||_{\Sigma_0}^2}_{\text{prior factor}}) \Big) \quad (3)
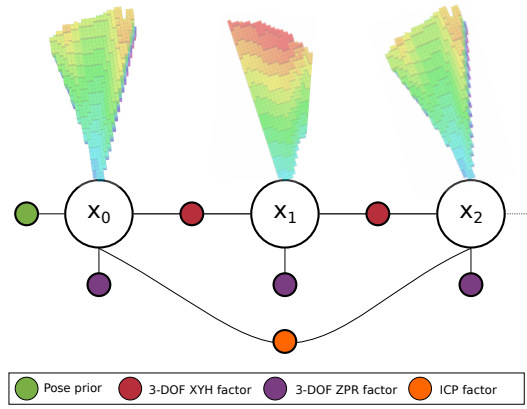$$

Fig. 4: The pose graph formulation we use, based on Teixeira et al. [8]. The nodes x_i represent vehicle pose, with odometry constraints and a pose prior. Loop closure constraints between poses correct for vehicle drift and corresponding local occupancy grid maps are adjusted [9].

The measurement prediction functions for the XYH, ZPR, and loop closure factors are $\mathcal{U}(\cdot)$, $\mathcal{V}(\cdot)$, and $\mathcal{L}(\cdot)$ respectively. Their covariances are Ψ_i, Φ_i, and Γ_{ik} respectively. $\mathbf{LC}$ is the set of all pairs of poses $(\mathbf{x}_i, \mathbf{x}_k)$ for which pairwise loop closure constraints exist. The pose prior $\mathbf{p}_0$ is in the $SE(3)$ Lie group, and $\ominus$ is the logarithm map of the relative transformation between the elements [33]. $\|w\|_\Lambda^2 = w^T \Lambda^{-1} w$ represents the Mahalanobis distance of w. We solve the system in Equation 3 with iSAM [34], through both efficient incremental updates and regular batch updates.

B. Allowable Navigation Uncertainty

We propose a dual-behavior navigation policy based on maximum allowable pose uncertainty, similar to Kim et al. [35]. For a detailed description, we refer the reader to [36]. Revisits are executed when the current state estimate has drifted, reflected in the robot's marginal pose covariance $\Sigma_r(t)$. This is expressed with respect to the initial vehicle pose that is aligned with the world origin. To obtain this, we use an efficient method to recover a part of the full covariance matrix [37]. As shown in Section IV-A, a 3×3 covariance matrix sufficiently captures the robot's drifting state.

We use the D-opt criterion [38] as a mapping from $\Sigma_r(t) \to \mathbb{R}$. Kiefer et al. [39] provides evidence that D-opt is proportional to the uncertainty ellipse of the state parameters. The D-opt criteria ($\mathcal{D}$) has certain desired properties in the context of active SLAM [40]. For our case:

$$
\mathcal{D}(\Sigma_r(t)) = \det(\Sigma_r(t))^{1/3} \quad (4)
$$

Its ratio with a maximum allowable uncertainty Σ_{allow} is:

$$
\mathcal{U}_r(t) = \frac{\mathcal{D}(\Sigma_r(t))}{\mathcal{D}(\Sigma_{\text{allow}})} \quad (5)
$$

C. Policy Algorithm

Algorithm 1 describes the dual-behavior policy of the vehicle. This is called after adding a new submap and optimizing the pose-graph. The GloSSy scores (Section III) are computed in parallel. When the uncertainty ratio exceeds the threshold ($\mathcal{U}_r(t) > 1$) we execute a *revisit policy* in an attempt to re-localize, else we follow the NEXTBESTVIEW

