

Building 3D Mosaics from an Autonomous Underwater Vehicle, Doppler Velocity Log, and 2D Imaging Sonar

Paul Ozog, Giancarlo Troni, Michael Kaess, Ryan M. Eustice, and Matthew Johnson-Roberson

Abstract—This paper reports on a 3D photomosaicing pipeline using data collected from an autonomous underwater vehicle performing simultaneous localization and mapping (SLAM). The pipeline projects and blends 2D imaging sonar data onto a large-scale 3D mesh that is either given *a priori* or derived from SLAM. Compared to other methods that generate a 2D-only mosaic, our approach produces 3D models that are more structurally representative of the environment being surveyed. Additionally, our system leverages recent work in underwater SLAM using sparse point clouds derived from Doppler velocity log range returns to relax the need for a prior model. We show that the method produces reasonably accurate surface reconstruction and blending consistency, with and without the use of a prior mesh. We experimentally evaluate our approach with a Hovering Autonomous Underwater Vehicle (HAUV) performing inspection of a large underwater ship hull.

I. INTRODUCTION

Several tasks in ocean exploration require the visualization of a large set of images to understand underwater phenomena at a broad spatial scale. In recent years, techniques have been developed that allow for the reconstruction of visually rich 3D mosaics of the seafloor from thousands of optical images [1, 2]. Though these methods have been successfully applied in large-scale marine environments, underwater optical cameras have several limitations. For example, turbid waters make identification of visual features difficult, light attenuates much more in water than in air, and underwater cameras typically must provide their own light source. Acoustics are the preferred sensor modality for underwater robotics because they overcome several of those limitations. However, there are several challenges with this approach that need to be solved. These challenges include: (i) sonar technology is typically more expensive than optical cameras, (ii) the sensor’s vantage point strongly affects signal intensity, and (iii) high field of view (FOV) imaging sonars have non-standard geometrical properties of their projection from 3D to 2D. This paper will explore the third challenge: we present a method to create a textured 3D mosaic from an

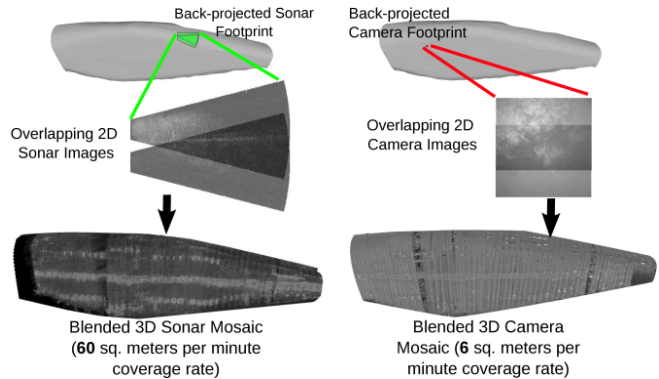


Fig. 1. In the context of ship hull inspection, one key benefit of our sonar-based mosaicing pipeline (left) over our camera-based mosaicing pipeline (right) is coverage rate. In this example, a 3D surface, shown above in gray, is either reconstructed from a typical underwater SLAM sensor payload, or given as a prior CAD model. Two overlapping sonar images back-project onto the 3D surface in the green region, which has a much larger footprint than an underwater camera’s, shown in red. Mesh faces within the overlapping region are blended to avoid the presence of seams in the final mosaic.

imaging sonar, coupled with a typical sensor payload on a small autonomous underwater vehicle (AUV). An overview of our method is illustrated in Fig. 1.

A. Related Work

In terms of attenuation, sonar is by far the preferred sensor for surveying the ocean seafloor [3], and imaging sonars in particular are a popular alternative to underwater cameras. Recent work related to imaging sonars has focused on the registration problem, where two overlapping images are geometrically matched. Solutions to this problem use either spectral methods [4, 5], or feature-based methods [6]. In either case, these techniques have several applications, such as underwater simultaneous localization and mapping (SLAM) [7, 8] and photomosaicing [9, 10]. However, previous work produces strictly 2D mosaics or 3-degree of freedom (DOF) motion estimates (relative x , y , and heading). Recently, Negahdaripour lifted feature tracking and motion estimation to 3D, but did not explore the applications to 3D mosaicing [11]. The paper additionally improves the results of full 3D pose estimation by zeroing higher dimensional parameters (as in the 3-DOF case).

Despite their limitations in marine environments, optical cameras have become a popular modality for the creation of underwater mosaics. Historically, most applications have used 2D mosaics for vision-aided navigation [12]. Johnson-Roberson et al. [2] argue that the rugged terrain of the

*This work was supported in part by the Office of Naval Research under awards N00014-12-1-0092 and N00014-14-1-0373, and in part by the American Bureau of Shipping under award number N016970-UM-RCMOP.

P. Ozog is with the Department of Electrical Engineering & Computer Science, University of Michigan, Ann Arbor, MI 48109, USA paulozog@umich.edu.

G. Troni is with the Monterey Bay Aquarium Research Institute, Moss Landing, CA 95039, USA gtroni@ing.puc.cl.

M. Kaess is with the Robotics Institute, Carnegie Mellon University, Pittsburgh, PA 15213, USA kaess@cmu.edu.

R. Eustice and M. Johnson-Roberson are with the Department of Naval Architecture & Marine Engineering, University of Michigan, Ann Arbor, MI 48109, USA {eustice,mattjr}@umich.edu.

