Relocating Underwater Features Autonomously Using Sonar-Based SLAM

Relocating Underwater Features Autonomously Using Sonar-Based SLAM

Maurice F. Fallon, John Folkesson, Hunter McClelland and John J. Leonard

Abstract

This paper describes a system for reacquiring features of interest in a shallow water ocean environment, using autonomous underwater vehicles (AUVs) equipped with low-cost sonar and navigation sensors. In performing mine countermeasures, it is critical to enable AUVs to navigate accurately to previously mapped objects of interest in the water column or on the seabed, for further assessment or remediation. An important aspect of the overall system design is to keep the size and cost of the reacquisition vehicle as low as possible, as it may potentially be destroyed in the reacquisition mission. This low-cost requirement prevents the use of sophisticated AUV navigation sensors, such as a Doppler velocity log (DVL) or an inertial navigation system (INS). Our system instead uses the Proviewer 900 kHz imaging sonar from Blueview Technologies, which produces forward look sonar (FLS) images at ranges up to 40 meters at approximately 4 Hz. In large volumes it is hoped that this sensor can be manufactured at low cost. Our approach uses a novel Simultaneous Localization and Mapping (SLAM) algorithm that detects and tracks features in the FLS images in order to renavigate to a previously mapped target. This feature-based navigation (FBN) system incorporates a number of recent advances in pose graph optimization algorithms for SLAM. The system has undergone extensive field testing over a

The authors are with the Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. Email: {mfallon,johnfolk,huntermc,jleonard}@mit.edu

Manuscript received Month 00, 2011; revised Month 00, 2011.

period of more than four years, demonstrating the potential for the use of this new approach for feature reacquisition. In this report, we review the methodologies and components of the FBN system, describe the system's technological features, review the performance of the system in a series of extensive in-water field tests, and highlight issues for future research.

I. INTRODUCTION

Numerous scientific, military, and commercial applications of autonomous underwater vehicles (AUVs) can benefit from the capability to renavigate to objects of interest that have been mapped during a previous mapping session. One such application is the neutralization of mines in shallow water — a task that has traditionally been carried out by human divers. The potential for casualties has motivated the development of unmanned systems to replace human divers. While a tethered robotic vehicle could be remotely controlled to perform Mine Countermeasures (MCM), a solution using untethered AUVs offers numerous mobility and independence advantages. Our goal for this work is to enable small, low-cost AUVs to robustly reacquire previously mapped features of interest, whether floating in the water column or lying on the seabed.

When mission requirements dictate that vehicle cost must be extremely low, the navigation problem for feature reacquisition is quite challenging. The primary challenge is to achieve good navigation performance using only very low cost sensors. To do this, we propose to use an *a priori* map, built by other vehicles with more capable (and expensive) sensing as input to our feature-based navigation software running on the low-cost reacquisition vehicle.

The ocean environment places severe challenges on the operation of autonomous vehicles. Communication with an untethered Autonomous Underwater Vehicle (AUV) via acoustic modems is typically restricted to low rates (bytes per second) and with several seconds of latency, such that interaction with the AUV is primary restricted to command and control of the vehicle [1]. Furthermore, due to the scattering and absorption effects of water, typical terrestrial sensors such as LIDARs and cameras are not suitable for use in many environments of interest. Finally, operational conditions in littoral waters are often challenging with high tides, wind-driven currents, and wave surge which each heavily affect the robot's position estimation as well as increasing operator risk during field operations. Access to the global positioning system (GPS) is not possible underwater and acoustic long-baseline transponders (and related technologies) can drift or be lost during the time lapse between deployment and use in a mission.

Nevertheless, unmanned underwater vehicles deployed with expensive inertial sensors, such as Doppler Velocity Loggers (DVLs) and Fiber Optic Gyroscopes (FOGs) can successfully negotiate complex underwater terrains [2], [3]. Sophisticated instruments of this type would enable an AUV to navigate to and relocalize a mine, however the neutralization would be prohibitively expensive. Less costly (and therefore expendable) AUVs, guided with only onboard inertial sensors, would typically encounter performance problems that would render them inefficacious in carrying out MCM.

Therefore, we have developed an approach based on the use of low-cost forward looking sonar technology to estimate vehicle motion as well as the location of features of interest so as to reach the *a priori* designated feature of interest (FOI). In this paper we describe the components of the system including Forward Look Sonar (FLS) feature detection and tracking, local map estimation, global map matching, and terminal homing. The structure of the remainder of this document is as follows: Section II provides an overview of the feature-based navigation (FBN) system for target reacquisition. Section III describes the Proviewer FLS and provides a detailed overview and performance analysis of the sonar processing tool-chain we have developed. Section IV discusses the SLAM engine at the heart of our AUV's navigation system before explaining how this map can be matched to an *a priori* map in Section V. Section VI presents some examples of the performance of our system in increasingly difficult operating scenarios. Finally, Section VII concludes this paper with a summary of our contributions and a discussion of issues warranting future research.

II. System Overview

Our mission scenario envisages revisiting a feature of interest (FOI) located on (or slightly above) a seabed and surrounded by significant clutter. To aid in this task the feature field is first uniformly surveyed by a high quality AUV (such as the REMUS 100 in Figure 1(c)) that navigates using acoustic beacons or an Inertial Navigation System (INS) with a Doppler Velocity Logger (DVL) to acquire accurate side-scan sonar data of the region of interest. Using the side-scan sonar data, a mosaic can be generated, as illustrated in Figure 2. Human operators then analyze this imagery to extract an *a priori* map of features as well as identifying the FOI. The accuracy of the feature locations in this map is typically 5m. These features may be man-made or natural objects on the sea floor as well as moored objects in the water column and represent the 'local universe' of the FOI. The features varied greatly in size and shape, and were typically less than one meter in diameter.

