Direct Ping-Level Landmark Detection for Side-Scan Sonar SLAM in Feature-Poor Seabed Environments

Jinho Im and Seonghun Hong†

***Abstract*—Acoustic perception in underwater environments presents fundamental challenges distinct from those encountered in terrestrial or aerial domains, including GPS-denied navigation, noisy sensor measurements, and sparse visual features in featureless seabeds. Side-scan sonar (SSS) is a particularly attractive sensing modality for such environments, offering wide-area seabed coverage and reliable long-range acoustic measurements. However, most existing SSS-based SLAM approaches operate in the image domain, extracting visual features from stacked acoustic images, and their performance deteriorates in feature-poor or homogeneous seabed environments where distinctive structures are absent. Furthermore, range-dependent intensity variations and speckle noise inherent in SSS measurements further degrade the reliability of feature extraction and data association. This paper proposes a direct ping-level landmark detection method that exploits anomalous deviations in raw backscatter intensity profiles as landmark features, without relying on image formation or appearance-based descriptors. By modeling the slant-range-dependent attenuation of backscatter intensity and applying a spatially adaptive threshold, the proposed method enables reliable landmark extraction directly from individual sonar pings, even in feature-sparse seabed environments where conventional approaches struggle. The proposed detection method is integrated into a landmark-based SSS SLAM framework that directly incorporates raw ping measurements without intermediate image construction. Real-world field experiments demonstrate improved detection precision over image-level and ping-level baselines and enhanced vehicle localization accuracy in feature-poor seabed environments.**

***Index Terms*—Side-scan sonars, Ping-level landmark detection, Feature-poor environments, Underwater SLAM**

I. INTRODUCTION

Underwater navigation and mapping have become increasingly important as unmanned underwater vehicles (UUVs) are deployed for various scientific and engineering missions where human intervention is inefficient or unsafe. However, precise vehicle pose estimation remains challenging due to the unavailability of global positioning system (GPS) under the water surface. Although dead-reckoning (DR) using proprioceptive sensors such as an inertial measurement unit (IMU) and a Doppler velocity log (DVL) serves as the primary navigation method, it inherently suffers from accumulated drift errors over time. Acoustic beacon-based positioning systems [1] can correct such drift, but require pre-deployment and in-situ calibration reducing overall operational efficiency. To overcome the drift accumulation inherent in DR, underwater simultaneous localization and mapping (SLAM) has been introduced as a promising alternative, enabling a vehicle to estimate its pose while constructing a map without prior infrastructure. Among exteroceptive acoustic sensors, side-scan sonar (SSS) is particularly well-suited for this purpose, providing a wide insonified swath and reliable long-range acoustic measurements. SSS generates acoustic imagery by assembling sequential one-dimensional (1D) acoustic returns (pings) [2], [3] along the vehicle trajectory, making it a valuable sensing modality for large-area seabed mapping and SLAM applications. In SSS-based SLAM, the reliability of navigation critically depends on the ability to extract consistent and informative landmark measurements from raw acoustic returns, which remains challenging due to range-dependent intensity variations and speckle noise, both inherent characteristics of SSS measurements.

Most existing SSS-based SLAM approaches operate in the image domain, where acoustic images are formed by stacking sequential pings and visual features are extracted for data association. Fallon et al. [4] demonstrated an early SSS-based SLAM system by combining range measurements from surface beacons with landmark position estimates derived from SSS imagery. However, since reliable landmarks are difficult to detect in SSS imagery, their system relied on manually annotated features, limiting its applicability to automated scenarios. Aulinas et al. [5] introduced a boosted cascade of Haar-like features for automatic landmark detection into an EKF-based submapping framework, employing the joint compatibility branch and bound (JCBB) algorithm to resolve data association during submap fusion. For real-time onboard applications, Siantidis [6] proposed a hybrid detection strategy combining local threshold-based segmentation with template matching, further refining data association through ambiguity checks and time-gating heuristics to mitigate artifacts inherent in SSS geometry. Zhang et al. [7] used FAST-detected corner points and SIFT descriptors to establish keypoint correspondences between compared SSS images, which were subsequently integrated into a pose-graph SLAM framework. Although the preprocessing step employed a canonical transformation [8] to reduce image distortions, their approach inherently relies on ground-range projection, introducing a coordinate transformation step dependent on accurate altitude information. Despite improvements in geometric modeling and feature extraction, these approaches inherently rely on the presence of sufficiently rich image-domain features, and their performance can deteriorate in feature-sparse or homogeneous seabed environments. Furthermore, range-dependent intensity variations and speckle noise inherent in SSS measurements can degrade the reliability of feature extraction and data association.

