Imaging Sonar-Aided Navigation for Autonomous Underwater Harbor Surveillance

Hordur Johannsson, Michael Kaess, Brendan Englot, Franz Hover, John Leonard

Abstract— In this paper we address the problem of drift-free navigation for underwater vehicles performing harbor surveillance and ship hull inspection. Maintaining accurate localization for the duration of a mission is important for a variety of tasks, such as planning the vehicle trajectory and ensuring coverage of the area to be inspected. Our approach only uses onboard sensors in a simultaneous localization and mapping setting and removes the need for any external infrastructure like acoustic beacons. We extract dense features from a forward-looking imaging sonar and apply pair-wise registration between sonar frames. The registrations are combined with onboard velocity, attitude and acceleration sensors to obtain an improved estimate of the vehicle trajectory. We show results from several experiments that demonstrate drift-free navigation in various underwater environments.

I. INTRODUCTION

Around the world there are underwater structures and areas that need to be routinely inspected for maintenance and security purposes. These include pier pilings, harbor seafloor, pipelines, oil platforms and ships. To address this need, Bluefin Robotics and MIT built a ship hull inspection vehicle (see Fig. 1), called the Hovering Autonomous Underwater Vehicle (HAUV) [1]. The HAUV is equipped with a Doppler Velocity Log (DVL) to measure velocity relative to a surface, a ring laser gyro for attitude measurements and a dual frequency identification sonar (DIDSON) [2] for imaging the structures being inspected.

Accurately keeping track of the vehicle position is crucially important, but difficult in harbor environments in particular. We need to ensure full coverage of the area being inspected, avoid obstacles and restricted areas, and also report an accurate location estimate for detected targets. It is difficult, however, to obtain a global position estimate underwater from an external source. GPS is only available at the surface, so acoustic beacons would need to be deployed. Moreover, a magnetic compass works poorly near large metal structures. Employing only rate gyros and odometry, over time sensor errors accumulate and the position estimate will drift. Using time of flight measurements with acoustic beacons has been commonly used in underwater navigation [3]–[5] to obtain a global position estimate, it has also proved successful in various applications like underwater archaeology [6] and ship hull inspection [7]. But in our
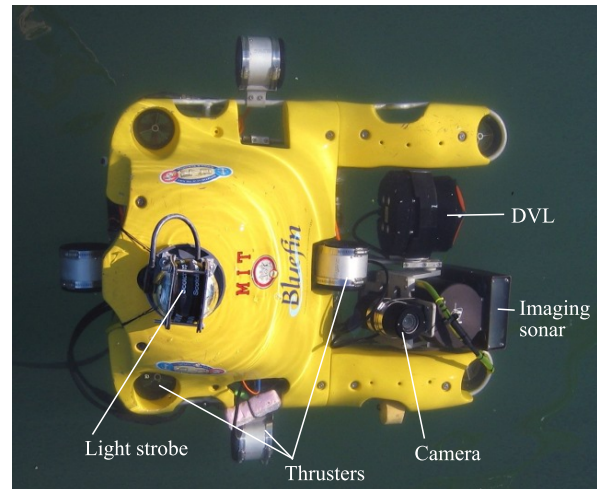
Fig. 1. Top view of the Bluefin-MIT Hovering Autonomous Underwater Vehicle (HAUV). The vehicle is equipped with a Doppler velocity log (DVL), an imaging sonar, an optical camera and a light strobe. The sonar and DVL can be actuated independently to optimally align the sensors to the surface being inspected.

work we are interested in providing drift-free navigation using the onboard imaging sonar, by identifying and aligning previously visited areas, combined with dead reckoning from the vehicle's other sensors.

Augmenting vehicle localization using sonars has been undertaken in a number of prior works. Walter et al. [8] used manually extracted landmarks, and later automatic feature detection [9] with the Exactly Sparse Extended Information Filter (ESEIF) to produce a map of the environment. An automatic feature detector and a landmark formulation using an EKF filter was used in [10]. Sekkati et al. used extracted corner features from DIDSON frames to estimate vehicle motion over several frames [11]. In related work Negah-daripour et al. combined the DIDSON with an optical camera for 3-D target reconstruction using opti-acoustic stereo [12]. Eustice et al. [13] used constraints from overlapping camera frames within a SLAM information filter to estimate vehicle pose. In work by Folkesson et al. [14], a forward-looking sonar was used with a prior map to track beacons in sonar images and use them to localize the vehicle. A full 360-degree sonar scanner has been used in partially structured underwater environments [15] for localization, by tracking line features in the environment using an EKF for the estimation process. Mallios et al. recently showed promising results in [16] using an mechanical scanning sonar and scan matching in an EKF framework.

Here, we use a pose graph formulation to combine onboard navigation information with sonar registration. Pose graphs [17], [18] represent a map of the environment as a graph, where the nodes are variables representing the vehicle states along its trajectory and edges are constraints between those variables. Efficient online methods have been developed to optimize this pose graph, such as incremental smoothing and mapping (iSAM) [19]; this is the approach used for the present work.

Our main contribution are i) an automated dense feature extraction technique that allows for effective registration and loop closures, ii) an efficient formulation and implementation of the map estimate, iii) a real-time implementation of the system, and iv) experiments in various underwater environments with the HAUV vehicle. We focus in this paper on imaged areas that are roughly flat and horizontal, like the large open areas of ship hulls, and the seafloor. Most importantly, our system allows for drift-free navigation without depending on any external infrastructure. The dense feature extraction allows us to detect and use a wide range of objects for this purpose, and is particularly useful to disambiguate between different places when searching for loop closures. We can deal with large areas having few features, because the sensor drift of our vehicle is fairly small. These scenarios are brought out in our analysis of the testing performed to date.