This map and the location of the FOI acts as input to a second *low-cost* relocalization vehicle. In the mission scenario we aim to release this vehicle at a distance of 100 to 1,500m from the center of

(a)

(b)

(c)

Fig. 1. *The AUVs used in the project. Top: iRobot Ranger — a low-cost single-man portable AUV. Middle: iRobot Transphibian equipped with fin-actuators to allow for 6 degree of freedom maneuverability. Bottom: Hydroid REMUS 100 used primarily as the survey vehicle.*

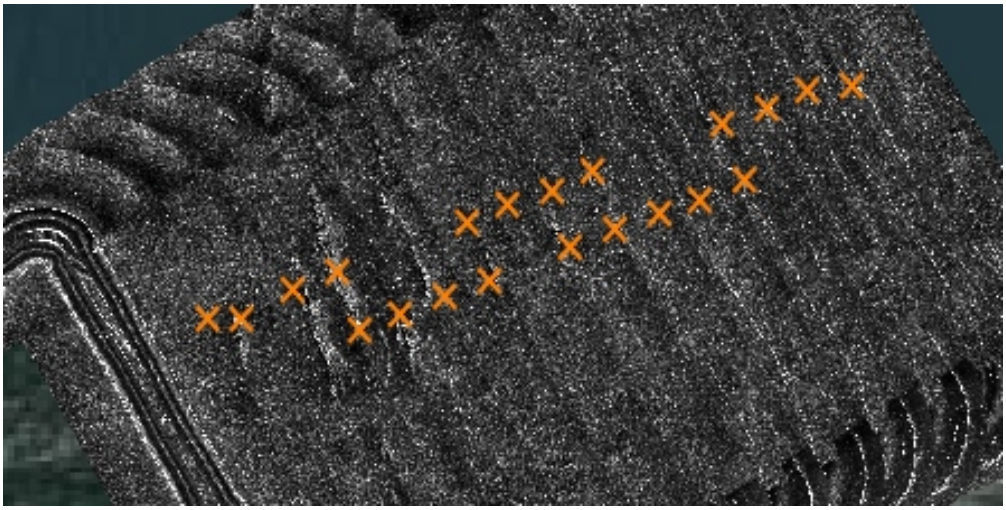

Fig. 2. *A typical prior map generated using a REMUS 100 equipped with a MarineSonics side-scan sonar. A series of features were extracted by trained human operators from the side-scan sonar imagery to produce this map for the target reacquisition mission. The distance between the features is approximately 20m. Figure courtesy of SeeByte Ltd.*

the prior map and have it swim at the surface close to the feature field before diving to the seabed. Upon re-entering this feature field, the vehicle will extract features from its sonar imagery and use these features to build a map, while concurrently estimating its position. This process is commonly known as Simultaneous Localization and Mapping (SLAM) [4].

Having re-observed a sufficient number of features, the AUV will localize relative to the *a priori* map and then home to the FOI. If successful the AUV will self-detonate or place an explosive charge. Because of this the vehicle is not intended to be recoverable. For these reasons a low-cost vehicle design requirement has had significant impact on the SLAM algorithms mentioned here. More detail about the re-attachment process can be found in Section V-C.

Test runs of these reacquisition operations have been conducted over several years on three different vehicles, shown in Figure 1: (1) an iRobot Ranger, (2) an iRobot Transphibian, and (3) a Hydroid REMUS 100.

A. Overview of Vehicles Used

The primary platform used during development has been the iRobot Ranger AUV [5]. This vehicle was equipped with a depth sensor, an altimeter, a GPS receiver, a 3D compass, an acoustic pinger, and

a Blueview Proviewer 900 kHz forward looking sonar. The vehicle's design was intended to be low-cost and lightweight. As indicated by Figure 1(a), it is single-man portable and deployable.

The design of the vehicle incorporates a servoed propeller which allows the vehicle to be highly maneuverable with a tight turning radius of 0.5m. This compares to 10m for the REMUS 100. This is of particular importance for the target homing at the end of the mission (see Section V-C) as a vehicle with a lower turning radius can more reliably seek to its intended target without needing to come about (and potentially loose track of the target).

Since dead-reckoning is the prime source of error growth, minimizing the distance to turn-and-capture increased the chance of a successful mission. The cruising speed of the AUV is quite low at about 0.6 m/s — comparable with typical ocean surface currents. Thus the dead-reckoning error due to the current can be quite significant. For that reason we developed a method to estimate the ocean current on-line. Our approach is discussed in the following section.

The vehicle specifically did not have a DVL, used for precise velocity estimation, due to cost limitations. Currently available DVLs cost approximately $20,000. It would be remiss of us not to mention that the current range of FLS devices are comparably expensive, however the manufacturer expects that mass production can reduce cost by an order of magnitude, as the current price is related to the initial cost of research and development. Nonetheless the utility of the capabilities outlined herein go far beyond this particular application.

Additionally, underwater vehicles typically use industrial processor designs which have more limited processing capability than typical desktop computers. This places severe computing restrictions on the operating vehicle — which we will discuss in following sections.

Other vehicles: While most of our testing focussed on the Ranger AUV, some experiments were also carried out with a biologically inspired AUV — the Transphibian [6] — which is equipped with fin-actuators to allow for 6 degree of freedom maneuverability and a variable buoyancy engine allowing for complete servoing around the FOI. Finally, while the Hydroid REMUS 100 was primarily used as a survey vehicle (as discussed in Section IV), it has also been used in the experiments to evaluate some sonar detector benchmarks in Section III-B. Doing this also demonstrates the generality of the algorithmic solution we have proposed.