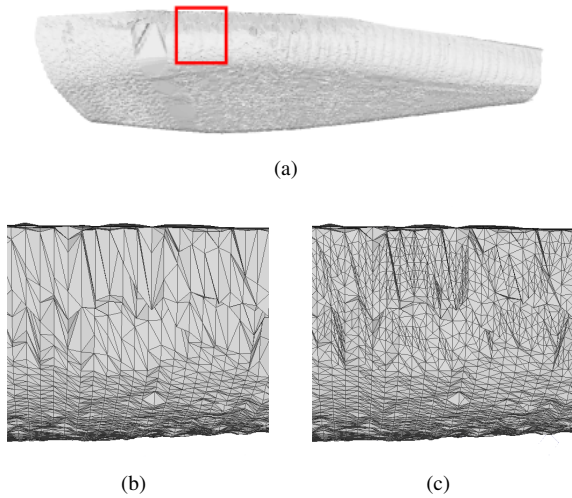


Fig. 2. A simple midpoint resampling method prevents pixel stretch for the Delaunay reconstruction of the mesh shown in (a). In (b), there are triangles that are ill-suited for texture blending. By recursively splitting the edge of these triangles at their midpoint, shown in (c), these triangles are divided into smaller faces while preserving the overall shape of the mesh.

seafloor necessitates projecting the imagery onto 3D models to properly account for the geometry, rather than force-fitting a plane to a non-planar environment.

Similar reasoning suggests that 3D mosaicing methods are a better choice for building mosaics of large man-made structures, which include dams, harbors, pipelines, and ship hulls [13]. We are particularly interested in autonomous ship hull inspection, and we believe that 3D mosaicing will be of great benefit for improved efficiency in maintenance, assessment, and security. 3D mosaics would help robots’ human supervisors to easily visualize the data, assist in cooperation between robots, or aid in automated tracking of structural changes or anomalies over time. Novel techniques for the generation of acoustic 3D mosaics is therefore the focus of this paper.

B. Outline

This paper is organized as follows. In §II we describe our 3D mosaicing approach, where a surface mesh is reconstructed from SLAM-corrected poses, and an empirical reprojection operation is used to assign images to triangles for texturing. In §III we describe our experimental setup and we offer several evaluations of the method’s performance. §IV summarizes and offers concluding remarks.

II. APPROACH

A. Correcting Navigation Drift with SLAM

A prerequisite of our 3D mosaicing pipeline is that the vehicle’s trajectory is already corrected from SLAM. There are several methods to accomplish this. Our past work primarily focused on camera-based techniques [14] but our recent work has shifted some attention on relaxing the reliance of an underwater camera [15]. In particular, by estimating surface normals from Doppler velocity log (DVL) range returns, we can constrain the normals of nearby planar patches and

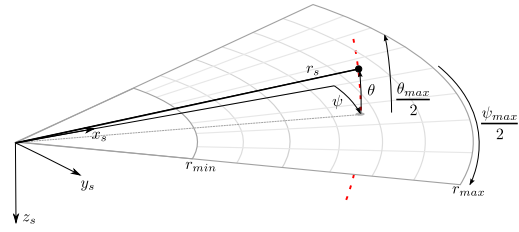


Fig. 3. This figure shows how a 3D point is projected into the DIDSON frame from its spherical coordinates. All points not contained in the volume bounded by r_{min} , r_{max} , θ_{max} , and ψ_{max} are not visible to the sonar.

produce more self-consistent maps. This approach also has tremendous benefits for performing long-term SLAM since it can be effectively combined with recent developments in graph sparsification techniques [16, 17].

This method, which we will call “piecewise-planar SLAM”, is of particular relevance to sonar mosaicing because it can be generalized to other AUVs that do not have a camera. Furthermore, the geometrical information provided by planes is beneficial to other applications besides ship hull inspection, such as underwater trenches or dam inspection.

B. Surface Reconstruction

The first step in creating a 3D mosaic is to estimate a surface reconstruction. Traditionally, this either relied on a stereo camera to merge individual keyframe meshes into a large 3D model [18], or from camera-derived point sets that contain large amounts of features [19]. Since we are interested in mosaics derived from imaging sonar, our work instead reconstructs the surface using the DVL range returns. By linearly interpolating a DVL-derived 3D point cloud, we can create suitable models of ship hulls.

By converting range returns from all DVL poses into a point cloud in the global frame, we can linearly interpolate the z -coordinates of the points over an evenly-spaced grid in the global frame’s x, y plane. We apply Delaunay triangulation to these points, producing a height-map. A well-known limitation with this technique is that it exposes stretched triangles in the near-vertical portions of the surface that will not fit in a single camera or sonar image. To mitigate this effect, we recursively inscribe triangles within triangles until all edge lengths in the mesh are below a threshold (0.5 m for the results shown in §III). We provide an illustration of this so-called “midpoint method” in Fig. 2.

Due to the sparsity of the DVL returns, a simple linear interpolation may not generalize to some non-ship hull applications. However, more advanced techniques exist that can interpolate DVL returns in more varied structure, like underwater terrain [20].

