

Bundle Adjustment from Sonar Images and SLAM Application for Seafloor Mapping

Young-Sik Shin*, Yeongjun Lee[†], Hyun-Taek Choi[†], Ayoung Kim*

*Department of Civil and Environmental Engineering, KAIST, S. Korea

email:{youngsik.shin, ayoungk}@kaist.ac.kr

[†]Korea Research Institute Ship and Ocean engineering (KRISO), S. Korea

email:{leeyoungjun, htchoi}@kriso.re.kr

Abstract—This paper reports on two-view bundle adjustment using sonar images, specifically focusing on feature detection and a sensor measurement model for imaging sonar. To overcome limited sensor information for underwater navigation, we use Dual frequency IDentification SONar (DIDSON) in the imaging mode to provide spatial constraints when a scene is revisited. Unlike terrestrial images, sonar images are usually low resolution with highly speckled noise. We found that exploiting features from nonlinear scale space improves feature detection. In this paper, we adopt KAZE features and use random sample consensus (RANSAC) to refine correspondences. Using these correspondences, we propose point-based relative pose estimation via bundle adjustment. The target application that this work focuses on is underwater seafloor mapping, and the proposed model assumes a fixed elevation. Through this work, we present (i) validation of nonlinear scale space features for sonar images and (ii) proposal of a sonar sensor measurement model for underwater simultaneous localization and mapping (SLAM). The proposed method will be validated through both synthetic data sets and a tank test for seafloor mapping.

I. INTRODUCTION

Underwater navigation is challenging due to the limited sources of sensors for mapping and localization. The medium of water prevents conventional communication and global positioning system (GPS) signal from penetrating through water and being used for navigation. Despite these challenges, the robotics community has searched for solutions in underwater navigation (e.g., underwater ship hull [1], dam [2], pipeline inspection [3] and under-ice exploration [4]).

For this GPS-denied environment, SLAM presents a navigation method. Visual recognition for reducing navigation uncertainty has been the main idea of SLAM. By recognizing previously visited places, a robot can reduce the navigation uncertainty and correct error. Despite challenges that the underwater optical camera poses, several successful implementations can be found in the underwater visual SLAM literature. Williams and Mahon [5] successfully mapped the Great Barrier Reef using bottom looking cameras. Mapping of the Titanic using visual SLAM was reported by Eustice et al. [6]. Medagoda et al [7] proposed a method that improves mid-water column navigation also using visual SLAM. More recently, Kim and Eustice [8] presented visual SLAM implemented on an autonomous underwater vehicle (AUV) to correct longitudinal drift on the ship hull.

For underwater application, sonar has been a popular sensor due to its robustness and independence of turbidity in

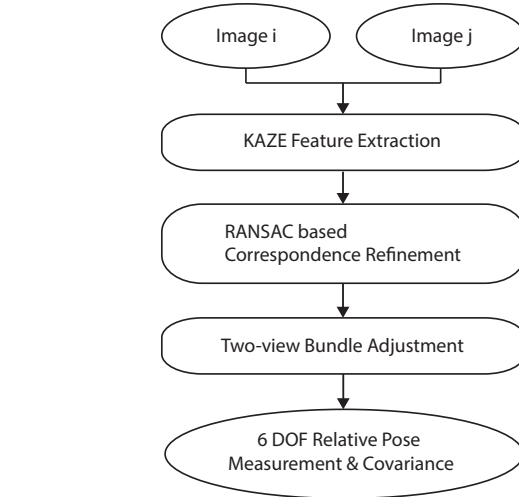


Fig. 1. Overview of two-view bundle adjustment algorithm. For each image, KAZE feature extraction is performed for keypoint and descriptor. Then NNDR matching process find initial matching pairs. In addition, RANSAC based outlier rejection is performed for robust correspondence. Finally, we obtain 6-DOF relative pose constraint in the two-view bundle adjustment step.

water. However, using sonar for direct navigation addresses many challenges. Sonar images are usually low resolution, and estimating motion from sonar sensor information is not straightforward. One challenge is at elevation ambiguity. For optical images, the epipolar geometry from multiple views offers the 3-D scene interpretation. However, there is scale factor ambiguity. In contrast, measurements of 2-D imaging sonar are composed of range and azimuth from 3-D scene, excepting the elevation. The unknown elevation generates the motion ambiguity problem of a plane as theoretically proved by Negahdaripour [9]. The 3-D motion estimation from imaging sonar is treated as a nonlinear optimization problem in [10], [11]. According to the result in [11], the uncertainty of translation is higher for z directional translation while rotation uncertainty is higher for x and y rotations (roll, pitch) due to elevation ambiguity.