B. Vehicle Dead Reckoning and Current Estimation

At the core of our SLAM system is a dead reckoning algorithm similar to others presented in the literature. High frequency depth estimates, altimeter altitudes, GPS fixes, and compass estimates of roll,

pitch, and heading are fused with actuation values (orientation of the servoed propeller and the estimated propeller RPM) using a typical EKF prediction filter to produce an estimate of the vehicle position and uncertainty. Please consider [7] for full overview of the state of the art. In benign current-free conditions, with careful tuning and excellent compass calibration this procedure produced a dead-reckoning estimate with about 1-2% error per distance traveled. However as we transitioned to more realistic test locations a more typical error was of the order of 3%. By comparison Fibre Optic Gyroscope-based (FOG) navigation systems can achieve drift rates below 0.1% in the open ocean. See [8] for a list of such systems.

In current-prone conditions in later stages of the project (as discussed in Section VI) a current estimation model was developed so as to reject the vehicle's drift in this situation. (Because of the nature of this project it was not possible to use the aforementioned DVL-enabled vehicle's estimate of the current profile.) Immediately prior to a mission, as the AUV approaches the feature field it was programmed to carry out a series of dives of increasing depth (between 3–5 in total). Between each dive it surfaced and used GPS fixes to compare its actual position to the position estimated without factoring in current modeling, and then used those comparisons to compute the ocean current estimate. No ADCP was present on the vehicle.

Estimating the current profile in this manner is problematic for a number of reasons — primarily because the parameters at greater depths are estimated using those at shallower depths. Additionally the following issues are not and cannot be modeled using our sensor suite:

1) Momentary surge currents
2) Currents rising and falling in the horizontal water column
3) Current profiles with numerous layers moving in different directions at unknown depths
4) Foreign objects altering the vehicle dynamics (such as seaweed entangling the propeller)

Aware of these limitations, we assumed a two current model: a surface current ($\vec{S}$, primarily due to wind) and a deeper laminar current ($\vec{D}$, due to the tide). These currents are estimated as 2D vectors in the horizontal plane.

At depths less than 0.5 m, the surface current is assumed to be dominant. Below this there is a transition between the two currents. Finally at the seabed the laminar current decays to zero — fitted as a square-root to the vehicle altitude. Numerically the combined current vector at a specific depth is

$$\vec{C}(a,d) = \sqrt{a/(a+d)}(w(d)\vec{S} + \vec{D}) \tag{1}$$

where a is the AUV altitude, d the AUV depth and $w(d)$ is the depth weighting of the surface current

and was calculated as follows

$$w(d) = \begin{cases} 1 & \text{for } d \leq d_0 \\ 1 - k(d - d_0)/|\vec{S}| & d_0 < d \text{ and } d \leq |\vec{S}|/k + d_0 \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

where $d_0 = 0.5m$, representing the boundary between the surface region and the rest of the water column, and $k = 0.2$ was empirically fixed to degrade this effect with increased depth.

This simplistic model performed reasonably well in smaller currents (below 0.3 m/s) and allowed the AUV to locate the feature-field after ingressing using solely dead-reckoning. After entering the field success was primarily due to the sonar-based SLAM algorithm (outlined in Section IV). At 0.3m/s we were able to enter the field approximately 85% of the time using this model and we estimate the error as about 5% per distance traveled.

However, when operating in high currents (close to or greater than the velocity of the vehicle), this model was not sophisticated enough to produce a sufficiently accurate position estimate for the AUV to find the feature field. While the vehicle could surface to reacquire GPS and then try again, the ocean-current-to-maximum-AUV-speed ratio of 50–70% represents a limitation to this or any other AUV perception system. This limitation is further discussed in Section VII.

As this system was developed in the absence of independent measurements of the currents at different altitudes, we refer the interested reader to a more fundamental treatment of ocean dynamics such as that presented in [9]. Given the limitations of our approach, a more accurately guided AUV is likely to have greater success in these challenging current regimes.

III. Proviewer Forward Look Sonar Processing

The sonar is our most important sensor: enabling the AUV to observe the features in its environment and to act based on these observations. During the project a series of Blueview Proviewer FLS sonars were used. In this section we give an overview of the sensor technology before presenting our sonar processing algorithms in Section III-A.

The Proviewer FLS operates using Blazed Array technology [10]. Typically the sonar consisted of two heads (horizontal and vertical) each with a field of view of 45°. Units with 90° and 135° field of view were later used and the statistics presented in Section III-B were calculated using the 90° version. Each head is made up of individual 22.5° sonar transducers as illustrated in Figure 3. Additionally, the horizontal transducers were tilted down 5° to best image the seabed as the vehicle typically flew horizontally and about 2m above the sea bottom.