In contrast, several studies have investigated landmark detection directly at the ping level, retaining the native acoustic

---

J. Im and S. Hong is with the Robotic Intelligence & Perception Lab., Keimyung University, Daegu, S. Korea `jinho_im@kmu.kr`, `sh.hong@kmu.ac.kr`

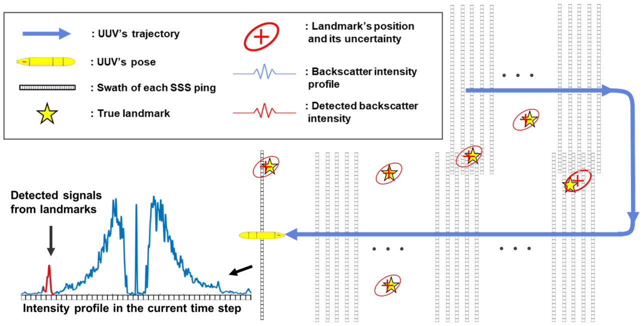


Fig. 1: Conceptual illustration of the SSS SLAM framework using the proposed ping-level landmark detection method. The framework jointly estimates the vehicle trajectory and landmark positions by incorporating peak-signal measurements detected from 1D acoustic intensity profiles (marked as blue waveforms), which represent reflected acoustic returns from the seabed.

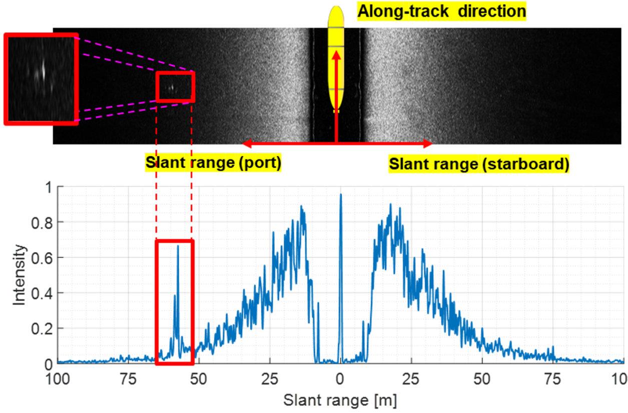


Fig. 2: Representative example of ping-level landmark detection based on backscatter intensity profiles. **Top:** Stacked SSS image. **Bottom:** Normalized backscatter intensity profile of a single ping, where the decaying baseline reflects the expected slant-range-dependent attenuation. Local high-intensity peaks that significantly deviate from this baseline are selected as potential landmark features (red boxes).

signal characteristics and avoiding geometric distortions introduced by ground-range projection and image construction steps. Chen and Guo [9] proposed a real-time method that detects salient features directly from raw backscatter signals and projects them onto a two-dimensional (2D) occupancy grid map, combining motion and measurement uncertainties for landmark localization. Al-Rawi et al. [10] developed a detector that applies a cubic smoothing spline to individual pings to model noisy backscatter data and extract peak-valley patterns corresponding to objects and their acoustic shadows. Expanding on signal-level processing, Hoff et al. [11] developed a real-time pipeline that transforms processed swaths into probabilistic 2D maps for semantic landmark detection. Davenport et al. [12] proposed a Bayesian estimation method that treats each sonar ping as a separate measurement, utilizing a particle filter to handle measurement origin uncertainty and the nonlinear nature of slant-range observations. While these studies move toward signal-level processing, many existing methods still depend on the presence of discernible acoustic shadows or specific structural priors, and the range-dependent noise chracteristics and signal attenuation inherent in SSS measurements are not explicitly addressed.

In this study, we propose a direct ping-level landmark detection method for SSS-based SLAM that exploits anomalous intensity deviations in raw backscatter profiles as landmark features, without relying on image formation or appearance-based descriptors. The proposed method is specifically designed for feature-poor seabed environments, where homogeneous backscatter returns and the absence of distinctive acoustic shadows limit the applicability of conventional image-level approaches. Fig. 1 conceptually illustrates the core idea of the proposed method. The primary contributions of this work are summarized as follows:

- We propose a ping-level landmark detection algorithm that models slant-range-dependent backscatter attenuation using a range-adaptive decay model with a spatially varying threshold, enabling reliable identification of salient intensity peaks directly from raw SSS returns without requiring image formation or appearance-based feature extraction.
- We design and implement an extended Kalman filter (EKF)-based landmark SLAM framework that directly incorporates raw SSS ping measurements, providing slant-range and azimuth observations without intermediate image construction or ground-range projection.
- We validate the proposed method through real-world field experiments, demonstrating improved detection precision over image-level and ping-level baselines and enhanced localization accuracy across two field trials conducted in feature-poor seabed environments.