II. PROBLEM STATEMENT

The goal of our work is to correct drift in the vehicle state estimate over time using the imaging sonar. In this section we describe which quantities need to be estimated, followed by a discussion of the imaging sonar geometry.

A. Vehicle State

The vehicle state we are interested in estimating is the so-called vehicle pose, which consists of position and attitude. The vehicle position in 3D is specified by Cartesian coordinates x, y, z with respect to some arbitrary reference frame, such as the starting point of the mission or a GPS frame acquired before diving. The attitude of the vehicle is specified by the standard Euler angles ϕ, θ, ψ that reference roll, pitch, and heading respectively.

While the vehicle state has six degrees of freedom, we only need to estimate three. We obtain absolute depth measurements based on water pressure, as well as absolute measurements of pitch θ and roll ϕ based on gravity. The estimate for the remaining three degrees of freedom will drift over time when only based on gyro and DVL measurements.

The ring laser gyro is used for estimating heading, because using a magnetic compass in close vicinity to a ship hull or other steel structure is often not a viable option. The vehicle estimates its heading by integrating the angular rotation as measured by the gyro. The x, y position is then estimated by dead reckoning using the velocities from the DVL and the heading estimate. The goal of our work is to bound drift in the position estimate by using the imaging sonar.
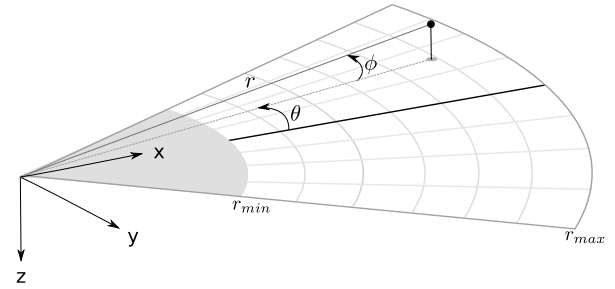

Fig. 2. Imaging sonar geometry.

B. Imaging Sonar Geometry

Using the imaging sonar to correct drift accumulated by dead reckoning requires understanding how the sensor functions. Following the formulation in [11], [12], [20], we define the geometry of the imaging sonar and derive a model that describes how the image is formed. To generate an image, the sonar emits a sound wave and then listens with an array of receivers to the returning sound wave, sampling the acoustic energy returned from different directions. The sonar samples the transducers at fixed intervals, providing time of flight, azimuth angle and intensity for each sample. Combining the returns from all the elements provides an image of the reflective surfaces in front of the sonar. For the imaging sonar we are considering, the vertical beam width is greater than the horizontal beam width. Note that for a given point in the image it can lie anywhere on an arc at a fixed range, spanning the vertical beam width.

Mathematically the imaging process can be described as follows. We define the coordinate system for the sonar as shown in Fig. 2. For simplicity let us assume that the sonar coordinate system coincides with the vehicle coordinate system. The x-axis is perpendicular to the sonar array, the y-axis is to the right and z-axis points down. Let us consider a point $\mathbf{p} = [x\ y\ z]^T$ in the sensor coordinate frame, where x, y, z are the Cartesian coordinates of the point. Now let $\mathbf{s} = [r\ \theta\ \phi]^T$ be the same point in spherical coordinates, where r is the range, θ is the azimuth and ϕ is the elevation of the point. We can relate the spherical and Cartesian coordinates with the following equations

$$\mathbf{p} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} r\cos\phi\cos\theta \\ r\cos\phi\sin\theta \\ r\sin\phi \end{bmatrix} \quad (1)$$

$$\mathbf{s} = \begin{bmatrix} r \\ \theta \\ \phi \end{bmatrix} = \begin{bmatrix} \sqrt{x^2+y^2+z^2} \\ \arctan 2(y,x) \\ \arctan 2\left(z, \sqrt{x^2+y^2}\right) \end{bmatrix} \quad (2)$$

In practice the image we measure of point $\mathbf{p}$ is $I(\mathbf{p}) = [r\ \theta]^T$. We define the Cartesian projection of the point $\mathbf{p}$ as

$$\hat{I}(\mathbf{p}) = \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} r\cos\theta \\ r\sin\theta \end{bmatrix} \quad (3)$$

This projection can be viewed as an approximation to an orthographic projection.

What is the effect of observing the same structure from two different poses? Let us first consider the effects of the approximation on translation. We start with point $p = [x \, y \, z]^T$, which yields the spherical coordinates $s = [r \, \theta \, \phi]^T$. Now the sonar moves by $[\delta x \, \delta y \, 0]^T$. The point in the new coordinate frame is

$$\mathbf{p}' = \begin{bmatrix} x - \delta x \\ y - \delta y \\ z \end{bmatrix} \qquad (4)$$

Using our projection we obtain

$$\hat{I}(\mathbf{p}) = \begin{bmatrix} r \cos \theta \\ r \sin \theta \end{bmatrix} \quad \hat{I}(\mathbf{p}') = \begin{bmatrix} r' \cos \theta' \\ r' \cos \theta' \end{bmatrix} \qquad (5)$$

From the projection we derive our estimated translation $\mathbf{t} = \hat{I}(\mathbf{p}') - \hat{I}(\mathbf{p})$. Then we can calculate the translation error $\epsilon = \mathbf{t} - [\delta x \, \delta y]^T$

$$\epsilon = \begin{bmatrix} \epsilon_x \\ \epsilon_y \end{bmatrix} = \begin{bmatrix} u(1 - \cos \phi) - u'(1 - \cos \phi') \\ v(1 - \cos \phi) - v'(1 - \cos \phi') \end{bmatrix} \qquad (6)$$

Now, if we assume a flat surface that is level with the world's x-y plane, at a distance h from the vehicle, then from equation (2) we can see that the elevation depends on the distance in the x-y plane and the vertical distance, z, to the point. Let us consider a situation where the targets are at 2 meter distance below us, then for 0.5 meter translation the error is around 0.14 meter when the target is close (around 2.25 meters), and decreases as the targets gets farther away (less then 3cm at 7m distance).