C. Projecting Mesh Vertices into Image Coordinates

When the surface is reconstructed, we must then project each vertex into the sonar images where the vertex is visible. Unlike a calibrated projective camera, which has a simple closed-form expression for the projection from 3D to 2D pixel coordinates, we empirically compute this relationship using several sensor-specific parameters.

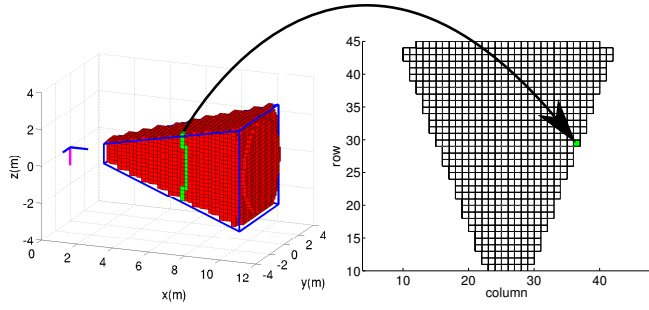


Fig. 4. This diagram illustrates the process of reprojecting 3D points into a pixel value in a sonar image. We discretize the Cartesian volume contained in the sonar’s field-of-view frustum as voxels. These voxels are mapped into a corresponding pixel in the 2D image with a simple look-up table. This lookup table is computed in the sonar’s frame, so it only must be stored once. In this example, all 3D points contained in the green arc of voxels (left) map to the single green pixel in the sonar’s 8-bit image (right). We over-pixelated this diagram for the sake of clarity.

The imaging sonar has a discrete number of beams, N_b , each containing a discrete number of ranges, N_r , as illustrated in Fig. 3. Let r_{min} and r_{max} be the minimum and maximum ranges that are observable by the sonar. For a given u, v pixel coordinate in the sonar image (such as the synthetic one in Fig. 4, right), the corresponding sensor-frame Cartesian coordinates, x_s, y_s , are given by:

$$x_s = \frac{u - \frac{w}{2}}{\gamma}$$

$$y_s = r_{max} - \frac{v}{\gamma},$$

where w is the sonar image width, h is the height, u is the column index of the pixel, v is the row index of the pixel, and

$$\gamma = \frac{w}{2r_{max} \sin(\psi_{max}/2)}$$

is a constant.

We convert these sensor-frame coordinates into range and bearing values in the sensor frame (assuming that $z = 0$) as follows:

$$r_s = \sqrt{x_s^2 + y_s^2}$$

$$\psi_s = \text{atan2}(x_s, y_s)$$

Finally, we assign these continuous range and bearings to one of the sensor’s discrete range bins and beam number:

$$n_b = \frac{(r_s - r_{min})(N_r - 1)}{r_{max} - r_{min}}$$

$$n_r = M_4(\psi_s)^\top \mathbf{a},$$

where n_b is the beam number, n_r is the range bin number, $M_4(\psi) = [1, \psi, \psi^2, \psi^3]$ is a fourth-degree vector of monomial bases and \mathbf{a} is a vector of sensor-specific coefficients provided by the manufacturer. To compute the inverse mapping from n_r and n_b to 3D Cartesian coordinates, we simply discretize the volume inside the sonar’s frustum into voxels, apply the above set of equations, and store the inverse map. For a given x_s, y_s coordinate, the voxels where $z \neq 0$ project

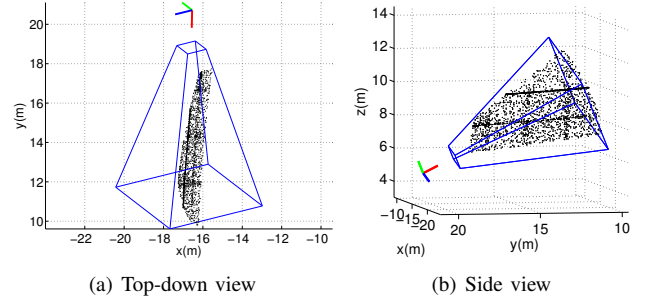


Fig. 5. Example subset of mesh vertices within an imaging sonar’s FOV. The frustum, computed from the method in §II-C, is shown in blue. Vertices from the surface reconstruction that lie within the frustum are visible to the sonar, and will be included in the blending step. These vertices are shown as black dots.

to the same pixel as the corresponding voxel where $z = 0$. This is illustrated in Fig. 4.

D. Blending Step

Our image blending pipeline is based on the previous work by Johnson-Roberson et al. [2] for creating large-scale 3D mosaics of seafloor environments using a stereo camera. The approach works by assigning a fixed number of sonar images to every face in the mesh (for the experimental results, shown in §III, we use a maximum of four images per face). For each mesh face, we compute the set of sonar poses such that the face is visible. Fig. 5 shows an example of which face vertices are visible. From this set, we pick the four best views of the face using a user-defined proxy for image quality. For underwater cameras, popular heuristics include choosing the smallest distance of the projected face to the center, or choosing the most orthogonal camera poses to the face’s surface normal.

Choosing the correct heuristic for an imaging sonar is a subjective matter, and we have found from our experimental results that picking images where the face projection is closest to the pixel coordinates $u_{best} = w/2$ and $v_{best} = 3h/4$ works well. For the ship hull mosaics presented in §III, we have found this typically corresponds to a face normal of approximately 7 degrees from orthogonal to the sonar frame’s x -axis.