II. PING-LEVEL LANDMARK DETECTION METHOD

The proposed ping-level landmark detection algorithm is inspired by the observation that raw SSS measurements inherently provide slant range information and exhibit a general trend where backscatter intensity decreases as the slant range increases. Based on this trend, the proposed algorithm identifies salient peaks in the backscatter profile that deviate from the expected attenuation pattern and considers them as potential landmark features. The corresponding slant range measurements associated with these peaks are then utilized as observation inputs for the SLAM estimation process. Fig. 2 illustrates the motivation behind this detection strategy, showing how the backscatter intensity generally decreases with increasing slant range while distinct high-intensity responses emerging from this decaying pattern serve as reliable landmark observations. Once backscattered signals are acquired from the SSS, they are first separated into port and starboard channels and processed independently. Each separated ping is subsequently passed into the detection module to extract the slant range and azimuth angle between the vehicle and the corresponding landmark. Note that the slant range is directly measured from the SSS returns, whereas the azimuth angle is assigned as either $\pi/2$ or $-\pi/2$, depending on whether the signal originates from the port or starboard channel, respectively.

A Gaussian filter is applied to a 1D ping profile to suppress high-frequency noise and improve the robustness of subsequent feature detection steps. A more detailed characterization of the general trend shown in Fig. 2 reveals that the backscatter intensity profile exhibits a distinct pattern as the slant range increases: starting from the sensor origin, the intensity generally increases toward a prominent peak near the nadir zone, after which it gradually attenuates. This transition point is defined as the near-nadir peak. In this study, the location of the near-nadir peak is automatically identified using the golden section search (GSS) method [13], which efficiently determines the extremum point where the rising intensity trend transitions into a falling trend. Compared to conventional gradient-based methods, the GSS offers better computational efficiency, making it well-suited for implementation on embedded computers commonly used in torpedo-shaped cruising underwater vehicles with limited onboard resources. Moreover, since it is a derivative-free iterative optimization method, it is particularly advantageous for our application, where numerical derivatives of noisy acoustic signals can become unstable and unreliable.

Once the near-nadir peak is identified, the subsequent portion of the profile, where the backscatter intensity exhibits a decaying trend, is extracted for further analysis. According to an existing study [14], the acoustic backscatter intensity in this region typically attenuates with increasing slant range, following a cubic polynomial pattern. To model this decay, a cubic polynomial curve is fitted to the post-peak segment of the profile, which is expressed as

$$I(r) = c_1 r^3 + c_2 r^2 + c_3 r + c_4 \tag{1}$$

where $I(r)$ denotes the backscatter intensity at slant range r. c_1, c_2, c_3, and c_4 are polynomial coefficients. The curve parameters are estimated using a variant of the random sample consensus (RANSAC) algorithm [15], which enhances robustness against measurement noise and signal irregularities. Unlike the standard RANSAC approach that uses a fixed residual threshold, the proposed method incorporates a slant range-dependent threshold. This modification accounts for the characteristic noise behavior observed in backscatter signals, where the noise level tends to be higher at slant ranges closer to the SSS sensor and decreases with increasing distance, as illustrated in Fig. 2. The residual threshold function $\epsilon\left(\cdot\right)$ is defined as a function of slant range, as shown in Eq. (2):

$$\epsilon\left(r;\ \epsilon_t, w_t\right) = \epsilon_t + w_t \frac{r_{\max} - r}{r_{\max}}. \tag{2}$$

Here, ϵ_t denotes the baseline residual threshold applied at the maximum slant range $r_{\max}$ of the SSS. The second term on the right-hand side is a correction component that relaxes the residual threshold for regions where the slant range is closer to the SSS. The parameter w_t represents the weighting factor that determines the influence of this correction term. Points that significantly deviate from the fitted decay curve, specifically outliers that do not conform to the general cubic trend, are regarded as salient intensity features and subsequently treated as landmark candidates for SLAM. Fig. 3 presents the result of applying the proposed near-nadir peak detection and the modified RANSAC algorithm to a real backscatter signal profile acquired from the starboard side of the SSS, illustrating how salient acoustic intensity peaks are extracted as landmark candidates.