Next we analyze how the projection affects rotation around the vertical axes of the sonar. Let $\mathbf{p}$ and $\mathbf{p}'$ be the point before and after rotation respectively. We rotate $\mathbf{p}$ counter-clockwise by angle α around the z-axis

$$\mathbf{p}' = \begin{bmatrix} \cos \alpha & -\sin \alpha & 0 \\ \sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathbf{p} \qquad (7)$$

$$= \begin{bmatrix} r \cos(\theta + \alpha) \cos \phi \\ r \sin(\theta + \alpha) \cos \phi \\ r \sin \phi \end{bmatrix} \qquad (8)$$

and

$$\mathbf{s}' = \begin{bmatrix} r \\ \theta + \alpha \\ \phi \end{bmatrix} \qquad (9)$$

Assuming only rotation of the sonar, we let $\hat{\alpha}$ be the estimate of the rotation of the sonar. We obtain the estimate by calculating the angle between the images of the two points.

$$\|\mathbf{p}\| \, \|\mathbf{p}'\| \cos(\hat{\alpha}) = \hat{I}(\mathbf{p}) \cdot \hat{I}(\mathbf{p}') \qquad (10)$$

$$= \begin{bmatrix} r \cos \theta \\ r \sin \theta \end{bmatrix} \cdot \begin{bmatrix} r \cos(\theta + \alpha) \\ r \sin(\theta + \alpha) \end{bmatrix} \qquad (11)$$

$$= r^2 \cos \alpha \qquad (12)$$

$$\Rightarrow \hat{\alpha} = \alpha \qquad (13)$$

In conclusion, the projection preserves the change in azimuth angles. This holds for rotation around the sonar's z-axis, but normally the sonar is tilted relative to the vehicle's coordinate frame and it is the heading change in that frame that we would like to measure.

Rotating the sonar around its y-axis (pitch) only changes ϕ and has no effect on the projection of the points. The rotation only affects the intensity of the returned values, and eventually the points will go out of the sonar's field of view. Roll of the sonar will tend to cause compression of the points along the y-axis. In our case the error is minimal as the roll angle is always small. However, if the imaged surface deviates significantly from our horizontal surface assumption, then we would have to model this deviation. One crude approximation is to assume a plane that is aligned with the sonar and positioned at the sonar origin. A better approximation is to use range measurements from the DVL to estimate the plane. This is feasible when the DVL is used to track the surface being imaged, i.e. when doing ship hull oriented tracking or bottom tracking for the bottom surveys.

III. SONAR REGISTRATION

The key to bounded navigation drift is to periodically detect revisits to places the vehicle has been to before. In this section we will describe our method for detecting those revisits and deriving pose constraints by registration of the acoustic images from those places. Extraction of features needed for registration is described next.

A. Feature Extraction

Unlike a laser scanner, which gives the range to the closest obstacle, the imaging sonar returns multiple intensity values along the beam. An example sonar image is shown in Fig. 3(a). A strong return usually represents an object standing above the sea floor or the surface we are imaging, and it is then usually followed by a shadow. Also a hole or a depression on the surface can show up as a shadow but then there is no associated bright return. Variations in the returned signal are also caused by changes in material properties, the strength of the transmitted signal, receiver sensitivity, distance to target, and the grazing angle, among other things.

We propose to extract reasonably stable features based on sharp transitions. We identify sharp transitions in the measurement, and check for a possible return from an object followed by low signal caused by the object shadowing the background. The main steps of the algorithm are:

1) Smooth image
2) Calculate gradient
3) Threshold a top fraction as features
4) Cluster points and throw out small clusters

First the image is smoothed using a median filter, significantly reducing noise, while still preserving edges, as shown in Fig. 3(b). Next, the gradient is calculated by computing the difference between the current value and the mean of the last few values (Fig. 3(c)). The number of previous values used to calculate the mean around the current values affect the type of objects that are detected. We then mark points with gradient exceeding a given threshold as features (Fig. 3(d)). The threshold is adaptively selected, such that
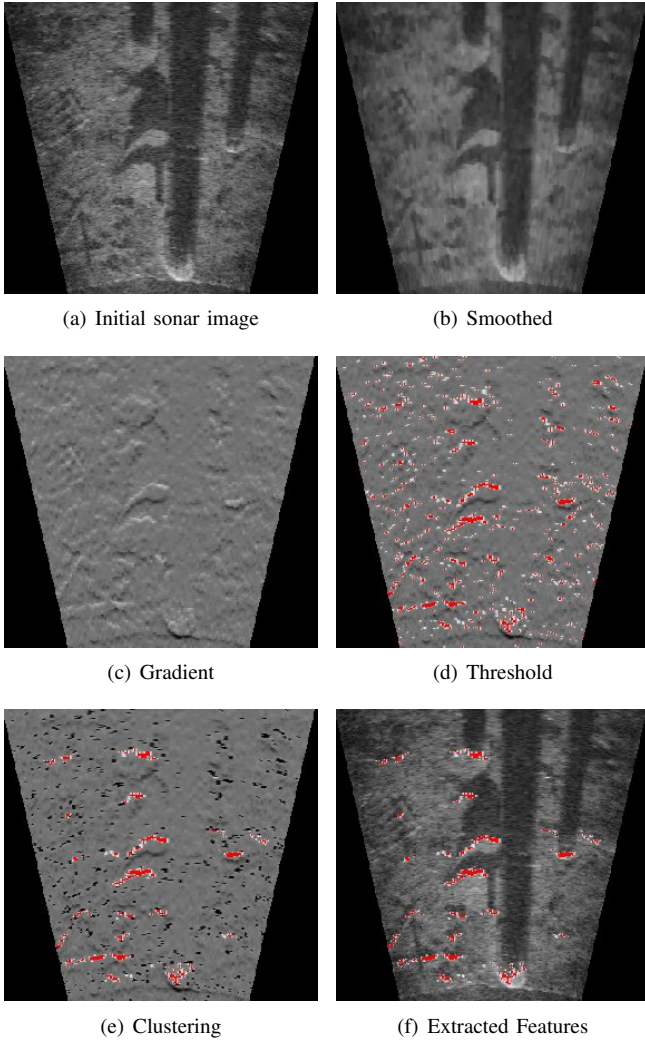
(a) Initial sonar image