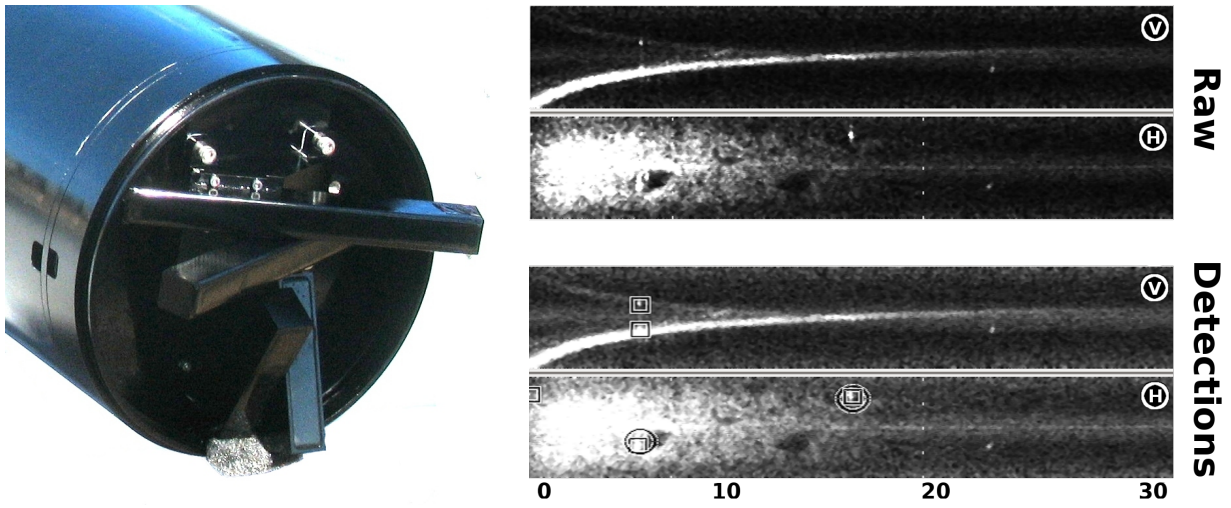

Fig. 3. *This 45° Blazed Array sonar consists of two perpendicular heads each made up of a pair of 22° sonar transducers (left). Each head produces a raw image — the upper pair (right) which we then automatically process to detect the features marked in the lower pair (right). In each case the vertical image is above the horizontal image. Both images are presented in polar coordinates with increasing ranges to the right, and right-to-left angle going vertically up. Note that the distinctive line in the vertical image is the observation of the seabed (which appears curved due to the use of polar coordinates here).*

The sensor works by transmitting an ensonifying signal (colloquially known as a 'ping') which reflects off of objects of incidence (in particular metal and rock) and is received by the transducer. The phase, amplitude, and delay of the returned signals are processed to produce the pattern indicated in Figure 4. This return is evaluated for each array element with one degree resolution in the plane of the head, and the output is then fused together by Blueview's onboard software to produce the image in Figure 5.

Sonar image formation is dramatically different from regular optical image formation [11]. Due to the reflected ensonification, the primary visual cue of the location of a feature is often not at its location, but rather beyond it — as indicated by a shadow of intensity values below the noise floor, as indicated in Figure 4. This, coupled with some increased amplitude and a gradient at the feature location, are the primary cues used in Section III-A to extract features. The length of the shadow can be indicative of the height of the feature, although moored features result in features detached from their shadows.

The outgoing sonar signal also has a significant lobe width, $\phi \sim 20°$, which means that there is significant ambiguity as to the location of the object in the axis normal to the sensed plane. In Section IV-B and Figure 9 we explain how we fused these observations so as to estimate the elevation of detections

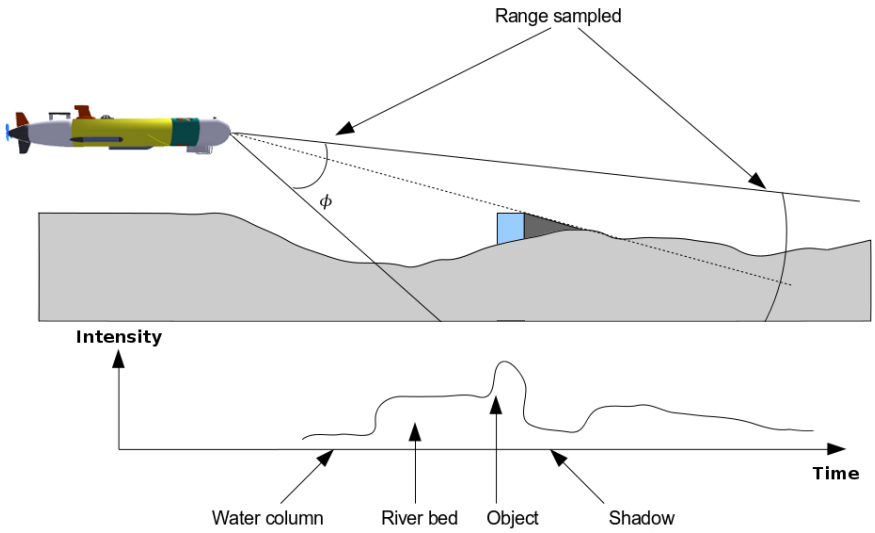

Fig. 4. *This figure illustrates the sonar image generation process for a single sonar beam. Each beam return is a vector of intensities of the returned sonar signal with objects of high density resulting in high returns and shadows formed by lower intensities.*

using only the horizontal image.

A. Sonar Feature Detection

In this section we outline our algorithms which extract point features from regions of high contrast. Williams *et al.* [12] implemented a similar processing pipeline to ours: beginning with raw data and progressively processing it to extract point features which are then tracked to carry out SLAM. Related previous work in our laboratory [13] also implemented a similar tool-chain, but was focused on dense registration rather than explicit point feature detection and extraction.

In related work, Ruiz *et al.* [14] performed on-line sidescan-based SLAM, but as they suggest this approach is not suitable for single pass operations such as the proposed mission. Fairfield *et al.* [15] utilized pencil-beam sonars to map flooded tunnels and sink holes. Using a particle filter, this approach can avoid explicit data association but the environment considered is quite different than the open ocean. Meanwhile, forward looking sonar has also been used for obstacle detection and path planning [16]. In this application the feature extraction is focused on conservative estimation of all detected objects given the noisy output of the FLS systems. Finally, [17] describes multi-target tracking of multiple features from a FLS using a probability hypothesis density (PHD) filter.