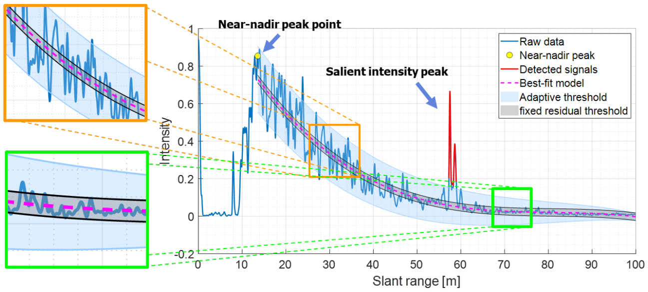


Fig. 3: Example of ping-level landmark detection results applied to real SSS data acquired on the starboard side. The decay model estimated via the modified RANSAC algorithm serves as a baseline for identifying salient intensity peaks as landmark candidates. The adaptive residual threshold accounts for range-dependent noise characteristics, in contrast to a conventional fixed threshold. Insets highlight the variation of noise levels with slant range, motivating the use of an adaptive threshold.

These salient intensity features, identified through the aforementioned procedures, may appear either at multiple scattered locations along the backscatter intensity profiles on each side of the SSS or form clusters within small local regions. To group meaningful responses and suppress spurious detections, the density-based spatial clustering of applications with noise (DBSCAN) algorithm [16] is applied to the extracted landmark candidates. The clustered signals are aggregated by computing the average slant range ρ_i for ith group, which is then incorporated into the SLAM framework as a measurement.

III. SSS SLAM FRAMEWORK

An EKF-based SLAM framework tailored to the SSS sensing modality, along with its key formulations, is presented in this section. The full state vector of the proposed SSS SLAM with n_ℓ landmarks at time step k is defined as follows:

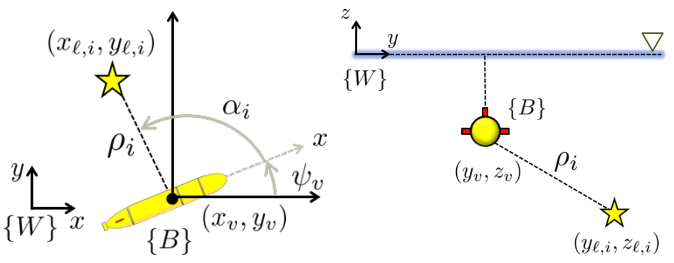
$$\mathbf{x}_k = \begin{bmatrix}\mathbf{x}_v^\top & \mathbf{x}_{\ell,1}^\top & \mathbf{x}_{\ell,2}^\top & \cdots & \mathbf{x}_{\ell,i}^\top & \cdots & \mathbf{x}_{\ell,n_\ell}^\top\end{bmatrix}^\top \tag{3}$$

where $\mathbf{x}_v$ denotes the vehicle's pose vector, and $\mathbf{x}_{\ell,i} = \left[x_{\ell,i}\ y_{\ell,i}\ z_{\ell,i}\ \right]^\top$ represents the three-dimensional (3D) position vector of the ith landmark. The coordinate systems and measurement geometry for SSS SLAM are illustrated in Fig. 4. The stochastic measurement model using the slant range and azimuth angle obtained from the ith landmark is expressed as follows:

$$\begin{aligned}\mathbf{z}_i &= \begin{bmatrix}\rho_i \\ \alpha_i\end{bmatrix} = \begin{bmatrix}h_1(\mathbf{x}_v, \mathbf{x}_{\ell,i}) \\ h_2(\mathbf{x}_v, \mathbf{x}_{\ell,i})\end{bmatrix} + \mathbf{v} \\ &= \begin{bmatrix}\sqrt{(x_{\ell,i} - x_v)^2 + (y_{\ell,i} - y_v)^2 + (z_{\ell,i} - z_v)^2} \\ \operatorname{atan2}(y_{\ell,i} - y_v, x_{\ell,i} - x_v) - \psi_v\end{bmatrix} + \mathbf{v}\end{aligned}. \tag{4}$$

The random vector $\mathbf{v} \sim \mathcal{N}\left(0, \mathrm{R}\right)$ refers to the measurement noise vector whose elements are assumed to follow a zero-mean Gaussian distribution.