To find matched images from different poses, various imaging sonar registration techniques based on local feature-based, region-based, and Fourier-based approaches have been proposed in [1], [12]–[14]. As a region-based approach, the Nor-

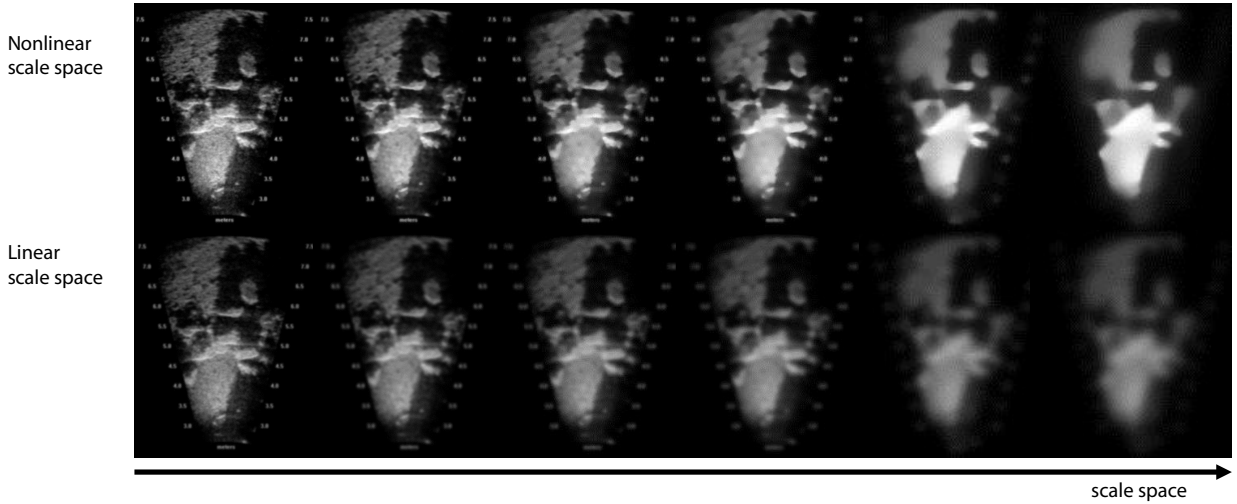


Fig. 2. Comparison between the linear and nonlinear scale space in accordance with scale size. First row is nonlinear scale space by nonlinear diffusion filtering and second row is linear scale space using Gaussian filtering. As the scale increases, nonlinear scale space approach better preserves texture detail, while efficiently reducing speckle noise.

mal Distribution Transform (NDT) was successfully applied by [15]. Similarly, [13] propose a method using a cluster-based Gaussian map. On the other hand, the image transformation parameters were calculated in Fourier domain in [12]. In [14] and [16], authors employ the method based on a point feature for optical images. However, they are highly sensitive to low-resolution sonar images with speckle noise. Other successful sonar-based navigation is reported in [1] and [17]. [1] uses DIDSON in two different mode to map the environment and localize the robot to the built map. [17] incorporated SLAM to build a 3-D photomosaic using DIDSON 2-D imaging sonar data.

To overcome aforementioned issues, we propose two solutions to the aforementioned research questions. First we examine nonlinear scale space features in sonar images and their robustness for an image matching purpose. Second, we propose a simplified motion model for imaging sonar to be used in underwater SLAM.

II. TWO-VIEW BUNDLE ADJUSTMENT

A. Overview

For SLAM implementation, we use the geometric constraint from a pair of sonar images with overlap to reduce navigation uncertainty and correct navigation drift. The geometric constraints are acquired by two-view bundle adjustment using the image pair. Fig. 1 represents the flow of two-view bundle adjustment for imaging sonar, which consists of the following three steps.

- 1) Feature extraction: For feature extraction and description, we use KAZE features. Computation efficiency is increased by using Accelerated-KAZE (A-KAZE) [18].
- 2) Correspondence refinement: Initial matching pairs are searched by NNDR matching process, using a k-dimensional tree (kd-tree) approach. For robust estimation, we perform the RANSAC-based homography for

outlier rejection. This is appropriate because we assumed a planar surface model.

- 3) Two-view bundle adjustment: We obtain 6-DOF relative pose and its covariance by two-view bundle adjustment using our reprojection error model that is described in the next section §II-D.