(b) Smoothed

(c) Gradient

(d) Threshold

(e) Clustering

(f) Extracted Features

Fig. 3. Intermediate steps of the feature extraction process. The extracted features are shown in red.

(a) Frame A

(b) Frame B

(c) Registration of A and B

Fig. 4. Two scans before and after registration. Red points show the model scan, green features the current scan after registration.

B. Registration

We align two overlapping sonar images by registration of the extracted features as shown in Fig. 4 using the Normal Distribution Transform (NDT) algorithm [21]. The NDT algorithm works by assigning points to cells of a grid spanning the area. For each cell we calculate the mean and variance of the points that end up in the cell. This is done for four overlapping grids, where each grid is shifted by half a cell width along each axis. Using multiple shifted grids alleviates the effect of discontinuities resulting from the discretization of space. Two of the benefits using the NDT are that it gives a compact representation of the scan, and we do not need to get exact correspondences between points. This is useful in our case, because the movement of the vehicle causes variation in the insonification of the surfaces, which causes some points to drop in and out of the extracted feature set.

The NDT serves as our model for registration. With our current scan we calculate a score for each point by evaluating Gaussians parameterized using the values in the corresponding cells that contain the points. This gives a measure of how likely it is we would measure a given point from our data. We define a cost function as the sum of the negative score of all the points in the current view. Minimizing the cost function with respect to the parameters (x, y, ψ) gives the transformation between the two scans.

Because the main goal of the registration method is to localize the vehicle, we do not use any initial estimate of the vehicle location when searching for the match. Instead, we repeat optimization with several initial values to try to find the global minimum. We are also very conservative in accepting a match. The match has to have a normalized score over a given threshold, and the current scan has to include a minimum number of points.

IV. SONAR-AIDED NAVIGATION

We now use the geometric measurements resulting from sonar registration to correct the vehicle drift over time. For that purpose, we adopt a pose graph formulation of the simultaneous localization and mapping (SLAM) problem to obtain a least-squares estimate based on all measurements.

To achieve a drift-free estimate of the vehicle position we maintain an estimate of the vehicle's complete trajectory. As

we retain a fixed fraction of the features. Next we eliminate spurious features by clustering the points and eliminating small clusters (Fig. 3(e)). The extracted features are shown in Fig. 3(f), which typically consist of 1000 to 2000 points.

How do the extracted features translate into 3D Cartesian coordinates? Because the elevation of a measured point is unknown, there is an ambiguity in the projection into Cartesian coordinates. We assume that the points lie in a plane that is level with the vehicle. This approximation gives a reasonable result when we have a roughly level surface. We tend to stay around 1-2 meters away from the surface for the best imaging results, so the objects we are imaging are around 0.5-2 meters away from the projection plane. Registration is never performed on frames that are far away from each other. As a consequence, the error in the registration caused by our approximation can be up to 10-15cm, while typically being below 5cm. We could improve on this approximation by using the range measurements from the DVL to estimate the ground plane.
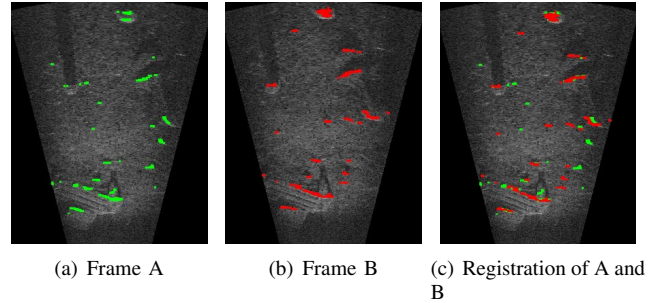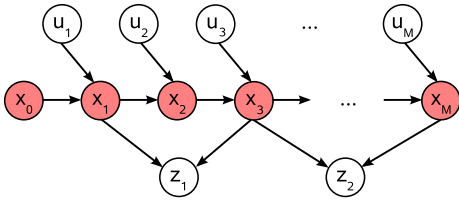
Fig. 5. Bayes net formulation of the estimation problem. x_i is the vehicle state at time i, u_i the control input and z_k a registration between arbitrary poses along the trajectory.

the vehicle moves around the environment it collects sensor information along the trajectory. This information is then used to detect when the vehicle reaches a place that it has seen before. This is commonly referred to as loop closure in the SLAM literature.