Algorithm 1 : Active SLAM policy algorithm

Require:
    map $\mathcal{M}(t)$, pose $x(t)$, pose-graph $\mathcal{G}(t)$, RRT path $\mathcal{R}(t)$
Ensure:
    vehicle trajectory $\mathcal{T}_{k*}$ **or** $\mathcal{T}_{\text{nbv}}$
    if $\mathcal{U}_r(t) > 1$ **then**
        $x_1^R \ldots x_N^R \leftarrow$ GETSALIENTREVISITPOSES(N)
        for $k \leftarrow 1$ to N **do**
            $\mathcal{T}_k \leftarrow$ GETREVISITTRAJECTORY$(\mathcal{R}(t),\ x(t),\ x_k^R)$
            $\mathcal{U}_k \leftarrow$ PROPAGATEVIRTUALODOM$(\mathcal{G}(t),\ \mathcal{T}_k)$
            Gain$_k \leftarrow$ GETSENSORINFOGAIN$(\mathcal{M}(t),\ \mathcal{T}_k)$
            $\mathcal{P}_k \leftarrow$ GETREVISITUTILITY$(\mathcal{U}_k,\ \text{Gain}_k,\ \alpha)$
        $\mathbf{k}^* \leftarrow$ GETMAX$(\mathcal{P})$
        EXECUTE$(\mathcal{T}_{k*})$
    else
        $\mathcal{T}_{\text{nbv}} \leftarrow$ NEXTBESTVIEW$(\mathcal{M}(t),\ x(t))$
        EXECUTE$(\mathcal{T}_{\text{nbv}})$

Function: GETREVISITTRAJECTORY$(\mathcal{R}(t),\ x(t),\ x_k^R)$

    $n^* \leftarrow$ GETCLOSESTNODE$(\mathcal{R}(t),\ x_k^R)$
    $\mathcal{T}_k \leftarrow$ RETRACETREE$(\mathcal{R}(t),\ x(t),\ n^*)$
    $\mathcal{T}_k \leftarrow$ INTERPOLATE$(\mathcal{T}_k,\ x_k^R,)$

exploration policy [1]. GETSALIENTREVISITPOSES returns the best loop closure candidates as computed in Section III.

For each candidate we call GETREVISITTRAJECTORY from the cached RRT path, a method similar to Stenning et al. [41]. For each revisit candidate we compute the closest node in the tree (GETCLOSESTNODE). We retrace our path down the tree and locally interpolate between the closest point and revisit candidate. Thus, the vehicle follows a similar trajectory and sensor coverage, while avoiding circuitous routes (Fig. 5). We generate an inbound path, and the vehicle subsequently continues exploration.

D. Revisit Utility Function

Kim et al. formulates active SLAM as optimizing a function based-on propagated uncertainty and redundant area coverage [24]. This heuristic, or *utility* function, represents the effect of an action on localization and planning. Our utility function is similar, and considers **(i)** propagated pose covariance at the revisit pose, and **(ii)** sensor information gain from the path. When choosing a revisit policy, we evaluate this function for each candidate. It is defined in Equation 6, and evaluated by GETREVISITUTILITY. We select the revisit candidate k^* that maximizes the function:

$$\mathcal{P}_{\mathbf{k}} = -\Big(\underbrace{\alpha \cdot \mathcal{U}_k}_{\text{uncertainty}} - \underbrace{(1-\alpha) \cdot \text{Gain}_k}_{\text{gain}} \Big) \qquad (6)$$

This is depicted in Fig. 5, with uncertainty ellipses and sensor ray directions. The relative weighting α can be adjusted, leading to a spectrum of behaviors.

1) Uncertainty Term: Given a candidate revisit trajectory $\mathcal{T}_k$, we compute the terminating covariance matrix Σ_k as a measure of uncertainty. This is computed by forward propagating current robot covariance $\Sigma_r(t)$ n steps via *virtual* odometry factors in the pose graph (Fig. 6). Here
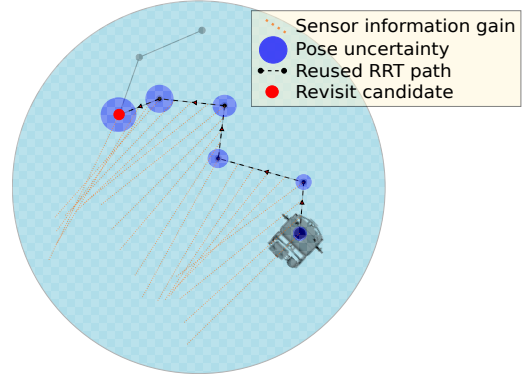