B. KAZE Features

Conventional scale-invariant local feature detectors build linear scale space by Gaussian blurring (e.g., Scale Invariant Feature Transform (SIFT) [19] and Speeded Up Robust Features (SURF) [20]). Blurring in linear scale space does not preserve texture detail because it smooths both noise and detail in equal scale. To cope with this issue, constructing nonlinear scale space has been addressed as in KAZE [21] and bilateral filter scale-invariant feature transform (BFSIFT) [22]. Focusing on the fact that sonar images have low resolution with highly speckled noise, we propose to use KAZE features [21] and extract features in a nonlinear scale space.

KAZE, however, is known to be slow for real-time implementation. It is computationally expensive to extract the KAZE feature for the Additive Operator Splitting (AOS) scheme to build nonlinear scale space and use Modified-SURF. According to [21], KAZE features require more expensive computation time than SURF or binary features, but are a little better than SIFT. To improve computation cost, a Fast Explicit Diffusion (FED) scheme and Modified-Local Difference Binary (M-LDB) descriptor apply to A-KAZE in [18]. By FED scheme, A-KAZE dramatically speeds up on building the nonlinear scale space. Furthermore, to enhance computational demand and lower memory requirement, they proposed an M-LDB descriptor to A-KAZE. It is rotation and scale invariant while the original LDB descriptor that is neither rotation nor scale invariant.

To better illustrate the effect of using nonlinear scale space, we provide comparison result of two scale spaces in Fig. 2.

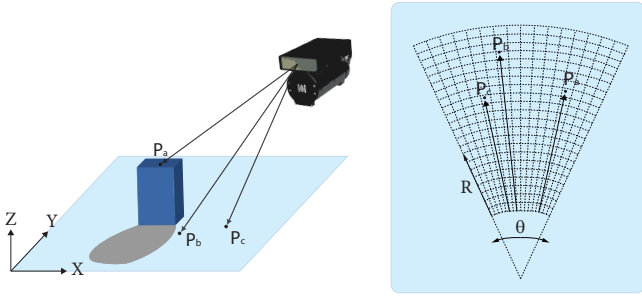


Fig. 3. Sonar coordinate definition. Three 3-D points (P_a , P_b , P_c) are projected on image plane with the zero elevation ($\phi = 0$). Left shows three points in Cartesian space and right shows the projected points in the R - θ polar plane

The series of filtered images depict a comparison between the linear scale space and the nonlinear one using FED. As can be seen, strong image edges remain unaffected in nonlinear scale space (top). This can improve robustness for image matching in case of sonar images that are corrupted by speckle noise.

C. Sensor Coordinate Description

We start derivation by defining sonar sensor coordinate system. Similar to conventional imaging sonar geometry [9], [23]–[25], we represent a point, $\mathbf{P}_s$, in a sonar sensor frame on both Cartesian and spherical coordinates.

$$\mathbf{P}_s = \begin{bmatrix} X_s \\ Y_s \\ Z_s \end{bmatrix} = \Re \begin{bmatrix} \cos \phi \sin \theta \\ \cos \phi \cos \theta \\ \sin \phi \end{bmatrix} \quad (1)$$

$$\begin{bmatrix} \theta \\ \phi \\ \Re \end{bmatrix} = \begin{bmatrix} \tan^{-1} \left(\frac{X_s}{Y_s} \right) \\ \tan^{-1} \left(\frac{Z_s}{\sqrt{X_s^2 + Y_s^2}} \right) \\ \sqrt{X_s^2 + Y_s^2 + Z_s^2} \end{bmatrix}$$

Using this relation, we define a point s on an image plane that is defined on the zero elevation ($\phi = 0$) in a spherical coordinate. Fig. 3 illustrates the projection of 3-D points onto the sonar image plane.

$$s = \begin{bmatrix} x_s \\ y_s \end{bmatrix} = \Re \begin{bmatrix} \sin \theta \\ \cos \theta \end{bmatrix} \quad (2)$$

We assume a point, $\mathbf{P}_s$ that lies on a plane with surface normal $\mathbf{n} = [n_x, n_y, n_z]^T$. The point on the plane with surface normal $\mathbf{n}$ satisfies the equation $\mathbf{P}_s \cdot \mathbf{n} = 1$. In the spherical coordinates, the surface equation can be expressed in the form:

$$(n_x \sin \theta + n_y \cos \theta) \cos \phi + n_z \sin \phi = \frac{1}{\Re} \quad (3)$$

Note that the elevation angle, ϕ , can be calculated from the plane representation in [25].