Following the formulation in [19], we model the problem as the joint probability distribution $P(X, Z, U)$ where $X = [x_1 x_2 \ldots x_N]$ is the vehicle trajectory, $Z = [z_1 z_2 \ldots z_M]$ are measurements between two poses and $U = [u_1 u_2 \ldots u_N]$ are the controls between two consecutive poses. The Bayes net corresponding to this model is shown in Fig. 5. From the model we can directly write the joint probability distribution as

$$P(X, Z, U) = P(x_0) \prod_{i=1}^{N} P(x_i|x_{i-1}, u_i) \prod_{k=1}^{M} P(z_k|x_{a_k}, x_{b_k}) \tag{14}$$

We assume Gaussian process and measurement models. For the motion model we define a function f that takes the previous state and the control input to predict our current state

$$x_i = f(x_{i-1}, u_i) + w_i \qquad w_i \sim N(0, \Sigma_i) \tag{15}$$

and for the measurement model we define a function h that relates two poses

$$z_k = h(x_{a_k}, x_{b_k}) + v_k \qquad v_k \sim N(0, \Lambda_k) \tag{16}$$

Given the measurements Z and U we obtain an estimate of our latent variables X. One such estimate is the *maximum a posteriori* (MAP) estimator. Let $\hat{X}$ be the MAP estimator, then

$$\hat{X} = \arg \max_X P(Z, U|X) P(X) \tag{17}$$

$$= \arg \max_X P(Z, U, X) \tag{18}$$

$$= \arg \min_X -\log P(Z, U, X) \tag{19}$$

Under the assumption that the measurement noise is Gaussian, we arrive at the following *nonlinear least-squares* problem

$$\hat{X} = \arg \min_X \sum_{i=1}^{N} \|f(x_{i-1}, u_i) - x_i\|_{\Sigma_i}^2 \tag{20}$$

$$+ \sum_{i=k}^{M} \|h(x_{a_k}, x_{b_k}) - z_k\|_{\Lambda_k}^2 \tag{21}$$

where $\|x\|_{\Sigma}^2 := x^T \Sigma x$.

By exploiting the sparseness of the problem it is possible to obtain an efficient solution to this problem with Gauss-Newton methods. Incremental smoothing and mapping (iSAM) [19] provides an efficient online solution to the problem that updates an existing solution rather than recalculating from scratch in each step.

V. EXPERIMENTS AND RESULTS

We performed several experiments in different settings. Our initial trials were tank experiments to verify that we could run the system online and stay localized over an extended period of time. Then we ran the vehicle in a more realistic setting: An inspection of the river bottom of the Charles River at the MIT Sailing Pavilion. Another experiment was performed on the *King Triton* vessel in Boston Harbor. The main difference between the harbor surveillance and ship hull inspection missions is the type of features that the vehicle encounters. The results allow us to show that the system works in a wide range of situations.

To evaluate the results we manually labeled several obvious objects in multiple frames over the duration of the mission. The purpose of this was for us to systematically track localization errors of those objects using different localization methods. It is important to note that these labels were never used in the navigation solution; all features used for navigation were automatically extracted during the missions. This metric tells us if a method gives a consistent estimate of an object position over time. We will not detect systematic errors like scale factor error in velocity measurements with this metric, but it is still useful for comparing different settings and algorithms. In Table I we summarize the main results from the experiments. We note that when the navigation accuracy is drifting the mean error might not be the best measure of accuracy, so we also include the maximum error encountered. The heading is also drifting, so the actual position error varies depending on the vehicle location. For example, given a rotation around the origin, the position error increases as the vehicle moves farther away from the origin.

The experiments demonstrate that our sonar registration can make use of a wide range of features for navigation. The features used for registration ranged from cooling pipes and zinc anodes on the ship, to pier pilings, planks and depressions in the Charles River. It is important not to focus on a specific type of feature, as the kind of features encountered in new underwater harbor environments are difficult to predict.

A. Tank

The first experiments that we report were performed in a testing tank at MIT. The tank is approximately 9 meters long, 3 meters wide, and 1 meter deep. We ran the vehicle for approximately one hour in the tank, keeping stationary while observing several cinder blocks as targets. Then we had the vehicle move in a box pattern so that it occasionally loses sight of the features but is able to reacquire them and

TABLE I

OVERVIEW OF THE RESULTS FROM OUR EXPERIMENTS.

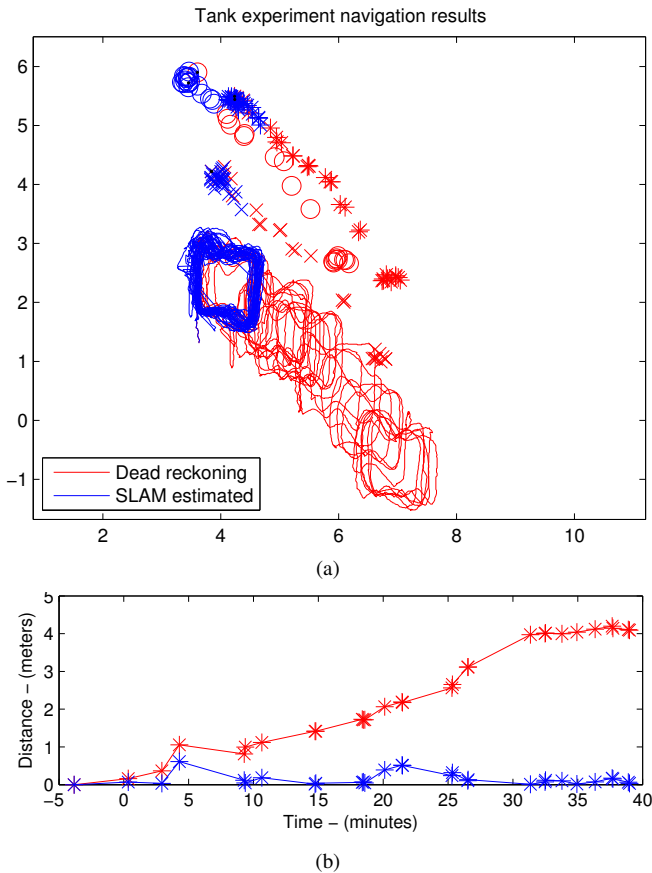| Data set	Length (m)	Duration (min)	Sonar			DVL		
			Err Mean (m)	Err StdDev	Err max	Err Mean	Err StdDev	Err max
Tank Box	173	43	0.2	0.17	0.82	2.4	1.48	4.3
Charles River	970	83	0.7	0.57	2.3	1.15	0.82	3.63
Charles River (post processed)	1459	133	0.31	0.23	1.2	1.6	1.18	4.88
King Triton	324	44	0.14	0.07	0.3	0.27	0.18	0.66