While a complete representation of UUV pose in 3D space nominally requires a 6-DOF state vector encompassing both



(a) Top-down view (b) Vertical view

Fig. 4: Coordinate frames and geometric relationship of SSS measurements. (x_v, y_v, z_v, ψ_v) and $(x_{\ell,i}, y_{\ell,i}, z_{\ell,i})$ denote the vehicle pose and the position of the ith landmark, respectively, both represented in the global frame W. ρ_i and α_i indicate the slant range and azimuth measurements obtained from the SSS.

position and orientation, we adopt a reduced 4-DOF representation comprising 3D position and yaw. This reduction is justified by the inherent observability limitations of SSS measurements. In conventional 3D landmark-based SLAM with range and bearing observations, pitch is geometrically coupled with landmark vertical positions, and roll is coupled with the lateral displacement of acoustic returns. However, in SSS-based SLAM, landmark elevation is not directly constrained by slant-range and azimuth measurements, which causes noticeable fluctuations in estimated landmark vertical positions and consequently degrades the reliability of pitch estimation. Similarly, roll induces asymmetric distortions in the across-track swath geometry that are not explicitly observable from slant-range measurements alone, making roll estimation unreliable in this sensing modality. Excluding both roll and pitch from the state therefore yields more consistent and stable SLAM estimation performance. This formulation further assumes that the UUV maintains roll and pitch angles close to zero during steady-state cruising, which is a reasonable assumption for the survey-class underwater vehicles considered in this work. Based on this analysis and assumption, we adopt a reduced 4-DOF vehicle pose representation comprising only the 3D position and yaw. The system dynamics are modeled as a nonlinear stochastic process, expressed as follows:

$$\dot{\mathbf{x}}_v = \begin{bmatrix} \dot{x}_v \\ \dot{y}_v \\ \dot{z}_v \\ \dot{\psi}_v \end{bmatrix} = \begin{bmatrix} v_x \cos\psi - v_y \sin\psi \\ v_x \sin\psi + v_y \cos\psi \\ v_z \\ w_\psi \end{bmatrix} + \mathbf{w} \tag{5}$$

where v_x, v_y and v_z denote the linear velocities along the x, y and z-axes respectively, and w_ψ denotes the yaw rate. These terms constitute the control input vector $\mathbf{u} = [v_x \ v_y \ v_z \ r]^\top$. The linear velocities can be measured from a DVL, and the yaw rate is provided by an IMU. $\mathbf{w}$ is the process noise vector that is assumed to follow a zero-mean Gaussian distribution.

For data association between the current measurements and predicted observations, a nearest-neighbor strategy based on the Mahalanobis distance is employed. For each detected feature, an innovation vector is computed as the difference between the measurement and its corresponding predicted observation, and the Mahalanobis distance is evaluated accordingly. Data association is then performed by selecting the predicted observation that yields the minimum Mahalanobis distance, provided that the distance lies within a predefined gating threshold.

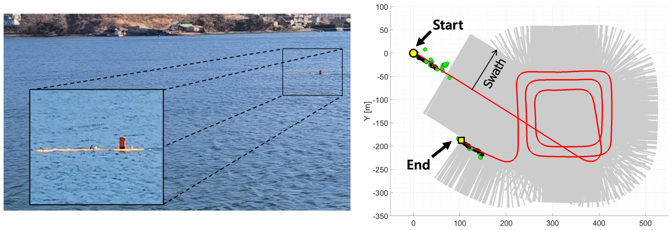


Fig. 5: Experimental setup for the field experiments. **Left:** The UUV operating at the sea surface prior to submergence. **Right:** A representative operational trajectory from `Field Trial 1`, obtained from real-time underwater navigation and control, showing the DR (red solid line), the footprint of the SSS swath (gray region), and GPS measurements acquired during surfacing intervals (green points).

IV. EXPERIMENTAL RESULTS

To evaluate the practical applicability of the proposed method, two field experiments were conducted in real sea environments. The seabed at the test sites consisted of generally gentle terrain with a few sparse protruding structures, representing feature-poor conditions. Fig. 5 presents the experimental setup and, as a representative example, the operational trajectory executed during `Field Trial 1`. The experimental UUV was equipped with an integrated sensor suite comprising an IMU (FiberPro, FI200P), a DVL (Teledyne, Pathfinder), an SSS (SonarTech, SEAVIEW), and a GPS receiver (NovAtel, OEM7600). The SSS operated at a center frequency of $455\,\mathrm{kHz}$ and provided an effective swath width of approximately $100\,\mathrm{m}$ on each side. A DR approach using proprioceptive odometry (IMU and DVL), a representative image-level SSS SLAM method (IL-SSS [7]), and a representative ping-level peak-valley detection method (PV-SSS [10]) were adopted as baseline methods.