Once we determine the four best image patches, we weight each pixel contained in the i^{th} mesh face by the distance r to the pixel coordinate (u_{best}, v_{best}) . This weight is determined from the expression

$$r = \sqrt{(u_{best} - u)^2 + (v_{best} - v)^2}$$

$$B_k^i(r) = \frac{e^{-k \frac{r}{R}}}{1 + e^{-2k \frac{r}{R}}},$$

where R is a reference distance (typically the maximum distance to be considered). This process is determined for three different resolution bands, where $k = 5, 10, 50$ is an appropriately chosen coefficient for each resolution band. Larger k indicates a sharper drop-off as r increases, and is useful for bands with higher resolution. During the actual

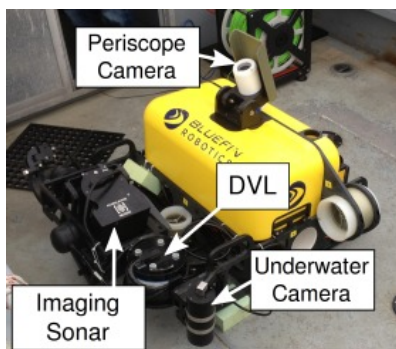


Fig. 6. The HAUV vehicle platform and sensor suite used in the experimental evaluation of our method. The HAUV is equipped with a DVL for hull-relative navigation, a periscope camera to globally localize to previous SLAM graphs, and a Sound Metrics DIDSON imaging sonar for collecting imaging data. Though the HAUV is equipped with an underwater camera, this was only used for evaluation purposes and our method does not require its use.

pixel blending computation, the four blending weights are normalized so that they sum to one [2].

III. FIELD EXPERIMENTAL EVALUATION

A. Robot Platform

We use data collected from the Hovering Autonomous Underwater Vehicle (HAUV) [21, 22] to experimentally evaluate our method for creating 3D mosaics. The imaging sonar used on the HAUV is a Sound Metrics Dual frequency Identification SONar (DIDSON) [23]. Other relevant sensors on the HAUV are shown in Fig. 6. Though the HAUV has an underwater stereo camera, it was only used for evaluation purposes and our method does not require it [15]. The periscope camera, however, was used to globally localize successive surveys into a single common hull-relative reference frame.

The datasets used in this section were collected on the 180 m *SS Curtiss* vessel in March 2014. The mosaics presented in §III-C consist of eight individual surveys, all aligned to a common hull-relative frame of reference. Sample mosaics, with and without the use of a prior computer aided design (CAD) model, are shown in Fig. 8.

B. Evaluated SLAM Techniques

As discussed in §II, the input to our 3D mosaicing pipeline is a set of SLAM-corrected sonar poses. We applied our 3D mosaicing pipeline using four different techniques of varying computational efficiency and practicality to assess our approach in each setting. In particular, we investigate using (i) piecewise-planar surface SLAM, (ii) piecewise-planar surface constraints with underwater camera constraints, (iii) bundle-adjusted underwater camera poses, and (iv) bundle-adjusted underwater camera poses rigidly aligned to a CAD model. We provide more details for each method in the following sections.

1) *Piecewise-planar SLAM*: This approach models the ship hull as a collection of many planar features. Co-registered planar patches are constrained so that their normals

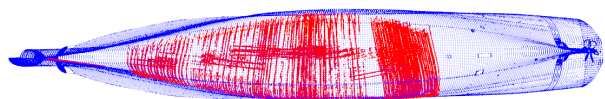


Fig. 7. Rigid alignment between CAD model vertices (shown in blue) and DVL point cloud computed from the bundle-adjusted underwater camera poses (shown in red). To perform the alignment, we use the outlier rejection approach described in §II-B and find the optimal rigid transformation using the GICP algorithm [28].

are similar, but not to the extreme that the curvature of the hull is lost in the representation. Essentially, this deviation measures the orthogonal distance from a point on one mesh to the closest face on the mesh to which it is being compared. This method has the most practical significance since it does not require an underwater camera for navigation correction.

Despite not relying on an underwater camera, this technique is applied across multiple SLAM sessions, where the initial global localization must be determined with a periscope camera. For the HAUV application, a global positioning system (GPS) will not suffice since we desire the SLAM maps to be expressed in a hull-relative reference frame. This requirement can be relaxed for other applications, where GPS or beacons provide world-frame localization.

2) *Piecewise-planar SLAM with underwater camera*: A major benefit of the piecewise-planar SLAM technique discussed in the previous section is its usability in an underwater visual SLAM framework. In particular, we explore the use of piecewise planar SLAM techniques with the saliency-informed visual SLAM approach developed by Kim and Eustice [14]. This method does not require full camera coverage, but can still constrain navigational drift even if the space between survey tracks results in low or zero image overlap.