(a)

(b)

Fig. 6. Evaluation of the tank experiment, comparing the dead-reckoned position estimate (red) and the smoothed trajectory (blue). (a) Comparison of the DVL and smoothed trajectories. The vehicle executed a box pattern for just under an hour. In the tank the water is very shallow and the DVL does not perform well, so that we see a considerable drift in the dead-reckoned trajectory. (b) Estimation error of a single target that was manually selected in several sonar frames. The dead reckoning errors grows with time, while the error in the SLAM estimated trajectory is bounded.

re-localize. The vehicle had multiple features in sight for most of the frames.

One interesting aspect of this experiment is that the tank is a poor environment for the DVL, both because the water is shallow and the vehicle is surrounded by walls. The poor performance of the DVL meant that the drift using only dead-reckoning was large, but the navigation system was able to correct for that by periodically acquiring registration with the map. Note that a similar problem occurs when the DVL is used in deep water near the limit of its range. It would not have been possible to operate using only the DVL, because the drift was large enough that the vehicle would have hit the

walls of the tank. So to collect this data set it was actually necessary to run the navigation online.

Fig. 6(a) provides a comparison of the dead-reckoning and the smoothed trajectory. To evaluate the performance of the system we manually labeled several objects and compared their estimated position, at each time, to the position of the object when it was first seen. Fig. 6(b) shows a plot of the position errors for one of the objects.

B. Charles River

The purpose of this experiment was to emulate a harbor bottom inspection. One reason for performing this type of inspection is, when a foreign ship comes into harbor it is often required to inspect the area where it will dock. For this mission the vehicle traversed in a lawn mower pattern, with the sonar looking under the pier and collecting imagery of various objects lying on the riverbed.

The mission length was around 1 km, see Fig. 7 for results. One difficulty of the environment is the sparsity of the features. Good features were only visible from one side of the rectangle covered by the vehicle. Nonetheless, for around an hour and a half the vehicle was able to get registrations to earlier sonar images and improve the localization accuracy. During that time the heading estimate from the vehicle was around 4 degrees off and we observed object prediction errors up to 3.6 meter. Using the SLAM estimate the maximum error was 2.3 meters and the mean was 0.7 meters compared to 1.15 meters for the dead reckoning.

After one and a half hour the navigation system accepted a wrong registration, which caused the estimated position to veer off from the true position. Building on the experience from these experiments we implemented a few improvements to the navigation system, that we subsequently tested on the recorded data. The key components we improved upon are:

- Improved feature extraction
- More selective registration process
- Feature decimation