Fig. 5: We reuse the next-best-view RRT path to navigate to a revisit location. To propagate uncertainty, we add virtual odometry along the tree vertices, and compute the pose uncertainty at the candidate. To compute sensor information gain, we interpolate our path and accrue the number of unmapped voxels along the path.
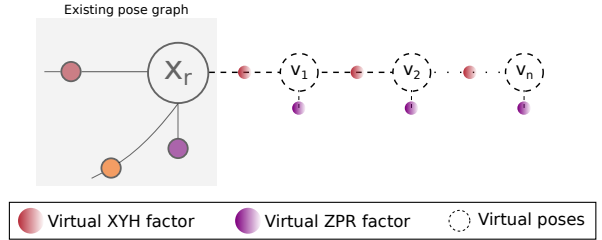

Fig. 6: Addition of virtual nodes to the existing pose graph. Here x_r is the current robot pose, v_i is a virtual pose node with the odometry factors. The graph terminates at the candidate pose v_n.

$v_1 \ldots v_n$ are the RRT path nodes in the revisit path $\mathcal{T}_k$. This considers the worst-case (no loop closures) along the virtual trajectory $\mathcal{T}_k$. We assume a constant velocity model and scale the virtual odometry noise proportional to distance between nodes. Finally, we run a batch optimization, and recover the marginal pose covariance Σ_k [37]. The propagated uncertainty can now be computed and is normalized by $\max(\mathcal{U}_k)$:

$$\mathcal{U}_k = \frac{\mathcal{D}\big(\Sigma_k\big)}{\mathcal{D}\big(\Sigma_{\text{allow}}\big)} \qquad (7)$$

Intuitively, PROPAGATEVIRTUALODOM penalizes revisits to far-off candidates as the drift incurred from that action can make an imperfect state estimate worse.

2) Gain Term: GETSENSORINFOGAIN incentivizes trajectories that increase free-space exploration along our revisit path. This is computed by discretizing $\mathcal{T}_k$ and summing up the sensor information gain according to Bircher et al. [1].

V. RESULTS AND EVALUATION

The framework is evaluated over multiple runs in a simulation and real-world environment. The objective is for the HAUV to explore and map the environment safely, with low pose uncertainty. The metrics we assess are the quality of the global map, pose uncertainty, number of loop closures, and volumetric exploration. The experiments are run on an Intel Core i7-7820HQ CPU @ 2.90GHz with 32GB RAM. We compare our system against a set of baselines:

1. No revisits (Ho et al. [9]),
2. Revisits to a randomized candidate,
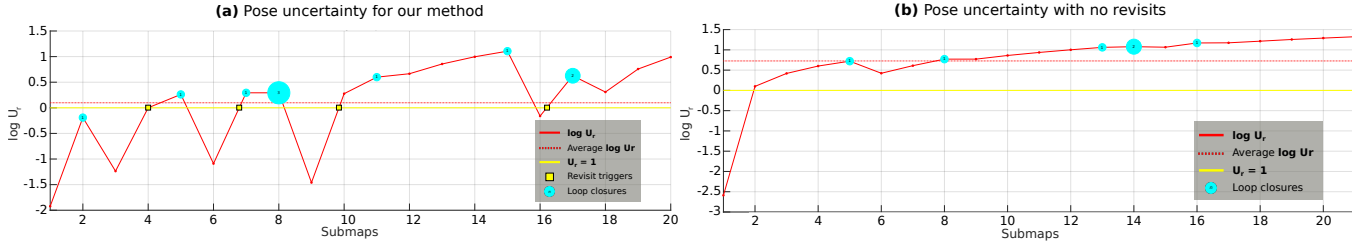3. Dead-reckoning (Bircher et al. [1])