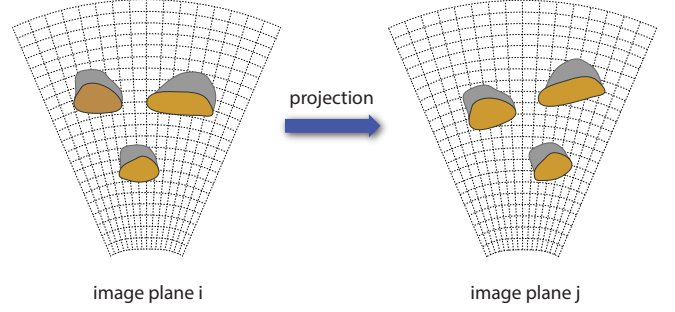


Fig. 4. Image-to-image transformation based projection between two sonar frames

D. Sonar Measurement Model

Using the coordinate defined in the previous section, we now derive the sensor measurement model under motion. The motion is described as a transformation between the two sensor poses and is defined in terms of rotation matrix, $\mathbf{R}$, and 3-D translation vector, $\mathbf{t}$.

The reprojection error is evaluated to estimate the relative pose. A point $s = (x_s, y_s)$ in the first sonar image frame can be projected to the mapped point $s' = (x'_s, y'_s)$ on the second sonar image's plane using the image-to-image transformation below.

$$\begin{pmatrix} x_s \\ y_s \end{pmatrix} \xrightarrow{\mathbf{n}} \begin{pmatrix} X_s \\ Y_s \\ Z_s \end{pmatrix} \xrightarrow{\mathbf{R}, \mathbf{t}} \Re' \begin{pmatrix} \cos \phi' \sin \theta' \\ \cos \phi' \cos \theta' \\ \sin \phi' \end{pmatrix} \quad (4)$$

$$\xrightarrow{\phi=0} \Re'_s \begin{pmatrix} \sin \theta'_s \\ \cos \theta'_s \end{pmatrix} = \begin{pmatrix} x'_s \\ y'_s \end{pmatrix}$$

The point s is transformed to a 3-D coordinate using a surface equation with a normal vector $\mathbf{n}$. Then the 3-D point is transformed using the rotation $\mathbf{R}$ and the translation $\mathbf{t}$ and then the expression can be represented with spherical coordinates. By changing the elevation ϕ for zero, we obtain image-to-image transformation $h(\cdot)$. This transformation is described in Fig. 4.

Using this transformation, we define objective function to optimize for the relative pose ($\mathbf{R}, \mathbf{t}$) and the surface normal $\mathbf{n}$. The objective function is to minimize the reprojection error between a sonar point s and the estimated $\hat{s}$.

$$\min_{\hat{\mathbf{h}}, \hat{\mathbf{x}}} \sum_{i=1}^{N_c} d(s_i, \hat{s}_i)^2 + d(s'_i, \hat{s}'_i)^2 \quad (5)$$

$$\text{s.t.} \quad \hat{s}'_i = h(\hat{\mathbf{h}}, \hat{s}_i)$$

$$\text{where,} \quad \hat{\mathbf{h}} = [\hat{\mathbf{R}}, \hat{\mathbf{t}}, \hat{\mathbf{n}}]^T$$

In this optimization problem, s and $\hat{s}$ indicates the measured points and the estimated points, and the notation $d(\cdot)$ represent the Euclidean distance between a pair of points. A set of estimated rotation $\hat{\mathbf{R}}$, translation $\hat{\mathbf{t}}$ and surface normal $\hat{\mathbf{n}}$ is denoted as $\hat{\mathbf{h}}$. The image-to-image transformation $h(\cdot)$ is given in (4).

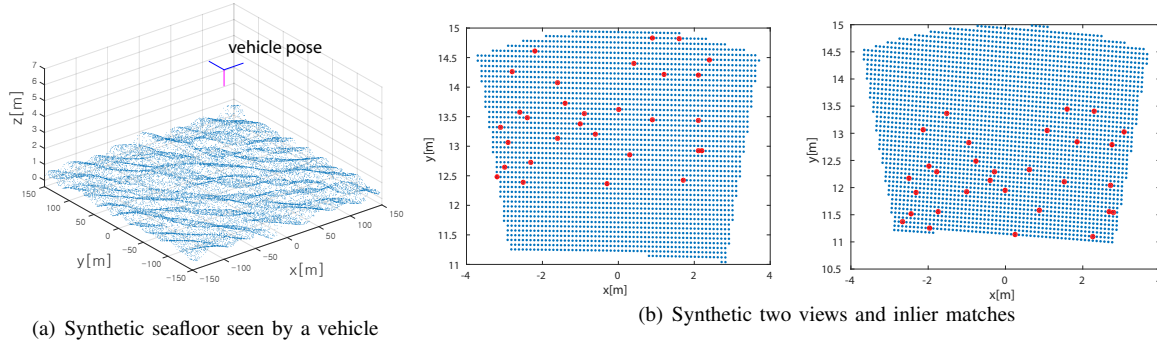


Fig. 5. Depiction of synthetic data generation. (a) Synthetic 3-D points with vehicle pose. (b) Correspondences projected on sonar image plane. Synthetic points are in blue (small) dots, and the red (larger) dots indicate correspondences.

