ICRA2026 Late Breaking Results Poster presented at
2026 IEEE International Conference on Robotics and Automation (ICRA 2026)
June 1-5, 2026. Vienna, Austria

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

Jinho Im and Seonghun Hong†

I. Introduction

Unmanned underwater vehicles (UUVs) typically rely on dead-reckoning (DR) using inertial measurement unit (IMU) and Doppler velocity log (DVL), which inevitably leads to accumulated drift during long-term operations. Due to its wide-area coverage and long sensing range, side-scan sonar (SSS) has been widely used as an exteroceptive sensing modality [1], [2] and is considered a promising option for underwater simultaneous localization and mapping (SLAM). However, most existing SSS SLAM approaches [3]–[5] operate on image-level representations, which can suffer from unstable data association in feature-poor or homogeneous seabed environments. To address these limitations, this study explores a ping-level interpretation of raw backscatter signals as an alternative to conventional image-level processing. Specifically, we propose a direct ping-level landmark detection method for SSS SLAM that exploits structurally salient deviations in raw backscatter intensity profiles as landmark features, without relying on image formation or appearance-based descriptors.

II. Methodology

Raw SSS ping measurements are processed through a four-stage detection pipeline consisting of preprocessing, near-nadir peak search, salient intensity detection, and clustering to generate landmark observations for the SLAM estimation process. In the preprocessing stage, raw SSS signals are first separated into port and starboard channels and processed independently. A Gaussian filter is applied to each one-dimensional backscatter intensity profile to suppress high-frequency noise and improve the robustness of subsequent feature detection steps. As illustrated in Fig. 1, the backscatter intensity profile exhibits a characteristic pattern with respect to increasing slant range: 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, referred to as the near-nadir peak, is automatically identified using the golden section search method, which efficiently locates the peak separating the rising and attenuation-dominant intensity trends without relying on derivative computations of noisy acoustic signal intensities. 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.

J. Im and S. Hong are with the Robotic Intelligence & Perception Lab., Keimyung University, Daegu, S. Korea. (e-mail: jinho_im@kmu.kr; sh.hong@kmu.ac.kr)

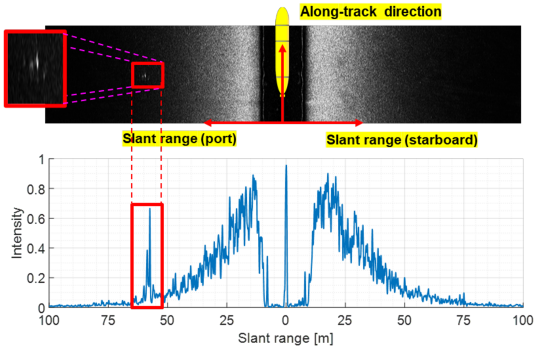

Fig. 1: Representative example of ping-level landmark detection based on backscatter intensity profiles. **Top:** Stacked SSS image. **Bottom:** Normalized backscatter intensity profile of a single ping, where the decaying baseline reflects the expected slant-range-dependent attenuation. Local high-intensity peaks that significantly deviate from this baseline are selected as potential landmark features (red boxes).

According to an existing study [6], 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, 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. 1. The residual threshold function $\epsilon(\cdot)$ is defined as a function of slant range, as shown in Eq. (2):

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

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

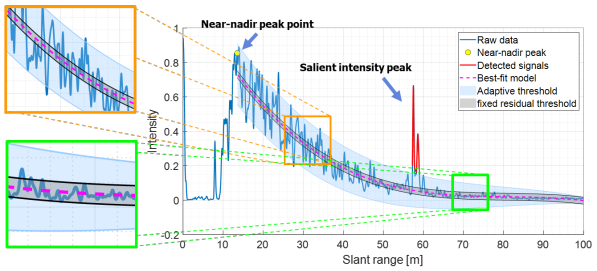

Fig. 2: Example of ping-level landmark detection results applied to real SSS data acquired on the starboard side. The decay model estimated via the modified RANSAC algorithm serves as a baseline for identifying salient intensity peaks as landmark candidates. The adaptive residual threshold accounts for 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.

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. 2 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.

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 is applied to the extracted landmark candidates. The clustered signals are subsequently aggregated by computing the average slant range for each group, which is then incorporated into the SLAM as a measurement.

III. Experimental Results

To validate the effectiveness of the proposed approach, a landmark-based SLAM framework incorporating the proposed ping-level landmark detection method was implemented and evaluated through real sea experiments. For comparison, a DR approach based on proprioceptive sensors, including an IMU and a DVL, and a representative image-level SSS SLAM approach, referred to as IL-SSS [5], were adopted as baseline methods. A qualitative comparison of the resulting trajectory estimates is presented in Fig. 3. Although ground-truth information regarding seabed geometry or landmark locations was not available in the real sea environment, absolute global positioning system (GPS) position fixes obtained immediately before submergence

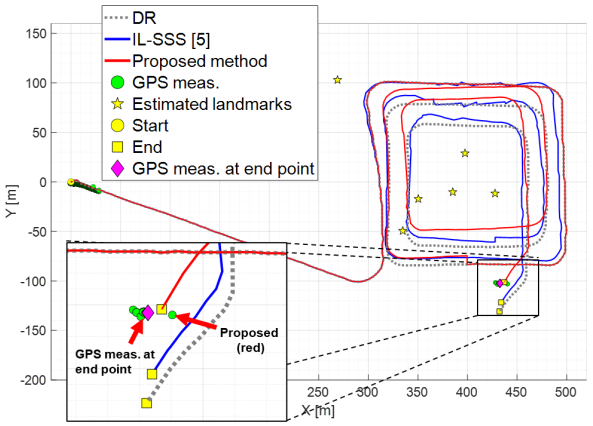

Fig. 3: Trajectory comparison for the field experiment conducted in a real sea environment. DR (gray dashed line), IL-SSS [5] (blue solid line), and the proposed method (red solid line) are compared against discrete GPS surfacing measurements (green dots). Star markers indicate the landmark positions estimated by the proposed method.

and after resurfacing were used as reference positions for evaluating localization accuracy. The experimental results show that DR suffers from substantial accumulated drift, leading to a final position error of $28.36\,\mathrm{m}$. The IL-SSS method reduced this error to $19.31\,\mathrm{m}$, indicating partial drift correction but still limited localization consistency. In contrast, the proposed SSS SLAM approach achieved a final position error of only $4.47\,\mathrm{m}$, demonstrating significantly improved drift suppression and yielding the trajectory most consistent with the GPS reference measurements.

IV. 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] 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.

[2] P. Blondel, *The Handbook of Sidescan Sonar*. Springer, 2009.

[3] 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.

[4] Y. Yang, C. Pang, C. Wu, and Z. Fang, "Geometry-aided underwater 3D mapping using side-scan sonar," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2024, pp. 1038–1045.

[5] 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.

[6] 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.