Fig. 7: Log-scale uncertainty ratio vs. submaps for an example run with 20 submaps. The cyan circles denote loop closures and yellow line is the allowable uncertainty threshold. (a) **Our method:** Yellow markers indicate the triggering of a revisit policy. The mean uncertainty ratio lies close to this threshold as a result of informative loop closures. (b) **No revisit method:** Here, the mean uncertainty ratio is moves away due to lack of informative loop closures.
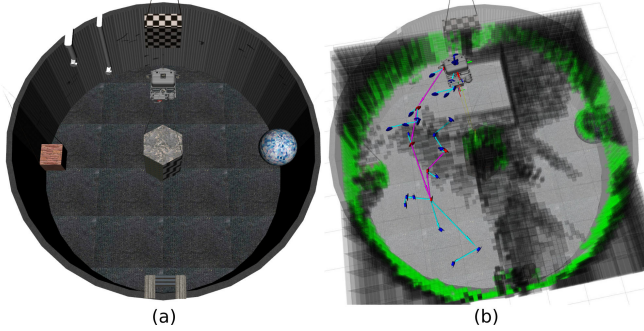

Fig. 8: **(a)** Gazebo simulation with the HAUV. The environment is a metric model of the real-world tank environment (Fig. 12), with targets of different geometries. **(b)** We grow an RRT tree based on sensor information gain, and the Octomap maintains free, unknown, and occupied space.
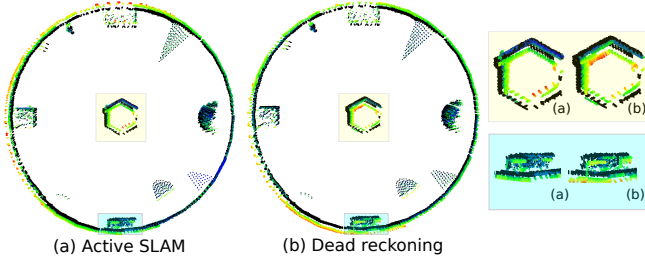

Fig. 9: Ground truth point cloud and resultant map, with heatmap indicating the cloud-to-cloud error. This global map is a collation of 20 submaps in the simulation environment. There is better alignment with our method for structures such as the central piling and the ladder.

A. Experimental Setup

Our simulator must have capabilities for mapping, planning, low-level vehicle control, and state estimation. We use the *UUV Simulator* [42] based on Gazebo and modify the vehicle to emulate the HAUV. In the real-world tests, we use the HAUV from Bluefin Robotics [43] (Fig. 12). Its payload consists of a Doppler velocity log (DVL), attitude and heading reference system (AHRS) and depth sensor, with measurement properties described in Section IV-A.

Here, we get direct readings of Z, roll and yaw (ZPR) from the depth sensor and AHRS. X, Y and yaw (XYH) is indirectly estimated through the accumulation of IMU/DVL odometry. To mimic this behavior, we add Gaussian noise to the relative odometry in XYH, as well as to the absolute ZPR measurements. The noise and covariance parameters are in Table III. Thus, our pose estimate drifts in the plane, similar to cheaper underwater sensing payloads. We treat the uncorrupted odometry as ground truth. The vehicle
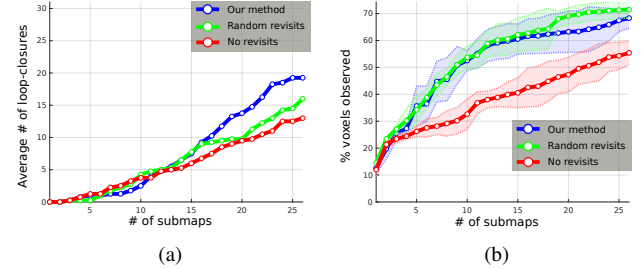

Fig. 10: **(a)** Average number of loop closures over submaps. All three methods show comparable performance < 10 submaps. Beyond that, our method tries to guide the vehicle back to good loop closure locations, and thus surpasses the baselines. **(b)** Average % of observed voxels in the unknown volume over submaps. Our method is able to perform more loop closures without compromising on the number of voxels observed. In the no revisit case, a reason for the lower % could be due to geometry of our sensor. With a line sensor, we miss observing voxels in an initial traverse, which are then observed in a revisit trajectory.

has a profiling sonar with 96 beams and $[29°, 1°]$ FoV. Our simulation approximates the sensor as a 1D laser line scanner, with range readings perturbed by Gaussian noise consistent with that of our real-world sonar ($\sigma_r = 0.01$m).