We use this sensor measurement model in bundle adjustment [26] to generate a relative pose between two poses. The resulting pose constraints are then fed into SLAM back-end. We use incremental smoothing and mapping (iSAM) [27] to solve for optimal trajectory given the odometry and sonar constraints.

III. RESULTS

In this section, we evaluate the proposed sensor model using both synthetic and tank test data. A seafloor mapping mission is simulated with synthetic data by adjusting the floor flatness. For the tank test, sensor system is equipped with navigation sensors and DIDSON. For both datasets, a proposed sensor model is applied to provide sensor measurements to SLAM frameworks.

A. Synthetic Data Validation

A synthetic dataset consists of 3-D points on the floor which are then mapped into two synthetic sonar images. In the generation of synthetic data, we control the flatness of the ground plane while the vehicle performs seafloor mapping following a circular trajectory. The corresponding keypoints are synthetically generated using field of view (FOV) and vehicle poses as described in Fig. 5. We randomly sample 30 correspondences (red dots in Fig. 5(b)). The final synthetic data consists of 140 consecutive image frames.

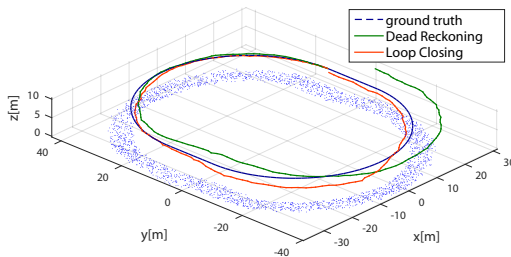


Fig. 6. Comparison of two trajectories. The red line represent the iSAM optimized trajectory with a sonar loop-closure, and the green line is result based on dead-reckoning alone. The ground truth is the blue line.

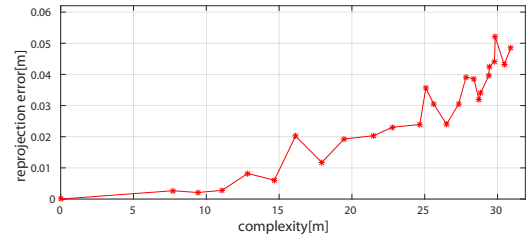


Fig. 7. The performance evaluation with respect to environment complexity. The reprojection error is evaluated as increasing the peak of seafloor (complexity).

Lastly, we fed these correspondences into the two-view image registration introduced in §II. The relative sonar measurements are then optimized in the SLAM back-end. To validate our sensor model in a SLAM framework, we corrupt synthetic navigation data with noise. Fig. 6 shows the comparison between trajectory with and without a sonar loop-closure. Odometry measurements tracking a circular path of path length 210 m is generated. The SLAM trajectory is consistent with the ground truth while the dead-reckoned (DR) trajectory diverges. More quantitative comparison can be found in Table I where the SLAM trajectory shows 10 times smaller error than the DR.

B. Sensitivity to Floor Flatness

In this paper, we assume a floor mapping scenario. To evaluate our assumption of flat floor plane, this section provides a short discussion on the pose estimation accuracy with respect to the floor flatness. In two-view bundle adjustment from the synthetic data, we evaluate reprojection error by changing the seafloor complexity.

Fig. 7 illustrates the change of the reprojection error as increasing the peak of seafloor height. The reprojection error