For evaluating landmark detection performance, Ground truth (GT) landmark positions were established by manual annotation. Both #GT and #F are defined on a per-landmark basis: #GT denotes the total count of GT landmarks encountered along each trajectory, and #F denotes the total count of landmarks detected by each method. Multiple detections in close spatial proximity to a single landmark were grouped prior to evaluation. Evaluation metrics include true positives (TP), false positives (FP), false negatives (FN), and precision (Pr.); true negatives (TN) are omitted as a clearly delineated negative region cannot be defined within a continuous ping intensity profile. For evaluating vehicle localization performance of the SSS SLAM system, GT trajectory information was not available in the real sea environments. Instead, GPS position fixes obtained when the UUV surfaced at the end of each survey mission were used as positional references. In addition, salient landmarks visible in the SSS imagery were manually annotated and incorporated into the proposed SLAM framework to construct a reference trajectory, which served as an approximate upper bound for comparative performance analysis.

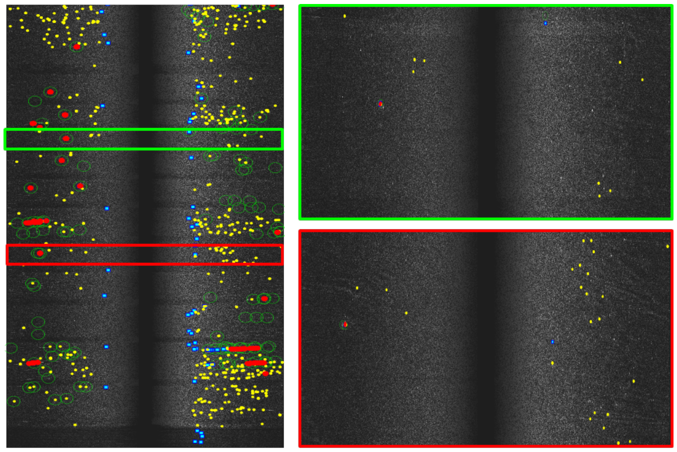
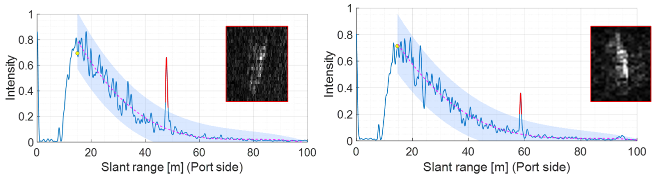


Fig. 6: Qualitative comparison of landmark detection results in Field Trial 1. **Left:** Detection results from the compared three methods overlaid on the SSS image (green circle: manually annotated GT landmark, yellow: IL-SSS, blue: PV-SSS, red: proposed) **Right:** Zoomed-in views of the highlighted regions at two representative along-track positions.

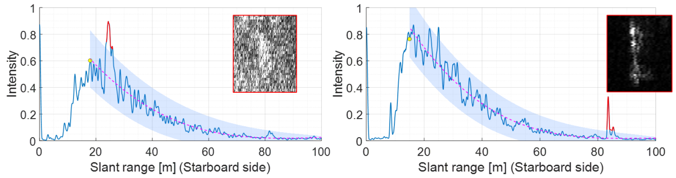
TABLE I: Landmark detection performance of each method in two field trials conducted in real sea environments.

Method	Field Trial 1 (#GT=120)					Field Trial 2 (#GT=166)				
	#F	TP	FP	FN	Pr.	#F	TP	FP	FN	Pr.
IL-SSS	337	76	261	44	0.23	134	80	54	86	0.60
PV-SSS	68	31	37	89	0.46	91	46	45	120	0.51
Proposed	46	35	11	85	**0.76**	100	66	34	100	**0.66**

Fig. 6 qualitatively compares the landmark detection results of IL-SSS, PV-SSS, and the proposed method for Field Trial 1, with detection points overlaid on the SSS imagery. IL-SSS generates a large number of detections scattered across the image, many of which are associated with weak backscatter fluctuations or speckle noise rather than physically meaningful landmarks. In contrast, both PV-SSS and the proposed method produce sparser and more spatially concentrated detections. However, PV-SSS tends to become less reliable at longer ranges where signal attenuation reduces the prominence of true landmark peaks relative to background noise, whereas the proposed method maintains more consistent detection through its range-adaptive threshold. The quantitative evaluation results are summarized in Table I. IL-SSS produces many false positives in both field trials due to its sensitivity to image-level intensity variations. Notably, the precision of IL-SSS is substantially lower in Field Trial 1 than in Field Trial 2, which may reflect differences in the acoustic characteristics of the two survey environments. PV-SSS shows unsatisfactory reliability in both trials, particularly for landmarks at longer ranges. The proposed method achieves the highest precision in both field trials, demonstrating that the range-adaptive detection criterion generalizes well across different field conditions. To further illustrate the behavior of the proposed range-adaptive threshold under varying noise conditions, Fig. 7 presents representative detection results at different slant ranges from the two field trials, all generated using the same parameter settings. In the close-range example from Field Trial 2 (approximately 25 m), where the noise level is relatively high, the adaptive threshold increases accordingly, allowing the detector to suppress surrounding noise peaks while preserving salient landmark responses that sufficiently deviate from the fitted decay model.