B. Simulation Environments

We first evaluate our method in our custom simulation environment, a metric mock-up of an underwater tank. Fig. 8 (a) shows the Gazebo model of the tank and vehicle, and Fig. 8 (b) shows the exploration of free-space. As each run is stochastic, we average the results from five runs each. We accumulate 25 submaps (except one with 20), and qualitatively ensure the environment is sufficiently explored. We use the simulation parameters described in Table III.

In Fig. 9, we compare the global map from a single run of our method against the dead-reckoning map. We can see structures are better aligned when incorporating the loop closures. In addition, Fig. 7 (a) and 7 (b) plot the uncertainty ratio $\mathcal{U}_r$ over submaps. We see that in our method the mean uncertainty ratio is close to the allowable threshold, while it is higher without revisitation. This is because we are able

TABLE I: Average cloud-to-cloud error over five runs. From left to right, the final map error decreases and number of loop closures increase.

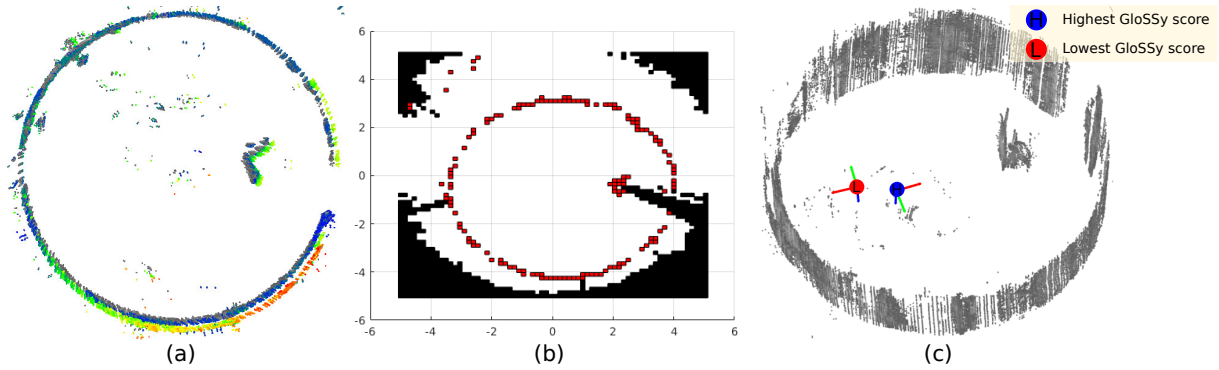| Avg. statistics	Dead-reckoning	No revisits	Random revisits	**Our method**
cloud error (m)	0.0831	0.0807	0.0798	**0.0734**
# loop closures	-	12.6	16.6	**17.4**

Fig. 11: **(a)** Ground truth point cloud with our resultant map, where heatmap indicates the cloud-to-cloud error. **(b)** For the same run, we plot cross-section of the final occupancy grid map *(red: occupied, white: free-space)*. **(c)** Good and bad loop closure candidates overlaid inside the global map.
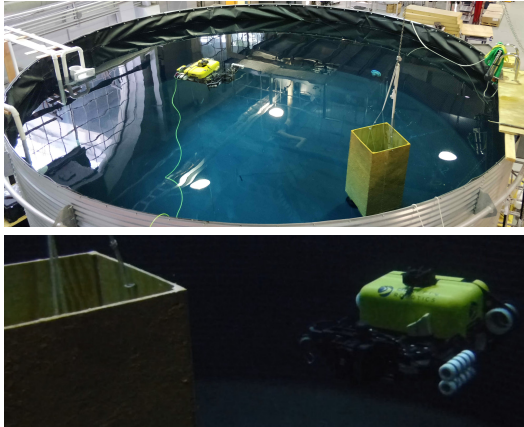

Fig. 12: Real-world tank environment and the HAUV running our method.