TABLE I
THE MAXIMUM TRAJECTORY ERROR OF SYNTHETIC DATA EXPERIMENT

	Dead Reckoning	SLAM (Loop closing)
maximum error (m)	11.733	1.098

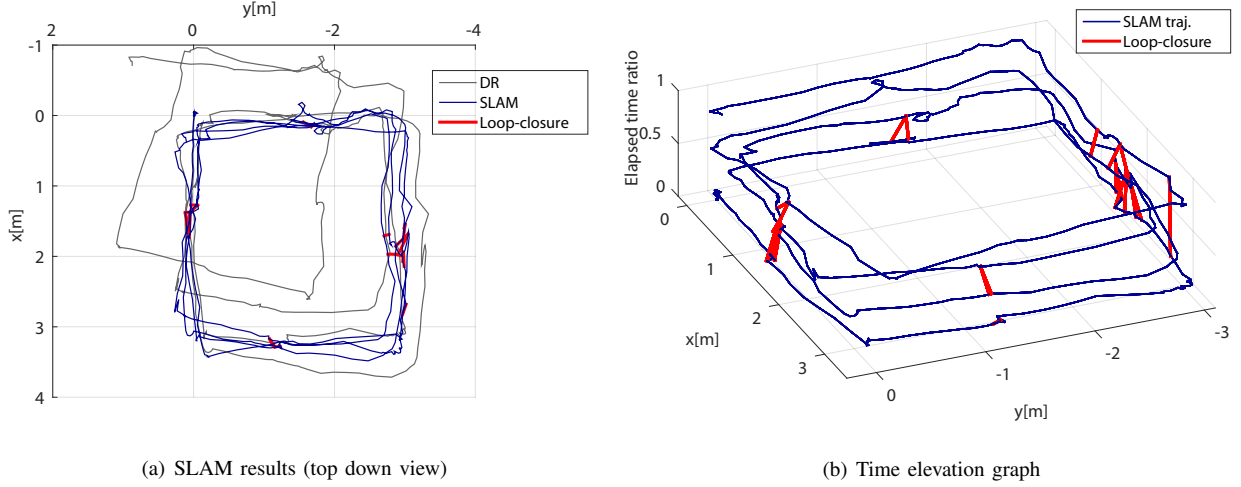


Fig. 8. SLAM and time elevation graph from tank test. (a) The blue line in the figure indicate the SLAM trajectory and the green line indicate DR result. The red links indicate the constraints derived from non-temporally sequential sonar image pairs. (b) The vertical axis indicates elapsed time ratio. The SLAM trajectory error is six times smaller than the DR

increases along with complexity due to assuming a planar surface model. We further examine the estimation sensitivity by analyzing the motion estimation. Table II describe uncertainty of the motion estimation of synthetic data. We investigate the uncertainty under a sample motion with a variance of seafloor height of 0.3m. The error shows somewhat differences depending on the type of motion and the direction. The rotational uncertainty is smaller in z direction than roll and pitch; there is less translation error in x - y direction than in z direction. This can be intuitively understood as the range and azimuth of the 3-D point are directly obtainable, and this direct information affects the estimation with elevation ambiguity.

TABLE II
THE MOTION ESTIMATION UNCERTAINTY IN SYNTHETIC DATA, INPUT TRANSLATION $\mathbf{T} = [1.5, 1.5, 1.5](m)$ AND ROTATION $\mathbf{R} = [3, 3, 3](deg)$, COMPLEXITY IS (30 CM)

	translation (m)	rotation (deg)
std of x directional	0.1046	2.8920
std of y directional	0.1162	2.2126
std of z directional	0.4989	0.7189

C. Tank Test Validation

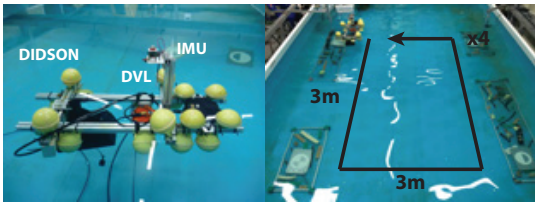


Fig. 9. Experimental prototype used for SLAM application (left) and trajectory that the robot followed (right)

To evaluate navigation performance using imaging sonar-based SLAM, we constructed a prototype with sensor packages

that takes the form of AUV in Fig. 9. To acquire navigation data, the robot is equipped with NavQuest600 Micro Doppler velocity log (DVL), Honeywell HG1700 inertial measurement unit (IMU), and 1.8MHz Soundmetrics DIDSON imaging sonar. The odometry measurements are generated using these navigation sensors.

The tank (as in Fig. 9 (right)) is 3m by 3m with woods and rocks deployed on the floor. For the tank test validation, the platform moves following the square trajectory four times. The system moves on the surface maintaining the sensor system maneuver at a fixed depth. The odometry from DVL suffered from shallow water depth and a walled environment, and we incorporated sonar measurements to correct navigation drift. Despite of poor performance DVL, the geometric constraint from a pair of sonar images can guarantee more superior navigation quality. Whenever the sensor systems returns to the previously traveled track, two-view image registration is performed to provide loop-closure for SLAM.