(a) Field Trial 1

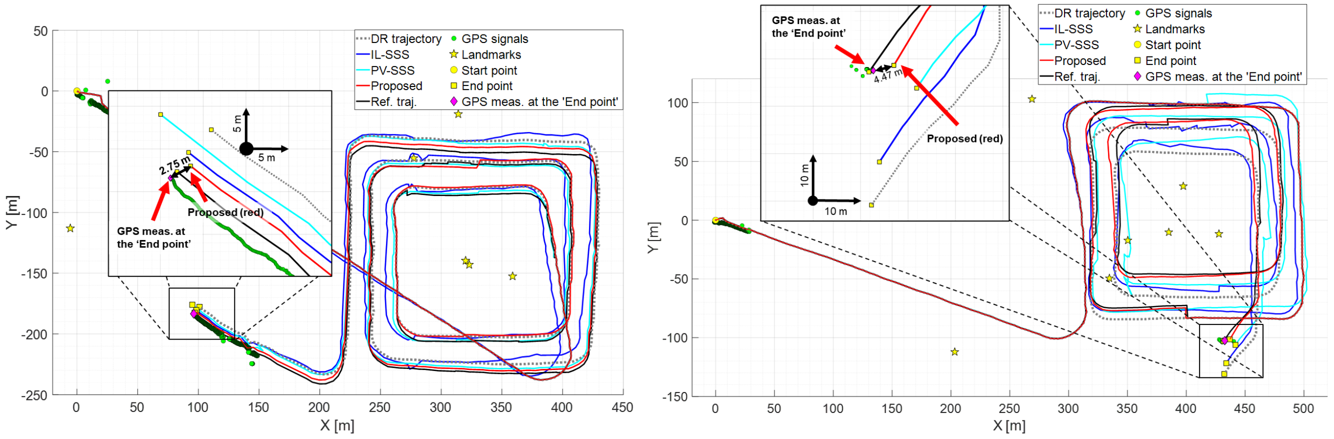


(b) Field Trial 2

Fig. 7: Representative detection results from two field trials, illustrating how the proposed range-adaptive threshold (Eq. (2)) behaves across different slant ranges using a fixed set of parameters. For each case, the corresponding cropped SSS image of the detected intensity peak is shown as an inset.

Fig. 8 compares the estimated trajectories produced by the baseline methods and the proposed approach for the two field experiments. In addition to the GPS position fixes obtained at surfacing, a reference trajectory (black solid line) is constructed using manually annotated landmark detections within the proposed SLAM framework to provide a more informative baseline for comparison. Note that the scattered GPS measurements observed near the end point reflect the discrete and intermittent nature of GPS fixes acquired only during surfacing intervals, rather than measurement noise. Although sparse, these references are sufficient to assess drift trends, as the final position discrepancies across methods are significantly larger than the GPS uncertainty. DR exhibits substantial accumulated drift in both trials. While IL-SSS and PV-SSS reduce this drift to some extent, both show noticeable deviations along repeated track segments and long-range detection regions. In contrast, the proposed method maintains the closest agreement with both the GPS measurements and the reference trajectory.

The quantitative comparison in Table II further supports these observations. The proposed method achieves the lowest FPE and normalized FPE across both field trials, yielding lower errors than DR, IL-SSS, and PV-SSS. The improved performance of the proposed method can be attributed to its measurement representation and feature extraction strategy. Unlike IL-SSS and PV-SSS, which are sensitive to range-dependent attenuation and local signal prominence respectively, the proposed method models the full backscatter intensity profile with a range-adaptive formulation, enabling more consistent landmark detection and improved drift suppression. For reference, the results obtained using manually annotated landmark detections are also reported, representing an approximate upper bound on achievable performance under ideal data association. The computational performance of the



(a) Field Trial 1

(b) Field Trial 2

Fig. 8: Trajectory comparison for the two field experiments. The proposed method shows the closest agreement with both GPS measurements and the reference trajectory constructed using manually annotated landmarks. While DR suffers from significant drift and IL-SSS exhibits limited correction capability, PV-SSS improves consistency but degrades in long-range regions. Insets highlight localization accuracy near the end points.