TABLE II: Average cloud-to-cloud error over three runs. *Our method gives better quality maps, but has a lesser number of loop closures. This is attributed to incorrect/degenerate matches from drift by the baseline method. In general, the poor quality of loop-closures with real-world data leads to only small improvement in map quality*

Avg. statistics	No revisits	Our method
cloud error (m)	0.0686	0.0628
# loop closures	12.67	8.34

TABLE III: System parameters used for real-world experiments

Parameter	Value	Parameter	Value
Octomap volume	10m x 10m x 3.75m	# Revisit Candidates	3
Initial free-space	3m x 3m x 4m	$\mathcal{U}_r$ threshold	10^{-1}
Voxel resolution	0.15m	α	0.6
RRT extension	1.5m	Virtual velocity	0.5 m/s
Dictionary size	50	$\Psi_i\ (\times 10^{-3})$	$4.14, 4.14, 0.027$
Submap size	100 scans	$\Phi_i\ (\times 10^{-5})$	$1, 0.001, 0.001$

to reduce vehicle drift through loop closures, while Ho et al. [9] performs uncertainty-agnostic exploration.

Table I compares cloud-to-cloud error compared to the three baselines. Our method has least error, with the largest number of loop closures. Random revisits performs better than no revisits, justifying the importance of returning to poses. Next, we evaluate the how each method accumulates loop closure constraints with time in Fig. 10 (a). Our method achieves a higher number of loop closures, while exploring a comparable volume of the environment (Fig. 10 (b)).

C. Real-world Experiments

We carry out real-time experiments with our HAUV in an underwater tank environment. The cylinderical tank has a depth of 3m, radius of 3.5m, and we submerge a 2m×0.6m×0.6m piling structure in it. Fig. 12 shows the HAUV autonomously navigating and mapping the environment. We run three trials each to compare our method with that of no revisits. The runs explore the tank sufficiently, and accumulate 15 to 25 submaps. These trials are autonomous, separate from the teleoperated ones in Section III-A.

Fig. 11 shows a result from one of the runs, and qualitatively highlights all the parts of our system. Fig. 11 (a) compares the final SLAM map with ground truth, where the colormap indicates the magnitude of error. A cross-section of the final 3D grid map (Fig. 11 (b)) illustrates the vehicle has not compromised on volumetric exploration. Finally, Fig. 11 (c) overlays the best and worst GloSSy scores with the global

map—the piling represents an interesting revisit pose for loop closures. In Table II we compare it to our baseline and observe better quality maps. While the no revisits baseline show a greater number of loop closures, the higher error metric is attributed to incorrect/degenerate matches. We note that across real and simulation tests, the lack of features leads to many instances of poor quality, degenerate constraints.

VI. CONCLUSION

We develop an active SLAM framework for volumetric exploration of 3D underwater environments with multibeam sonar. This enables safe long-term autonomy and the resulting dense maps provide valuable information for inspection tasks. We introduce the GloSSy metric for submap saliency, and use it to identify ideal revisit poses for loop closures. We choose from them based on propagated uncertainty and information gain. Finally, the policy is executed by retracing the cached RRT path. We evaluate our method in simulation and real-world, comparing against a few baselines. While demonstrated on the Bluefin HAUV, it can be used on hovering AUVs with similar sensing payloads.

Future work can draw inspiration from analogous LiDAR-based methods for map representation and place recognition [44]. Prior work on visual localization with this payload can be incorporated to provide sensing redundancy [24, 45]. We can also explore opportunistic methods that perform beneficial actions even below the uncertainty threshold [12].

REFERENCES

[1] A. Bircher, M. Kamel, K. Alexis, H. Oleynikova, and R. Siegwart, "Receding horizon "Next-Best-View" planner for 3D exploration," in *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, pp. 1462–1468, 2016.

[2] C. Papachristos, S. Khattak, and K. Alexis, "Uncertainty-aware receding horizon exploration and mapping using aerial robots," in *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, pp. 4568–4575, 2017.

[3] E. Vidal, J. D. Hernández, K. Istenič, and M. Carreras, "Online view planning for inspecting unexplored underwater structures," *IEEE Robotics and Automation Letters*, vol. 2, no. 3, pp. 1436–1443, 2017.

[4] M. Bosse, P. Newman, J. Leonard, and S. Teller, "Simultaneous localization and map building in large-scale cyclic environments using the Atlas framework," *Intl. J. of Robotics Research*, vol. 23, no. 12, pp. 1113–1139, 2004.