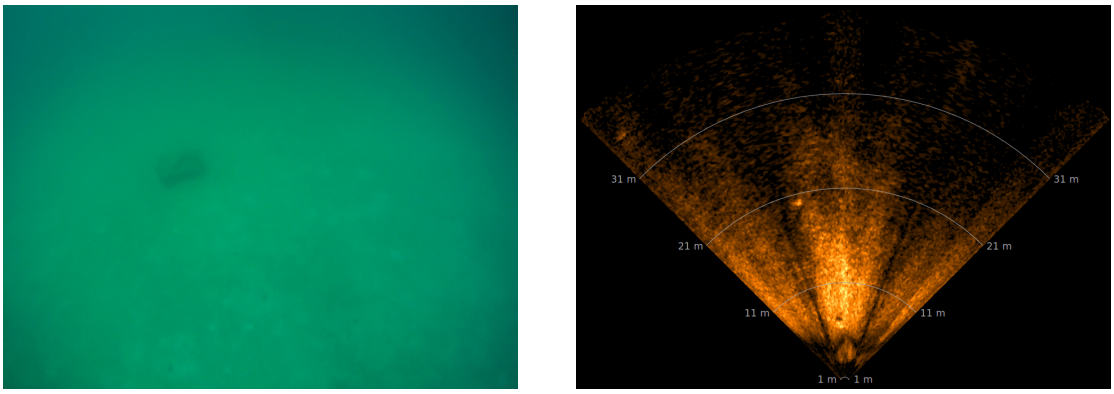

Fig. 5. *Typical underwater camera and sonar images (approximately synchronized). The clear water and well lit scenario represents some of the best possible optical conditions, nonetheless visibility is only a few meters. The 90 ° Blazed Array sonar horizontal image indicates 3 features (one at 5m in front; one at 20m and 5° to the left and one at 35m and 40° to the left) — which is more than typical.*

The processing tool-chain is carried out on the polar coordinate image produced by the Blueview SDK (Figure 3) rather than the Cartesian image (Figure 5). The maximum range of the Blueview FLS is 40m and images are typically processed at 3–4Hz. Objects were typically visible up to 20m away while brightly reflective objects are detectable at 40m.

Given the performance requirements of the platform, the algorithm is required to be adaptive to bottom reflective brightness. This is achieved by the online estimation of an average background image immediately after the vehicle levels out at its cruising depth. Estimating this background noise image, which is highly dependent on pitch and altitude as well the seabed material, is essential to achieve excellent performance in both sandy and muddy bottomed environments.

The steps used to process the images are as follows:

1) Form a 50 x 157 image of the vertical head sonar data (as in Figure 3).
2) Find sea bottom, altitude, and relative slope using line extraction.
3) If slope (i.e. vehicle pitch) is too far from level flight: stop processing.
4) Form a 50 x 157 image of the horizontal head sonar data.
5) Using the altitude (from step 2) select an altitude-specific averaged background image.
6) Low-pass filter the current horizontal image into the background image.
7) Subtract the background image from the horizontal image to form the shadow image.
8) Segment vertical image into bottom and open water regions.

9) Segment horizontal image into three regions: a) near field open water, b) ranges where the bottom is sufficiently bright to see shadows, and c) the remaining range out to 40 m.
10) Low-pass filter along each bearing to eliminate very small features and noise.
11) Search for edges as gradients in each segment with thresholds based on the average background image level. (This gives some adaption to different bottom types)
12) Search for a feature indicated by a bright area close to the sensor followed by a shadow (below the noise floor).
13) Select for output any features found in both heads as well as the strongest remaining detections in each segment.

This relatively simple detection algorithm is implemented using integer arithmetic and optimized for speed. This toolchain is similar to that presented previously in [13], including the figures presented therein.

In terms of computing load, the feature detector uses little processing power. The formation of the input image (using the Blueview SDK) — the input to this process — requires a substantial 180ms per image. The feature detector requires about 18ms while the remaining CPU power is used to fuse the measurements, to make high level mission decisions, and to control the AUV.

Ongoing work by the manufacturer to reduce this time (by forming the sonar image on the device) would allow a great increase in the processing frequency which in turn will improve the tracking and mapping performance.

B. Sonar Processing Performance Evaluation

To benchmark the performance of the feature detector we present a series of experiments in which a 90° FLS was attached to the FOG-enabled REMUS 100 in Figure 1(c). Patterns similar to the missions outlined in Section VI were carried out. Because of the shortness of the missions (10–15 min) and quality of navigation, the AUV accumulated negligible positioning error when compared to the Ranger platform and in what follows we assume its position estimate to be the 'Ground Truth' position.

The vehicle navigated a field of features arranged as in Figure 2 whose prior locations were accurate to about 5m. Four different missions on the same feature field were carried out from different directions and along different approaches. A total of 17 features were in the field: a mixture of natural and man-made, both moored and lying on the seabed. Time spent within the feature field was typically 180 seconds during which about 300 sonar pings were detected. The vehicle traveled at approximately 1.3m/s. Typical feature detection performance is shown in Figure 6.

1) Precision: The first metric we will consider, precision, is the ratio of correctly detected features to the total number of extracted features. We define a correctly detected feature as one which lies within a certain distance, r_{thres}, of a point feature, p_i, from the *a priori* map

$$C_i = \begin{cases} 1 & \text{if } ||d_i - p_j|| \leq r_{\text{thres}} \ \forall \ j \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

where $||d_i - p_j||$ represents the Euclidean distance between the two entities — evaluated in 2D for simplicity. The detection ratio is then simply

$$D = \frac{\sum_{i=1}^{F} C_i}{F} \tag{4}$$

where F is the total number of features detected. D lies in the range $0 \leq D \leq 1$.