3) *Bundle adjustment using underwater camera*: This approach minimizes the reprojection error of visual features in a calibrated camera [24]. In our application, we use a stereo underwater camera with 100% coverage of the hull surface. This approach is computed offline, and uses scale-invariant feature transform (SIFT) features to make visual correspondences [25]. Outliers are rejected by using random sample consensus (RANSAC) [26] with a least-squares point cloud alignment algorithm made popular by Arun et al. [27]. Along with optimizing a sparse set of SIFT features, we also include odometry measurements from the DVL and absolute constraints on depth, pitch, and roll from the HAUV's depth sensor and inertial measurement unit (IMU).

This approach is impractical for sonar mosaicing because it relies on 100% camera coverage, which is time-consuming and not possible in turbid water. However, we include this method so that we can compare its image blending consistency against the more practical approaches, discussed above.

4) *Rigidly aligning bundle-adjusted poses to CAD model*: For certain applications, like underwater ship hull inspection, a CAD model of the surveyed vessel may be available.

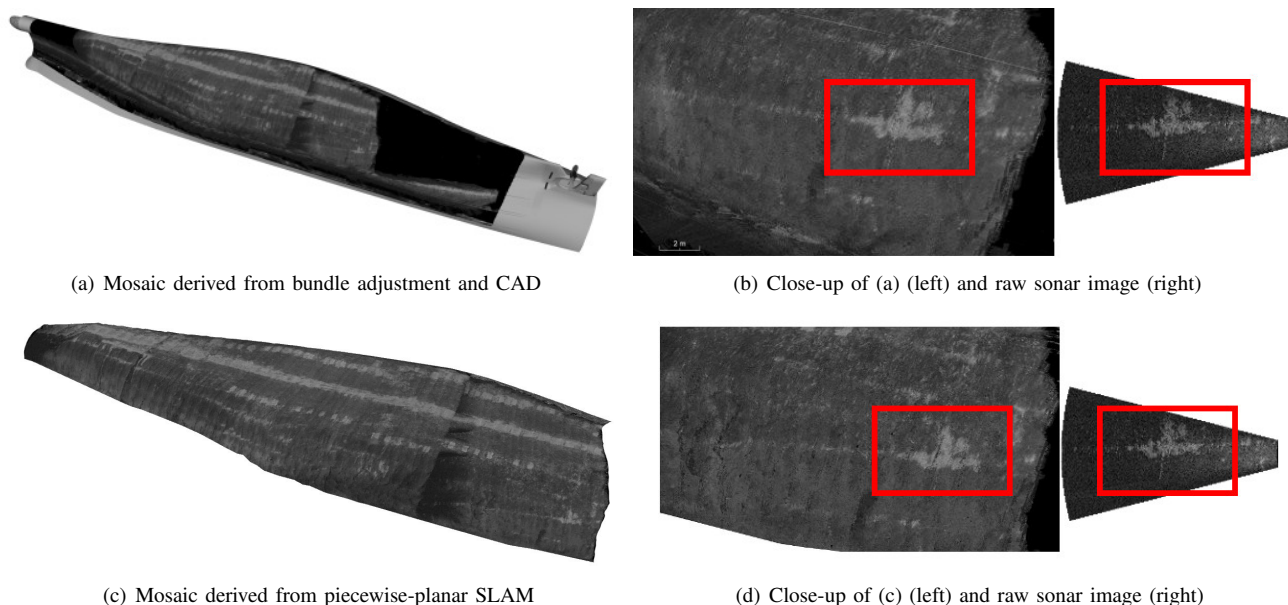


Fig. 8. Qualitative results of imaging sonar mosaics. The CAD model was used to generate (a) and (b). Even without a prior CAD model, as shown in (c) and (d), we can still produce a 3D mosaic that appears nearly identical in texture to the model from (a). The most apparent difference is that the model in (c) is smaller. The HAUV was not always able to measure valid DVL returns near the water surface. This explains why those portions of the model are missing.

TABLE I
“MESH ATTRIBUTE DEVIATION” SCORES OF 3D SURFACE RECONSTRUCTIONS COMPARED TO GROUND-TRUTH *SS Curtiss* CAD MODEL

METHOD	MIN. ABS. ERROR (m)	MAX. ABS. ERROR (m)	MEAN (m)	RMS ERROR (m)
CAD Model	—	—	—	—
Bundle Adjusted	1.56×10^{-5}	1.37	0.22	0.29
Piecewise-Planar+Visual SLAM	3.76×10^{-7}	1.03	0.18	0.23
Piecewise-Planar SLAM	9.53×10^{-7}	1.04	0.19	0.24

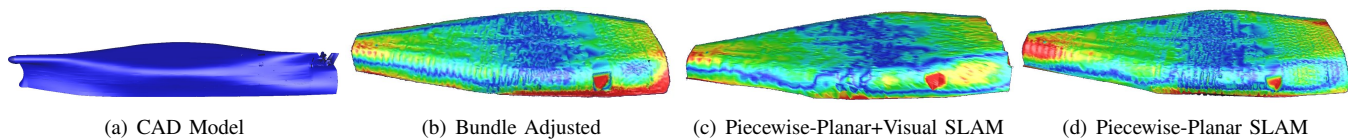


Fig. 9. False-color visualization of the results tabulated in Table I. Blue regions indicate little deviation, while red regions indicate substantial error. Blue-to-red represents a change of 70 cm. The RMS error for bundle adjustment is actually the highest because the ship hull surface itself is left out of the optimization. However, (c) and (d) account for this and therefore have less error compared to the CAD model.