[5] N. Fairfield and D. Wettergreen, "Active SLAM and loop prediction with the segmented map using simplified models," in *Field and Service Robotics*, pp. 173–182, 2010.

[6] K. Konolige, E. Marder-Eppstein, and B. Marthi, "Navigation in hybrid metric-topological maps," in *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, pp. 3041–3047, 2011.

[7] R. Wagner, U. Frese, and B. Bäuml, "Graph SLAM with signed distance function maps on a humanoid robot," in *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*, pp. 2691–2698, 2014.

[8] P. V. Teixeira, M. Kaess, F. S. Hover, and J. J. Leonard, "Underwater inspection using sonar-based volumetric submaps," in *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*, pp. 4288–4295, 2016.

[9] B.-J. Ho, P. Sodhi, P. Teixeira, M. Hsiao, T. Kusnur, and M. Kaess, "Virtual occupancy grid map for submap-based pose graph SLAM and planning in 3D environments," in *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*, pp. 2175–2182, 2018.

[10] N. Fairfield and D. Wettergreen, "Active localization on the ocean floor with multibeam sonar," in *Proc. of the IEEE/MTS OCEANS Conf. and Exhibition*, pp. 1–10, 2008.

[11] A. Kim, "Active visual SLAM with exploration for autonomous underwater navigation," tech. rep., Michigan Univ Ann Arbor Dept. of Mechanical Engineering, 2012.

[12] S. M. Chaves, A. Kim, E. Galceran, and R. M. Eustice, "Opportunistic sampling-based active visual SLAM for underwater inspection," *Autonomous Robots*, vol. 40, no. 7, pp. 1245–1265, 2016.

[13] N. Fairfield, G. Kantor, and D. Wettergreen, "Towards particle filter slam with three dimensional evidence grids in a flooded subterranean environment," in *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, pp. 3575–3580, 2006.

[14] S. Barkby, S. B. Williams, O. Pizarro, and M. V. Jakuba, "A featureless approach to efficient bathymetric SLAM using distributed particle mapping," *J. of Field Robotics*, vol. 28, no. 1, pp. 19–39, 2011.

[15] F. Dellaert and M. Kaess, "Factor graphs for robot perception," *Foundations and Trends in Robotics*, vol. 6, no. 1-2, pp. 1–139, 2017.

[16] S. M. LaValle and J. J. Kuffner Jr, "Rapidly-exploring random trees: Progress and prospects," pp. 293–308, 2000.

[17] A. Hornung, K. M. Wurm, M. Bennewitz, C. Stachniss, and W. Burgard, "Octomap: An efficient probabilistic 3D mapping framework based on octrees," *Autonomous robots*, vol. 34, no. 3, pp. 189–206, 2013.

[18] R. Bajcsy, "Active perception," *Proceedings of the IEEE*, vol. 76, no. 8, pp. 966–1005, 1988.

[19] R. Sim and N. Roy, "Global A-optimal robot exploration in SLAM," in *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, pp. 661–666, 2005.

[20] E. Galceran, S. Nagappa, M. Carreras, P. Ridao, and A. Palomer, "Uncertainty-driven survey path planning for bathymetric mapping," in *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*, pp. 6006–6012, 2013.

[21] E. Shtrom, G. Leifman, and A. Tal, "Saliency detection in large point sets," in *Intl. Conf. on Computer Vision (ICCV)*, pp. 3591–3598, 2013.

[22] F. Ponjou Tasse, J. Kosinka, and N. Dodgson, "Cluster-based point set saliency," in *Intl. Conf. on Computer Vision (ICCV)*, pp. 163–171, 2015.

[23] D. Cole, A. Harrison, and P. M. Newman, "Using naturally salient regions for SLAM with 3D laser data," in *International Conference on Robotics and Automation, SLAM Workshop*, 2005.

[24] A. Kim and R. M. Eustice, "Real-time visual SLAM for autonomous underwater hull inspection using visual saliency," *IEEE Trans. Robotics*, vol. 29, no. 3, pp. 719–733, 2013.