Results are presented in Figure 7. In summary approximately 70% of detections fall within 3m of a prior map feature location. The remaining 30% — falsely detected noise were uniformly distributed in the rest of the sonar sensor footprint. Mission 2 exhibited significantly more turning motion, we believe that timing latency is to blame for the performance reduction in this case.

2) Recall: A second metric is that of recall — how frequently is a particular feature f detected when the sonar senses the area in which it is located. To measure this parameter we count the number of times that a feature is detected within 5m of a prior map feature D_f and express it as a fraction of the total number of times that the sonar sensed the location of the prior map feature N_f. Which is simply:

$$R_f = \frac{D_f}{N_f} \tag{5}$$

If $R_f = 1$ then each time it was possible to sense feature f it was detected. Meanwhile $R_f = 0$ means the feature could not be detected at all (probably because it was indistinct, obscured or perhaps a mistake in the prior mapping process). This metric is presented in Figure 7 for the same 4 missions.

For each mission, a dot represents the recall rate of a particular prior map feature, while the bar indicates the median recall rate across all features. In the case of Mission 1, on average we detected each feature in 2 of every 3 scans that sensed the area in which it was located.

3) Field of View: Using the detections from the above experiments we also considered the effect of artificially varying the Field of View (FOV) of the sonar. We evaluated which feature extractions would have occurred for FOVs smaller than the 90° actually used.

In the case of a 45° sonar, the results show that on average 1.5 features would have been visible at any one time during these missions. This figure rose to 2.5 features for the 90° sonar. This gives an indication of the sparsity of information available to the navigation algorithm. In particular, for perfect

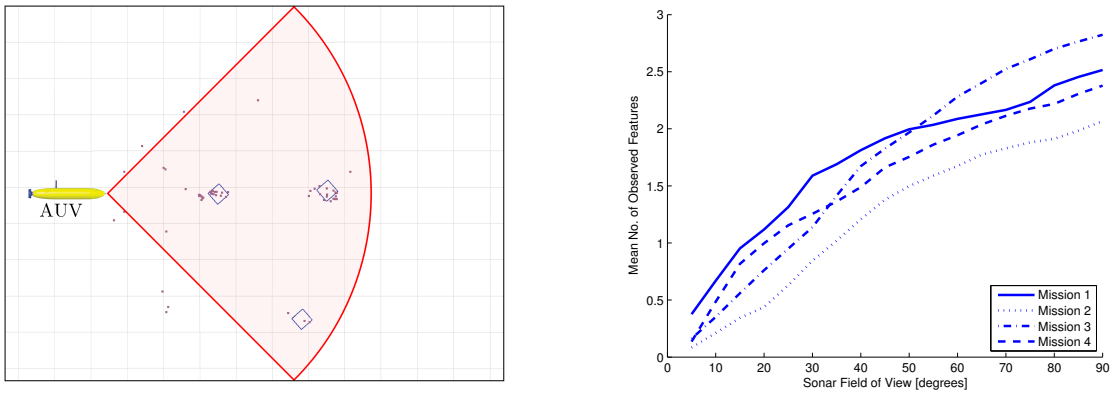

Fig. 6. *Left: Typical feature detection performance. The AUV is at the left hand side and its 90° sonar footprint is indicated by the red wedge. Each purple dot indicates a sonar point feature extraction. Two features at the center of the footprint have been detected several times. The third feature located at the side of the sonar footprint would not have been recognized using a 45° FOV sonar. The scale of the grid is 5m and the maximum range 40 meter. Right: Analysis of the average number of features detected at any instance during our missions — while varying the Sonar Field of View (FOV).*

data association it was typically not possible to uniquely determine the AUV's position using only a 45° sensor.

Due to the mechanical design of the sonar transducers and the small diameter of the Ranger AUV, all experiments in Section VI were run with a 45° sonar. Given this low aperture size, the sampling frequency, and the vehicle velocity, the AUV typically had to pass within 5m of a bottom feature so as to observe it frequently enough to reliably detect it.

While these results illustrate encouraging performance of the sonar detector; inaccurate dead-reckoning and a sparse feature field (often with one or no features visible at any particular time) mean that using feature detections alone to infer the vehicle's location in not possible. In the following section we discuss how the output of the detector is fused with the AUV navigation to jointly estimate its location and the relative position of the features.

IV. Marine Mapping and Localization

Simultaneous Localization and Mapping is a core skill of any intelligent autonomous robot: jointly estimating its position and the world around it as it actively explores. The SLAM field has developed a rich literature including EKF-based approaches such as [18], [19] and particle filter approaches including