The resulting SLAM trajectory is depicted in Fig. 8(a). It provides a comparison of SLAM trajectory estimate and DR. As shown in Fig. 8(a), the SLAM trajectory is a significant improvement over the DR trajectory. For easier visualization, Fig. 8(b) depicts a time elevation graph of SLAM trajectory with the constraint from image pairs.

Fig. 10 shows the change of pose uncertainty with respect to the mission time. We adopt a metric of pose uncertainty measure ($|\Sigma_{rr}|$) in [28]. As time goes on, the uncertainty of the pose using DR monotonically increases, whereas the SLAM result shows a sawtooth pattern with bounded uncertainty due to the constraints from sonar measurements. The increase and decrease of uncertainty is repeated whenever the sensor systems meet abundant features in a previously traveled region.

IV. CONCLUSION

This paper presented imaging sonar-based SLAM for seafloor mapping application. For two-view sensor measure-

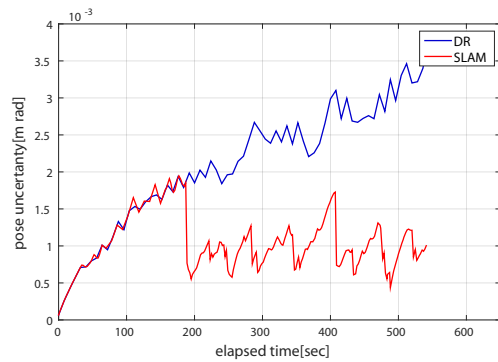


Fig. 10. The pose uncertainty with respect to mission time. The uncertainty of DR (blue line) shows monotonically increase, while the uncertainty of SLAM (red line) is bounded.

ments for imaging sonar, we introduce a KAZE feature that extracts features from nonlinear scale space. This allows us to extract robust features, especially when an image is low resolution with high noise. We found that KAZE features provide more reliable feature correspondences for two-view bundle adjustment. The resulting correspondences are refined using RANSAC, then piped into a two-view bundle adjustment algorithm. We use a sensor measure model derived for seafloor mapping to assume a flat floor. The proposed sensor measurement model is validated by using both synthetic and tank test dataset while a vehicle is performing seafloor mapping. Future work is aimed at improving the sensor model for a floor with high variation. We plan to focus on estimating distance from other sensors, such as DVL in order to remove elevation ambiguity.

ACKNOWLEDGMENTS

This work is supported through a grant from the Korea Research Institute Ship and Ocean engineering (KRISO) (Award #N04150009).