Having ground truth is a unique capability in marine robotics, especially for field experiments. We therefore draw attention to using bundle adjustment to take full advantage of this prior CAD model as a way to assess the quality of the proposed technique. In particular, the CAD model makes it possible to analyze the structural similarity of our 3D sonar mosaics to ground truth. It should be noted that this method’s reliance on camera-based bundle adjustment makes it impracticable for mosaicing with an imaging sonar.

For this approach, we substitute the CAD model surface in place of the one derived from the SLAM methods described previously. We align the reference frame of the CAD model to the SLAM reference frame using the well-known Generalized iterative closest point (GICP) algorithm [28]. The alignment between the bundle-adjusted DVL point cloud and

SS Curtiss CAD model is shown in Fig. 7.

C. 3D Mosaic Quality

We measure the quality of the 3D mosaic in two different ways. First, we consider the structural deviation from the mosaic’s surface to the ground truth CAD model. Second, we measure the consistency of the images used in the blending step described in §II-D.

For each of the SLAM techniques described in §III-B, we evaluate the structural similarity of the 3D surface to the ground truth CAD model. This method uses the “attribute deviation metric” developed by Roy et al. [29]. The false color visualization of this metric is shown in Fig. 9, and the results are tabulated in Table I.

The mesh deviation results demonstrate that our method can produce accurate models without the use of an underwa-

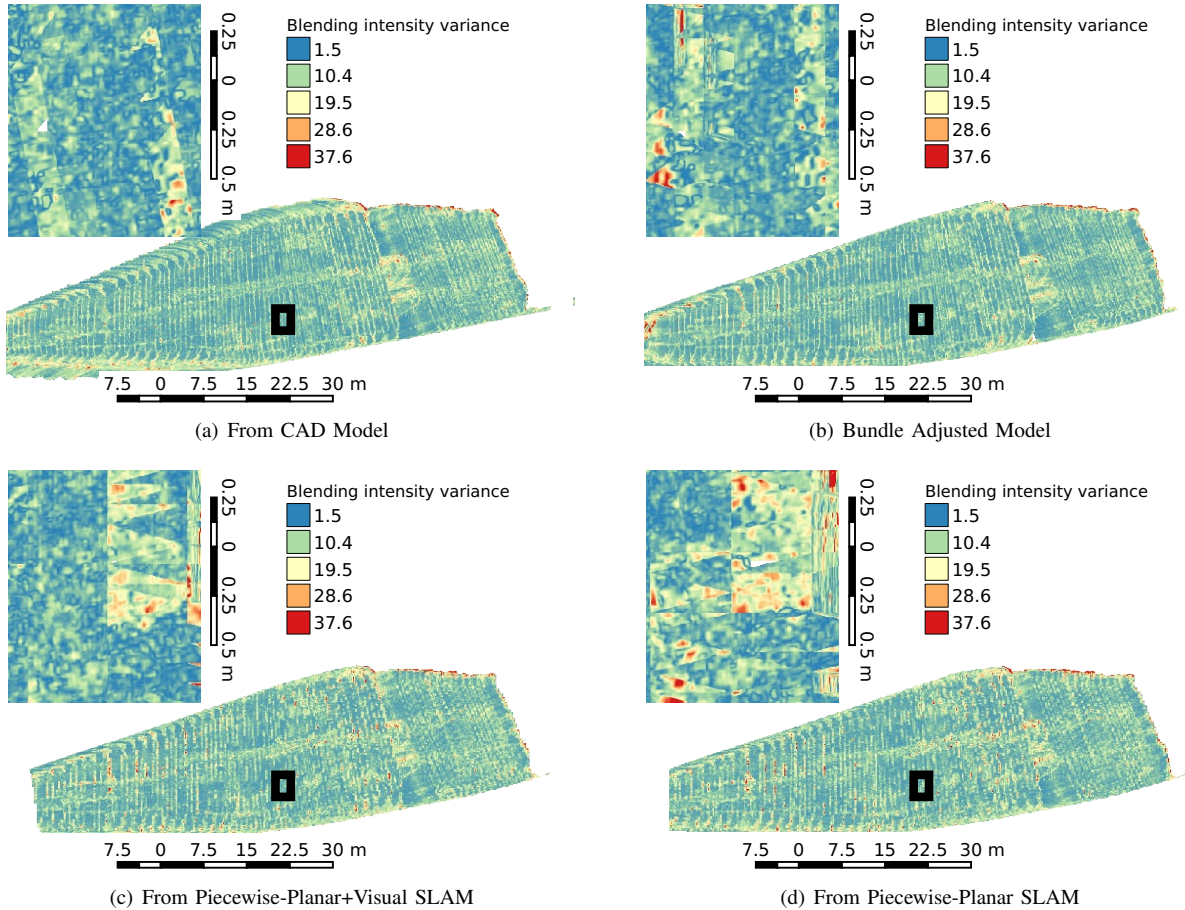


Fig. 10. These figures encode the variance of weighted pixel intensities used during the blending step. Blue values denote consistent pixel intensities, while red shows relatively large variation. For certain small-scale regions, the blending consistency is noticeably compromised if the poses are not bundle adjusted. As a whole, however, the large-scale features are captured in all cases, as shown in Fig. 8.

ter camera. Indeed, the methods leveraging our prior work with piecewise-planar SLAM out-performs bundle adjustment. This should not be surprising because our SLAM technique includes the ship hull surface itself as part of the optimization. Bundle adjustment, as implemented for this comparison, only uses visual feature correspondence to constrain the SLAM estimate.