By improving the feature extraction we were able to use a stricter acceptance criterion to determine if a given registration was a good match. The DIDSON frames arrive at 5 to 10 Hz and it is not possible to try to register every single incoming frame to all the potential frames at a given location. In the initial version of the registration, we always picked the most recent DIDSON frame and then tried to register it to all frames within a given range. However, by exploiting the fact that the vehicle has very high navigation accuracy over short distances, we can make better use of our computational resources by not trying to register to frames that are close
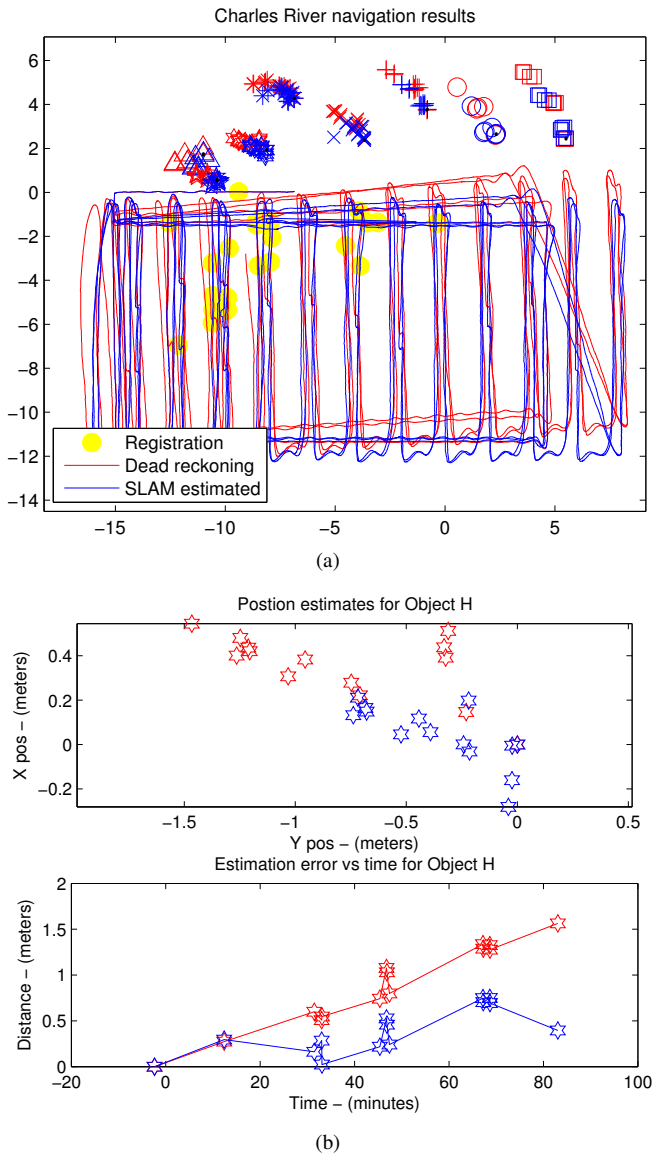
Fig. 7. Evaluation of the Charles River experiment. (a) Trajectories with manually labeled landmark positions shown for reference. Results based on dead-reckoning alone are shown in red, our corrected estimates are shown in blue. (b) Enlarged plot of position error for one of the objects. The effects of the gyro drift are clearly visible.
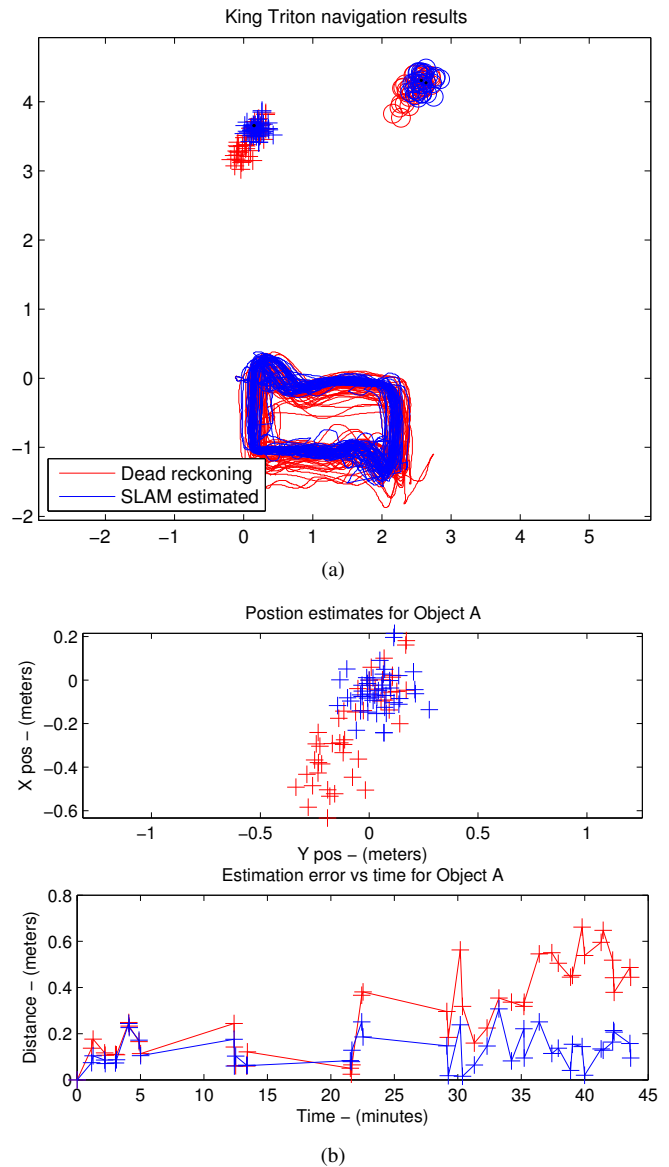

Fig. 8. Evaluation of the experiment on the *King Triton* vessel. (a) Trajectories with manually labeled landmark positions shown for reference. Results based on dead-reckoning alone are shown in red, our corrected estimates are shown in blue. (b) Enlarged plot of position error for one of the objects. The effects of the gyro drift are clearly visible.

in time, but instead choosing to match against frames that were recorded much earlier in the mission. This strategy improved our chances of finding large loop closures. Finally, we also incorporated feature decimation, which improved the execution speed of the registration.

After implementing these improvements we reprocessed the data we had collected. For the complete mission, the distance traveled was 1459 meters and the duration was 133 minutes. Now the mean error was down to 0.31 meters and the maximum error was 1.2 meters compared to a mean error of 1.6 meters and maximum error of 4.9 meters for the dead reckoning. The improvement was possible because more registrations were accepted while wrong registrations were avoided. The development of SLAM algorithms that

are robust to incorrect registrations is an important topic for future research.

C. King Triton

This experiment was targeted towards ship hull inspection, which is one of the main goals of the HAUV project. The acoustic imagery from a ship hull has different characteristics than the seafloor, in that it is often more repetitive and generally has a more complex shape. Therefore it is important to test the system on an actual ship hull. The *King Triton* is 20 meters long with a flat bottom, it has several features on the hull that system was able to use for localization, this included cooling pipes and zinc anodes. Note that the cooling pipes provide only limited constraints. Because of the small

size of the boat the DVL was pointed to the seafloor instead of toward the hull, as would be done for larger vessels. An overview of one of the runs can be seen in Fig. 8 together with the position error for one the verification objects. The HAUV navigated in a box pattern for about 45 minutes, and repeatedly localized using features on the hull to correct the drift in the dead reckoning, keeping the position error within 0.3 meters compared to 0.66 meters for the dead reckoning.