TABLE II: Quantitative comparison of localization performance in the field experiments. FPE and normalized FPE (FPE/L) are reported for the compared methods. L denotes the total trajectory length. Results obtained using manually annotated landmark detections are also included as a reference upper bound.

Method	Field Trial 1 (L=2,243 m)		Field Trial 2 (L=2,012 m)	
	FPE [m]	FPE/L [%]	FPE [m]	FPE/L [%]
DR	7.39	0.33	28.36	1.41
IL-SSS	3.66	0.16	19.31	0.96
PV-SSS	7.26	0.32	9.87	0.49
Proposed	**2.75**	**0.12**	**4.47**	**0.22**
Ref. Traj. (w/ GT landmarks)	1.07	0.05	1.21	0.06

proposed framework was evaluated on a laptop PC equipped with an Intel Core i7-11700 processor using a real-time C/C++ implementation. Averaged across the two field experiments, the detection module required approximately 4 ms per SSS ping, while the EKF module required 1 ms per ping. These results demonstrate the computational efficiency of the proposed approach and support its feasibility for real-time deployment on embedded platforms.

V. CONCLUSION

This paper presented a ping-level landmark detection method for SSS-based SLAM in feature-poor seabed environments. By detecting landmarks directly from raw backscatter intensity profiles without acoustic image formation, the proposed method preserves the native sensing geometry of SSS and enables robust landmark extraction even in feature-sparse conditions. The effectiveness and feasibility of the proposed method were validated through real-world field experiments.

REFERENCES

[1] L. Paull, S. Saeedi, M. Seto, and H. Li, "AUV navigation and localization: A review," *IEEE J. Ocean. Eng.*, vol. 39, pp. 131–149, 2013.
[2] S. Reed, I. T. Ruiz, C. Capus, and Y. Petillot, "The fusion of large scale classified side-scan sonar image mosaics," *IEEE Trans. Image Process.*, vol. 15, no. 7, pp. 2049–2060, 2006.
[3] P. Blondel, *The Handbook of Sidescan Sonar*. Springer, 2009.
[4] M. F. Fallon, M. Kaess, H. Johannsson, and J. J. Leonard, "Efficient AUV navigation fusing acoustic ranging and side-scan sonar," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2011, pp. 2398–2405.
[5] J. Aulinas, X. Lladó, J. Salvi, and Y. R. Petillot, "Feature based SLAM using side-scan salient objects," in *Proc. MTS/IEEE OCEANS*, 2010.
[6] K. Siantidis, "Side scan sonar based onboard SLAM system for autonomous underwater vehicles," in *Proc. IEEE/OES Auton. Underwater Veh.*, 2016, pp. 195–200.
[7] J. Zhang, Y. Xie, L. Ling, and J. Folkesson, "A fully-automatic side-scan sonar simultaneous localization and mapping framework," *IET Radar, Sonar & Navig.*, vol. 18, no. 5, pp. 674–683, 2024.
[8] W. Xu, L. Ling, Y. Xie, J. Zhang, and J. Folkesson, "Evaluation of a canonical image representation for sidescan sonar," in *Proc. MTS/IEEE OCEANS*, 2023.
[9] E. Chen and J. Guo, "Real time map generation using sidescan sonar scanlines for unmanned underwater vehicles," *Ocean Eng.*, vol. 91, pp. 252–262, 2014.
[10] M. Al-Rawi, A. Galdran, F. Elmgren, J. Rodriguez, J. Bastos, and M. Pinto, "Landmark detection from sidescan sonar images," in *Proc. IEEE Jordan Conf. Appl. Electr. Eng. Comput. Technol.*, 2017, pp. 1–6.
[11] S. Hoff, V. Haraldstad, B. R. Hogstad, and D. Varagnolo, "Side-scan sonar based landmark detection for underwater vehicles," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2024, pp. 11 723–11 729.
[12] E. Davenport, K. Nguyen, J. Jang, C. Ma, S. Fish, L. Lenain, and F. Meyer, "A landmark-aided navigation approach using side-scan sonar," *IEEE J. Ocean. Eng.*, 2025.
[13] J. Kiefer, "Sequential minimax search for a maximum," in *Proc. Amer. Math. Soc.*, 1953, pp. 502–506.
[14] G. R. Elston and J. M. Bell, "Pseudospectral time-domain modeling of non-Rayleigh reverberation: Synthesis and statistical analysis of a sidescan sonar image of sand ripples," *IEEE J. Ocean. Eng.*, vol. 29, no. 2, pp. 317–329, 2004.
[15] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, pp. 381–395, 1981.
[16] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *Proc. Int. Conf. Knowledge Discovery and Data Mining*, 1996, pp. 226–231.