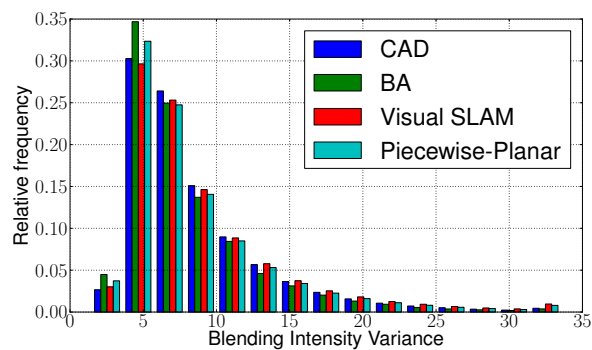
In addition to the results shown in Fig. 9 and Table I, the *texture quality* of our sonar mosaics will clearly be sensitive to the quality of poses from SLAM. To quantify this, we computed the variance of weighted pixel intensities during the image blending step described in §II-D for every pixel in the output mesh. These pixel intensities range from $[0, 255]$, corresponding to the value of an eight-bit unsigned integer. We provide a visualization of these variances in Fig. 10, where we highlight an area that was only correctly blended using a bundle adjustment step.

A histogram of these variances is shown in Fig. 11. Though the results are quite similar between each method, the results taken from poses that were *not* bundle-adjusted with a camera have noticeably heavier tails in the variance distribution (for variances greater than 20). That being said, these results show that, by and large, our mosaic blends

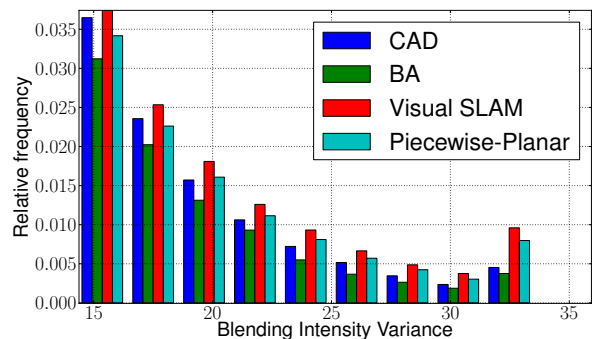
together images with relatively similar pixel intensities. Considering that the globally bundle-adjusted techniques do perform better, this suggests that our method will see improvement if we add a step to globally match features across sonar images and include these constraints in a SLAM framework.

IV. CONCLUSION

In summary, we have demonstrated a novel system to create 3D models using an imaging sonar, a DVL, and a small AUV. We provided a convenient empirical method to project 3D points onto a 2D pixel coordinate in a sonar image. We have shown that our mosaic pipeline can properly handle 3D geometry rather than requiring the environment to be entirely planar. We offered two ways to measure the quality of mosaic: structural deviation from ground truth and variance over pixel intensities that are chosen for blending. We provided quantitative evaluations for our approach using both classic and recently-introduced SLAM techniques, such as bundle adjustment and modeling smoothly curved surfaces as piecewise planar. These evaluations were experimentally analyzed using field data from an AUV performing ship hull inspection.



(a) Full distribution



(b) Tail of distribution

Fig. 11. Histogram of intensity variances for the results shown in Fig. 10. The tails in (b) are purposely clamped to emphasize that SLAM approaches that are not globally bundle adjusted do not perform as well as those that are. Importantly, however, the full histogram from (a) shows that as a whole our mosaic blends together images with relatively similar pixel intensities, even when using the piecewise-planar SLAM technique, which does not rely on constraints from an underwater camera.