[25] L. Itti and C. Koch, "Computational modelling of visual attention," *Nature Reviews Neuroscience*, vol. 2, no. 3, p. 194, 2001.

[26] E. Nowak, F. Jurie, and B. Triggs, "Sampling strategies for bag-of-features image classification," in *Eur. Conf. on Computer Vision (ECCV)*, pp. 490–503, 2006.

[27] M. Cummins and P. Newman, "FAB-MAP: Probabilistic localization and mapping in the space of appearance," *Intl. J. of Robotics Research*, vol. 27, no. 6, pp. 647–665, 2008.

[28] C. Redondo-Cabrera, R. J. López-Sastre, J. Acevedo-Rodriguez, and S. Maldonado-Bascón, "Surfing the point clouds: Selective 3D spatial pyramids for category-level object recognition," in *Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition*, pp. 3458–3465, 2012.

[29] I. Sipiran and B. Bustos, "Harris 3D: a robust extension of the Harris operator for interest point detection on 3D meshes," *The Visual Computer*, vol. 27, no. 11, p. 963, 2011.

[30] F. Tombari, S. Salti, and L. Di Stefano, "Unique signatures of histograms for local surface description," in *Eur. Conf. on Computer Vision (ECCV)*, pp. 356–369, 2010.

[31] H. Moravec and A. Elfes, "High resolution maps from wide angle sonar," in *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, vol. 2, pp. 116–121, 1985.

[32] E. Westman and M. Kaess, "Underwater Apriltag SLAM and calibration for high precision robot localization," tech. rep., Carnegie Mellon University, 2018.

[33] M. Agrawal, "A Lie algebraic approach for consistent pose registration for general Euclidean motion.," *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*, pp. 1891–1897, 2006.

[34] M. Kaess, A. Ranganathan, and F. Dellaert, "iSAM: Incremental smoothing and mapping," *IEEE Trans. Robotics*, vol. 24, no. 6, pp. 1365–1378, 2008.

[35] A. Kim and R. M. Eustice, "Next-best-view visual SLAM for bounded-error area coverage," in *IROS workshop on active semantic perception*, 2012.

[36] S. Suresh, "Localization and active exploration in indoor underwater environments," Master's thesis, 2019.

[37] M. Kaess and F. Dellaert, "Covariance recovery from a square root information matrix for data association," *J. of Robotics and Autonomous Systems*, vol. 57, no. 12, pp. 1198–1210, 2009.

[38] V. V. Fedorov, *Theory of optimal experiments*. Elsevier, 2013.

[39] J. Kiefer, "General equivalence theory for optimum designs (approximate theory)," *The annals of Statistics*, pp. 849–879, 1974.

[40] H. Carrillo, I. Reid, and J. A. Castellanos, "On the comparison of uncertainty criteria for active SLAM," in *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, pp. 2080–2087, 2012.

[41] B. E. Stenning, C. McManus, and T. D. Barfoot, "Planning using a network of reusable paths: A physical embodiment of a rapidly exploring random tree," *J. of Field Robotics*, vol. 30, no. 6, pp. 916–950, 2013.

[42] M. M. M. Manhães, S. A. Scherer, M. Voss, L. R. Douat, and T. Rauschenbach, "UUV simulator: A gazebo-based package for underwater intervention and multi-robot simulation," in *Proc. of the IEEE/MTS OCEANS Conf. and Exhibition*, pp. 1–8, 2016.

[43] J. Vaganay, M. Elkins, D. Esposito, W. O'Halloran, F. Hover, and M. Kokko, "Ship hull inspection with the HAUV: US Navy and NATO demonstrations results," *Proc. of the IEEE/MTS OCEANS Conf. and Exhibition*, pp. 1–6, 2006.

[44] R. Dubé, A. Cramariuc, D. Dugas, J. Nieto, R. Siegwart, and C. Cadena, "SegMap: 3d segment mapping using data-driven descriptors," 2018.

[45] S. Suresh, E. Westman, and M. Kaess, "Through-water stereo SLAM with refraction correction for AUV localization," *IEEE Robotics and Automation Letters*, vol. 4, no. 2, pp. 692–699, 2019.