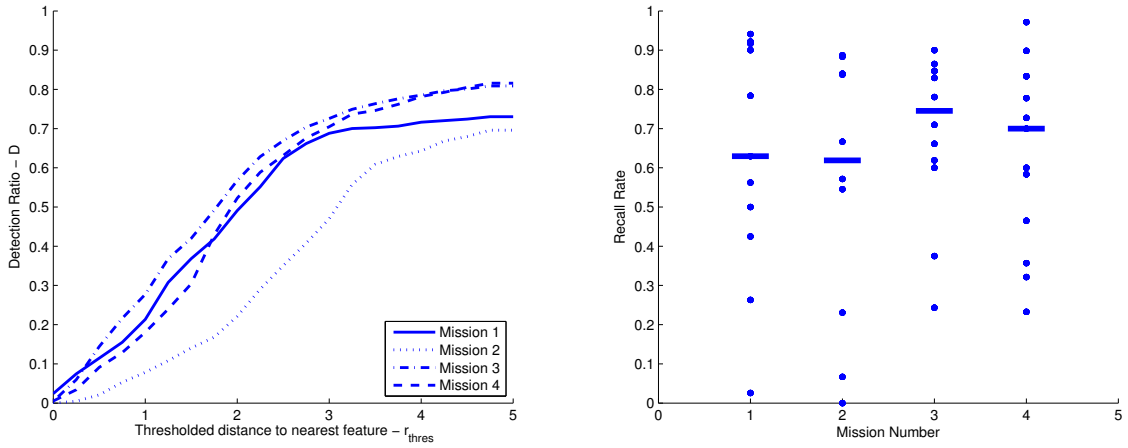

Fig. 7. *Sonar detection statistics for 4 typical missions. Left: the percentage of detections lying within a certain distance of the actual feature locations (in meters). Right: the frequency with which a feature is detected — given that the sonar sensor actually sensed that location. Each dot refers to a specific feature while the bar represents the median. See Section III-B for more detail.*

FastSLAM and GMapping [20], [21]. The former approaches suffer from a loss of information due to the linearization decision inherent in the EKF correction step [22]. Although the particle filter approaches solve this problem through Monte Carlo sampling, the loss of diversity during a loop closure means that for any large–sized exploration a prohibitively large number of particles is required.

The state of the art has emerged to be non-linear optimization of the graph of the vehicle trajectory and feature observations. In this approach the linearization points are regularly re-evaluated allowing us to maintain a consistent solution which can be efficiently solved on-line. Our implementation uses an approach similar to the Square Root Smoother [23], [24].

The joint probability distribution of the vehicle trajectory $X = [x_1\ x_2\ \ldots\ x_N]$, the sonar feature detections $Z = [z_1\ z_2\ \ldots\ z_M]$ and the dead-reckoning measurements between successive poses $U = [u_1\ u_2\ \ldots\ u_N]$ is given by

$$P(X, U, Z) = P(x_0) \prod_{i=1}^{N} P(x_i|x_{i-1}, u_i) \prod_{m=1}^{M} P(z_m|x_m, s_m) \tag{6}$$

where x_m represents the vehicle pose while observing the sonar feature s_m at a relative position z_m. $P(x_0)$ is a prior on the initial pose.

The *maximum a posteriori* (MAP) estimate of the vehicle trajectory can then be formed given the

measurements. Denoting this estimate $\hat{X}$, the resultant MAP estimator is given by

$$\begin{aligned}\hat{X} &= \arg\max_X P(U, Z|X)P(X) \\ &= \arg\max_X P(X, U, Z) \\ &= \arg\min_X -\log P(X, U, Z) \end{aligned} \tag{7}$$

Our solution, originally described in [25], incrementally solves this nonlinear least-square (NLS) optimization with some considerations for the marine environment and our limited processing capability. We start by initializing the graph with the GPS position of the AUV at dive and then add a new node to the graph representing each incremental change in AUV pose (which is synchronized to a sonar ping).

The dead-reckoning and sonar measurements are added to the measurement matrix A_i (at time i) which corresponds to the square root of the information matrix H_i. The information matrix is in turn the inverse of the covariance matrix for the measurement Σ_i

$$A_i^T A_i = H_i = \Sigma_i^{-1} \tag{8}$$

Solving Equation 7 is equivalent to minimizing the cost of the NLS optimization

$$E = \sum_i |(A_i \triangle \mathbf{x} - \mathbf{b}_i)^T (A_i \triangle \mathbf{x} - \mathbf{b}_i)| \tag{9}$$

With each additional measurement i, this cost grows as

$$\triangle E_i = (A_i \triangle \mathbf{x} - \mathbf{b}_i)^T (A_i \triangle \mathbf{x} - \mathbf{b}_i) \tag{10}$$

where $\triangle \mathbf{x}$ is the change in state from the current estimated state (where A_i and $\mathbf{b}_i$ are evaluated). $\mathbf{b}_i$ is the measurement mean and the A_i matrix characterizes the measurement covariance. This $\triangle E_i$ is added to all the previous terms and the resulting sum minimized by adjusting the state

$$\sum_j (A_j^T \sum_i (A_i \triangle \mathbf{x} - \mathbf{b}_i)) = A^T (A \triangle \mathbf{x} - \mathbf{b}) = 0 \tag{11}$$

The matrix A is incrementally formed by stacking the A_i while $\mathbf{b}$ is the state vector formed by stacking $\mathbf{b}_i$. As a result A has rows for each measurement component and columns for each state component.

The measurement matrix can then be decomposed as $A = QR$ with R upper–triangular Cholesky triangle and Q an orthonormal matrix. The minimum cost state is then found by back substitution

$$R^T (R \triangle \mathbf{x} - Q^T \mathbf{b}) = (R \triangle \mathbf{x} - \mathbf{b}') = 0 \tag{12}$$

where $\mathbf{b}' = Q^T \mathbf{b}$. Upon adding a new A_i we can append it to the R matrix and the subsequent decomposition step involves only the last few states.