VI. CONCLUSION

We have described an online navigation system that can keep an underwater vehicle localized over an extended period of time. This was achieved by using imagery from the inspection sonar to concurrently build a map of the environment and localize the vehicle. We described a sonar processing method that extracts features from a sonar image that are then used for registration of sonar frames. Then, we showed results from several real time experiments where we demonstrated improved localization compared to dead reckoning using the DVL and gyro.

In future work we plan to improve the robustness of our current method to handle a wider range of environments, to be able to detect and recover from erronoues registrations, and to extend the registration method to handle more complex geometry. There is ongoing work to use vision methods for inspection and improved navigation using the same vehicle [22]. Negahdaripour et al. [12] worked on using an optical and an acoustic camera for opto-acoustic stereo. Further work in fusing those two sensors is of interest. Finally, it would be useful to run experiments using an independent measurement system to obtain ground truth data for the vehicle trajectory. This could be done by placing objects at known locations using differential GPS. Another possibility is to operate in an area where we can deploy an acoustic tracking system, which would give us ground truth along the full vehicle trajectory.

ACKNOWLEDGMENTS

REFERENCES

[1] 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 *Proceedings of OCEANS MTS/IEEE Conference and Exhibition*, vol. 1, pp. 761–766, 2007.

[2] E. Belcher, W. Hanot, and J. Burch, "Dual-frequency identification sonar (DIDSON)," in *Underwater Technology, 2002. Proceedings of the 2002 International Symposium on*, pp. 187–192, 2002.

[3] D. Yoerger and D. Mindell, "Precise navigation and control of an rov at 2200 meters depth," vol. 92, 1992.

[4] L. Whitcomb, D. Yoerger, H. Singh, and D. Mindell, "Towards precision robotic maneuvering, survey, and manipulation in unstructured undersea environments," in *Proc. of the Intl. Symp. of Robotics Research (ISRR)*, vol. 8, pp. 45–54, 1998.

[5] L. Whitcomb, D. R. Yoerger, and H. Singh, "Combined Doppler/LBL based navigation of underwater vehicles," in *Proceedings of the International Symposium on Unmanned Untethered Submersible Technology (UUST)*, May 1999.

[6] D. Mindell, "Precision Navigation and Remote Sensing for Underwater Archaeology," *Remote sensing in archaeology*, p. 499, 2007.

[7] S. Harris and E. Slate, "Lamp ray: ship hull assessment for value, safety and readiness," vol. 1, pp. 493 –500 vol.1, 1999.

[8] M. Walter, F. Hover, and J. Leonard, "SLAM for ship hull inspection using exactly sparse extended information filters," in *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, pp. 1463–1470, 2008.

[9] M. Walter, *Sparse Bayesian information filters for localization and mapping*. PhD thesis, Massachusetts Institute of Technology, 2008.

[10] B. Englot, H. Johannsson, and F. Hover, "Perception, stability analysis, and motion planning for autonmous ship hull inspection," in *Proceedings of the International Symposium on Unmanned Untethered Submersible Technology (UUST)*, 2009.

[11] H. Sekkati and S. Negahdaripour, "3-D motion estimation for positioning from 2-D acoustic video imagery," *Lecture Notes in Computer Science*, vol. 4478, p. 80, 2007.

[12] S. Negahdaripour, H. Sekkati, and H. Pirsiavash, "Opti-acoustic stereo imaging: On system calibration and 3-D target reconstruction.," *IEEE transactions on image processing: a publication of the IEEE Signal Processing Society*, 2009.

[13] R. Eustice, H. Singh, J. Leonard, M. Walter, and R. Ballard, "Visually navigating the RMS titanic with SLAM information filters," in *Robotics: Science and Systems (RSS)*, Jun 2005.

[14] J. Folkesson, J. Leederkerken, R. Williams, A. Patrikalakis, and J. Leonard, "A feature based navigation system for an autonomous underwater robot," in *Field and Service Robotics (FSR)*, vol. 42, pp. 105–114, 2008.

[15] D. Ribas, P. Ridao, J. Neira, and J. Tardós, "SLAM using an imaging sonar for partially structured underwater environments," in *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*, 2006.

[16] A. Mallios, P. Ridao, E. Hernandez, D. Ribas, F. Maurelli, and Y. Petillot, "Pose-based slam with probabilistic scan matching algorithm using a mechanical scanned imaging sonar," May 2009.

[17] E. Olson, J. Leonard, and S. Teller, "Fast iterative alignment of pose graphs with poor initial estimates," in *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, pp. 2262–2269, May 2006.

[18] F. Lu and E. Milios, "Globally consistent range scan alignment for environment mapping," *Autonomous Robots*, pp. 333–349, Apr 1997.

[19] M. Kaess, A. Ranganathan, and F. Dellaert, "iSAM: Incremental smoothing and mapping," *IEEE Trans. Robotics*, vol. 24, pp. 1365–1378, Dec 2008.

[20] S. Negahdaripour, P. Firoozfam, and P. Sabzmeydani, "On processing and registration of forward-scan acoustic video imagery," in *Computer and Robot Vision, 2005. Proceedings. The 2nd Canadian Conference on*, pp. 452–459, 2005.

[21] P. Biber and W. Strasser, "The normal distributions transform: a new approach to laser scan matching," in *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*, vol. 3, pp. 2743–2748, Oct 2003.

[22] A. Kim and R. Eustice, "Pose-graph visual SLAM with geometric model selection for autonomous underwater ship hull inspection," in *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*, 2009.