REFERENCES

- [1] F. S. Hover, R. M. Eustice, A. Kim, B. Englot, H. Johannsson, M. Kaess, and J. J. Leonard, "Advanced perception, navigation and planning for autonomous in-water ship hull inspection," *International Journal of Robotics Research*, vol. 31, no. 12, pp. 1445–1464, Oct. 2012.
- [2] P. Ridao, M. Carreras, D. Ribas, and R. Garcia, "Visual inspection of hydroelectric dams using an autonomous underwater vehicle," *Journal of Field Robotics*, vol. 27, no. 6, pp. 759–778, 2010, special Issue: State of the Art in Maritime Autonomous Surface and Underwater Vehicles, Part 1.
- [3] E. Gustafson, B. Jalving, ystein Engelhardtson, and N. Burchill, "HUGIN 1000 Arctic class AUV," in *The Polar Petroleum Potential Conference & Exhibition*, 2011, pp. 714–721.
- [4] C. Kunz, C. Murphy, H. Singh, C. Pontbriand, R. A. Sohn, S. Singh, T. Sato, C. Roman, K. Nakamura, M. Jakuba, R. Eustice, R. Camilli, and J. Bailey, "Toward extraplanetary under-ice exploration: Robotic steps in the Arctic," *Journal of Field Robotics*, vol. 26, no. 4, pp. 411–429, Apr. 2009.
- [5] S. Williams and I. Mahon, "Simultaneous localisation and mapping on the Great Barrier Reef," in *Proceedings of the IEEE International Conference on Robotics and Automation*, vol. 2, 2004, pp. 1771–1776.
- [6] R. M. Eustice, H. Singh, J. J. Leonard, and M. R. Walter, "Visually mapping the RMS Titanic: Conservative covariance estimates for SLAM information filters," *International Journal of Robotics Research*, vol. 25, no. 12, pp. 1223–1242, 2006.
- [7] L. Medagoda, S. B. Williams, O. Pizarro, and M. V. Jakuba, "Water column current profile aided localisation combined with view-based SLAM for autonomous underwater vehicle navigation," in *Proceedings of the IEEE International Conference on Robotics and Automation*, 2011, pp. 3048–3055.
- [8] A. Kim and R. M. Eustice, "Real-time visual SLAM for autonomous underwater hull inspection using visual saliency," *IEEE Transactions on Robotics*, vol. 29, no. 3, pp. 719–733, 2013.
- [9] S. Negahdaripour, "Visual motion ambiguities of a plane in 2-d fs sonar motion sequences," *Comput. Vis. Image Underst.*, vol. 116, no. 6, pp. 754–764, Jun. 2012.
- [10] —, "On 3-d motion estimation from feature tracks in 2-d fs sonar video," vol. 29, no. 4, pp. 1016–1030, Aug 2013.
- [11] S. Negahdaripour, H. Assalih, Y. Petillot, and L. Brisson, "Performance and accuracy in visual motion computation from fs sonar video sequences," in *Proceedings of the IEEE/MTS OCEANS Conference and Exhibition*, Sept 2010, pp. 1–7.
- [12] N. Hurtos, X. Cufí, Y. Petillot, and J. Salvi, "Fourier-based registrations for two-dimensional forward-looking sonar image mosaicing," in *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on*, Oct 2012, pp. 5298–5305.
- [13] M. D. Aykin and S. Negahdaripour, "On feature matching and image registration for two-dimensional forward-scan sonar imaging," *Journal of Field Robotics*, vol. 30, no. 4, pp. 602–623, 2013. [Online]. Available: <http://dx.doi.org/10.1002/rob.21461>
- [14] S. Negahdaripour, P. Firoozfam, and P. Sabzmejdani, "On processing and registration of forward-scan acoustic video imagery," in *Computer and Robot Vision, 2005. Proceedings. The 2nd Canadian Conference on*, May 2005, pp. 452–459.
- [15] H. Johannsson, M. Kaess, B. Englot, F. Hover, and J. Leonard, "Imaging sonar-aided navigation for autonomous underwater harbor surveillance," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, Taipei, Taiwan, Oct. 2010.
- [16] K. Kim, N. Neretti, and N. Intrator, "Mosaicing of acoustic camera images," *Radar, Sonar and Navigation, IEE Proceedings -*, vol. 152, no. 4, pp. 263–270, Aug 2005.
- [17] P. Ozog, G. Troni, M. Kaess, R. M. Eustice, and M. Johnson-Roberson, "Building 3D mosaics from an autonomous underwater vehicle, doppler velocity log, and 2d imaging sonar," in *Proceedings of the IEEE International Conference on Robotics and Automation*, Seattle, WA, USA, May 2015, pp. 1137–1143.
- [18] P. F. Alcantarilla, J. Nuevo, and A. Bartoli, "Fast explicit diffusion for accelerated features in nonlinear scale spaces," in *British Machine Vision Conf. (BMVC)*, 2013.
- [19] D. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.
- [20] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Speeded-up robust features (SURF)," *Computer Vision and Image Understanding*, vol. 110, no. 3, pp. 346–359, 2008.
- [21] P. F. Alcantarilla, A. Bartoli, and A. J. Davison, "KAZE features," in *Eur. Conf. on Computer Vision (ECCV)*, 2012.
- [22] S. Wang, H. You, and K. Fu, "BFSIFT: A novel method to find feature matches for sar image registration," *Geoscience and Remote Sensing Letters, IEEE*, vol. 9, no. 4, pp. 649–653, July 2012.
- [23] M. Aykin and S. Negahdaripour, "On feature extraction and region matching for forward scan sonar imaging," in *Proceedings of the IEEE/MTS OCEANS Conference and Exhibition*, Oct 2012, pp. 1–9.
- [24] H. Johannsson, M. Kaess, B. Englot, F. Hover, and J. J. Leonard, "Imaging sonar-aided navigation for autonomous underwater harbor surveillance," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, Taipei, Taiwan, Oct. 2010, pp. 4396–4403.
- [25] S. Negaharipour, "On 3-d scene interpretation from f-s sonar imagery," in *Proceedings of the IEEE/MTS OCEANS Conference and Exhibition*, Oct 2012, pp. 1–9.
- [26] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2000.
- [27] M. Kaess, H. Johannsson, and J. Leonard, "Open source implementation of iSAM," <http://people.csail.mit.edu/kaess/isam>, 2010.
- [28] A. Kim and R. M. Eustice, "Active visual SLAM for robotic area coverage: Theory and experiment," *International Journal of Robotics Research*, vol. 34, no. 4-5, pp. 457–475, 2015.