REFERENCES

- [1] T. Nicosevici, N. Gracias, S. Negahdaripour, and R. Garcia, "Efficient three-dimensional scene modeling and mosaicing," *J. Field Robot.*, vol. 26, no. 10, pp. 759–788, 2009.
- [2] M. Johnson-Roberson, O. Pizarro, S. B. Williams, and I. Mahon, "Generation and visualization of large-scale three-dimensional reconstructions from underwater robotic surveys," *J. Field Robot.*, vol. 27, no. 1, pp. 21–51, 2010.
- [3] R. D. Ballard, L. E. Stager, D. Master, D. Yoerger, D. Mindell, L. L. Whitcomb, H. Singh, and D. Piechota, "Iron age shipwrecks in deep water off Ashkelon, Israel," *Amer. J. Archaeol.*, pp. 151–168, 2002.
- [4] N. Hurtós, D. Ribas, X. Cufi, Y. Petillot, and J. Salvi, "Fourier-based registration for robust forward-looking sonar in low-visibility underwater environments," *J. Field Robot.*, vol. 32, no. 1, pp. 123–151, 2015.
- [5] H. Bülow and A. Birk, "Spectral registration of noisy sonar data for underwater 3D mapping," *Autonomous Robots*, vol. 30, no. 3, pp. 307–331, 2011.
- [6] M. D. Aykin and S. Negahdaripour, "On feature matching and image registration for two-dimensional forward-scan sonar imaging," *J. Field Robot.*, vol. 30, no. 4, pp. 602–623, 2013.
- [7] H. Johannsson, M. Kaess, B. Englot, F. Hover, and J. Leonard, "Imaging sonar-aided navigation for autonomous underwater harbor surveillance," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots and Syst.*, Taipei, Taiwan, Oct. 2010, pp. 4396–4403.
- [8] D. Ribas, P. Ridaó, and J. Neira, *Underwater SLAM for structured environments using an imaging sonar*, ser. Springer Tracts in Advanced Robotics. Springer, 2010, vol. 65.
- [9] N. Hurtós, X. Cufi, and J. Salvi, "A novel blending technique for two-dimensional forward-looking sonar mosaicing," in *Proc. IEEE/MTS OCEANS Conf. Exhib.*, vol. 1, no. 7, San Diego, CA, Sep. 2013, pp. 23–27.
- [10] N. Hurtós, S. Nagappa, N. Palomeras, and J. Salvi, "Real-time mosaicing with two-dimensional forward-looking sonar," in *Proc. IEEE Int. Conf. Robot. and Automation*, Hong Kong, China, Jun. 2014, pp. 601–606.
- [11] S. Negahdaripour, "On 3D motion estimation from feature tracks in 2D FS sonar video," *IEEE Trans. on Robot.*, vol. 29, no. 4, pp. 1016–1030, 2013.
- [12] N. Gracias and J. Santos-Victor, "Underwater mosaicing and trajectory reconstruction using global alignment," in *Proc. IEEE/MTS OCEANS Conf. Exhib.*, vol. 4, Honolulu, HI, Nov. 2001, pp. 2557–2563 vol.4.
- [13] P. Ridaó, M. Carreras, D. Ribas, and R. Garcia, "Visual inspection of hydroelectric dams using an autonomous underwater vehicle," *J. Field Robot.*, vol. 27, no. 6, pp. 759–778, Nov. 2010.
- [14] A. Kim and R. M. Eustice, "Real-time visual SLAM for autonomous underwater hull inspection using visual saliency," *IEEE Trans. on Robot.*, vol. 29, no. 3, pp. 719–733, 2013.
- [15] P. Ozog and R. M. Eustice, "Real-time SLAM with piecewise-planar surface models and sparse 3D point clouds," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots and Syst.*, Tokyo, Japan, Nov. 2013, pp. 1042–1049.
- [16] N. Carlevaris-Bianco, M. Kaess, and R. M. Eustice, "Generic node removal for factor-graph SLAM," *IEEE Trans. on Robot.*, vol. 30, no. 6, pp. 1371–1385, 2014.
- [17] P. Ozog and R. M. Eustice, "Toward long-term, automated ship hull inspection with visual SLAM, explicit surface optimization, and generic graph-sparsification," in *Proc. IEEE Int. Conf. Robot. and Automation*, Hong Kong, China, June 2014, pp. 3832–3839.
- [18] M. Johnson-Roberson, M. Bryson, B. Douillard, O. Pizarro, and S. B. Williams, "Out-of-core efficient blending for underwater georeferenced textured 3D maps," in *IEEE Comput. for Geo. Res. and Appl.*, San Jose, CA, Jul. 2013, pp. 8–15.
- [19] R. Campos, R. Garcia, P. Alliez, and Y. M., "A surface reconstruction method for in-detail underwater 3D optical mapping," *Int. J. Robot. Res.*, vol. 34, pp. 64–89, 2014.
- [20] S. Barkby, S. Williams, O. Pizarro, and M. Jakuba, "Bathymetric SLAM with no map overlap using gaussian processes," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots and Syst.*, San Francisco, CA, Sep. 2011, pp. 1242–1248.
- [21] 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," in *Proc. IEEE/MTS OCEANS Conf. Exhib.*, Boston, MA, Sep. 2006, pp. 1–6.
- [22] 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," *Int. J. Robot. Res.*, vol. 31, no. 12, pp. 1445–1464, 2012.
- [23] Sound Metrics Corp. (2014) Didson 300 m Imaging Sonar. Specification sheet and documentations Available at www.soundmetrics.com.
- [24] B. Triggs, P. McLauchlan, R. Hartley, and A. Fitzgibbon, "Bundle adjustment – a modern synthesis," in *Vision Algorithms: Theory and Practice*, ser. Lecture Notes in Computer Science, 2000, vol. 1883, pp. 298–372.
- [25] D. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, 2004.
- [26] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Comm. of the ACM*, vol. 24, no. 6, pp. 381–395, 1981.
- [27] K. S. Arun, T. S. Huang, and S. D. Blostein, "Least-squares fitting of two 3-D point sets," *IEEE Trans. on Patt. Anal. and Mach. Intell.*, no. 5, pp. 698–700, 1987.
- [28] A. Segal, D. Haehnel, and S. Thrun, "Generalized-ICP," in *Proc. Robot.: Sci. & Syst. Conf.*, Jun. 2009.
- [29] M. Roy, S. Foufou, and F. Truchetet, "Mesh comparison using attribute deviation metric," *Int. J. of Image and Graphics*, vol. 4, no. 1, pp. 127–140, 2004.