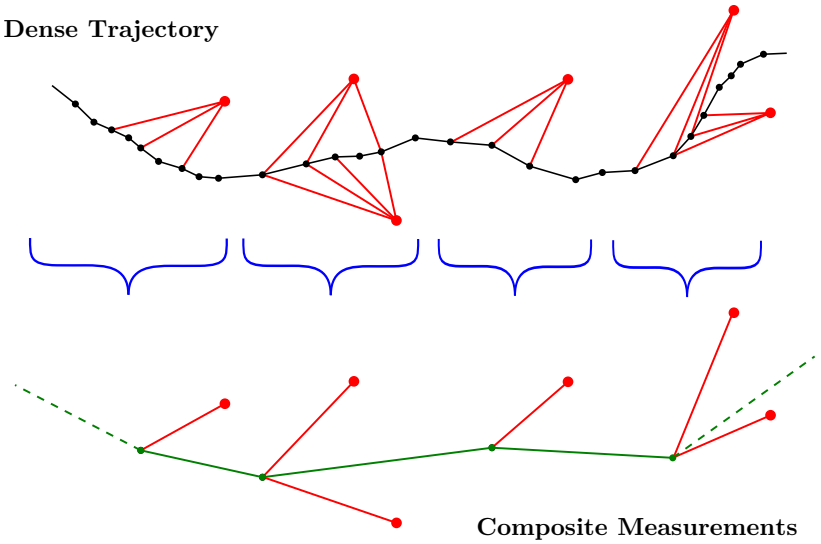

Fig. 8. *The estimation problem can be viewed as a graph. As the robot explores, a pose (black) is added to the graph at each iteration while feature detections (red) are also added to produce a dense trajectory. This dense trajectory is very large so we periodically marginalize portions of the trajectory and the feature observations into composite measurements (green) at a much lower rate. The composite measurements are the input to the matching algorithm in Section V.*

The complexity of this system of equations is tied to the sparseness of the A matrix which is itself dependent on the fill–in caused by loops in the graph structure. We explicitly avoid carrying out loop closures in this filter so as to maintain sparsity. All of this ensures that the matrices remain sparse and computation complexity predictable. Decomposition will not grow in complexity at each iteration while the computational cost of back substitution will grow, but it is linear.

A. Composite Measurements

So as to avoid computational growth due to an ever increasing graph size and to produce an input to the map matching stage, we periodically rationalize the oldest measurements from this graph to form a *composite measurement*. To do this we marginalize out all the poses that have occurred during the previous (approximately) 10 second period to produce a single node for the relative motion for that period as well as nodes for fully detected features and the associated covariances. This approach is similar in concept to key-frames in Vision SLAM and is illustrated in Figure 8.

The specific choice of 10 seconds is related to the speed of the AUV and the separation of the features.

Also doing so at this rate (corresponding to 1 node every approximately 100 velocity measurements) keeps the optimization problem relatively small.

We time this marginalization step to occur after a feature has left the sonar field of view as this allows us to optimally estimate its relative location given all available information. This composite measurement is then added to a lower-frequency higher level graph. This low-frequency graph is used as input to the prior map-matching algorithm in Section V. Meanwhile the high frequency graph begins to grow again by the insertion of newer constraints into A_i.

An alternative approach would be to maintain the dense trajectory of the robot pose at all times. This is the approach taken by iSAM [26], however given the size of the resultant graph, we are not certain that such an approach would have been able to yield a computationally constant solution required for our low-powered embedded CPU onboard the AUV.

Additionally and unlike most land-based systems, the underwater domain is characterized by extended periods where the seabed is featureless for long distances. In that case, the resultant composite measurement is simply the relative trajectory of the distance traveled.

B. Feature Tracking

While the section above explains how the graph of the trajectory and sonar observations is optimized and efficiently solved, we have not discussed the way in which sonar features are proposed.

The sonar detector passes point extractions to a feature nursery which maintains a vector of all recent detections. The nursery feature projects the detections into a local coordinate frame using the recent vehicle dead-reckoning and uses a probabilistic distance threshold to associate them with one another. Should a sufficiently large number of detections be clustered together (approximately 7–8, dependent on the spread and intensity of detections) it is inferred that a consistent physical feature is present.

At this stage the nursery feature is added to the Square Root smoother. All of the relative AUV-to-point constraints for that feature are then optimized which results in improved estimation of the feature and the AUV trajectory. Subsequent point detections, inserted directly into the pose graph, result in an improved estimate via further Square Root smoothing.

This approach also estimates the height/altitude of the sonar feature using the 2D sonar intensities measured at each iteration. This amplitude distribution is non-Gaussian and we model it instead as a cubic-normal distribution

$$A(\phi) = \mathcal{N}((\phi/\phi_{\text{nom}})^3; 0, 1) A_{\text{meas}} \tag{13}$$

where A_{meas} is the actual measured intensity. This distribution is characterized as being relatively flat within the nominal lobe width ϕ_{nom} before dropping off quickly a few degrees beyond the lobe edge. This allows us to differentiate bottom from moored features. See Figure 9 for a pictorial showing the height estimate of a feature position converging.

Finally it should be noted that the input to this feature tracker are point features characterized only by their location and covariance — due to the low resolution of the sonar sensor. This makes it difficult to robustly associate non-consecutive observations and hence to infer SLAM loop closures on the graph structure.

Additionally, clusters of nearby targets were prone to being mistakenly clustered into a single combined target. More careful estimation of target clusters is an interesting avenue of future work as well as being a possible avenue for successful loop-closures.

V. Global Estimation and Map Matching

Given the high level graph of the robot trajectory and observed feature locations, it still remains for the autonomous system to make a critical judgment of where it is relative to the *a priori* map and to decide if this relative match is certain enough to be declared convincingly. To do this we maintain a set of match hypotheses in parallel. We compare them probabilistically so as to quantify the quality of the map match.

This comparison is implemented using a bank of estimators working in parallel — each tracking a different match hypothesis. The relative likelihood of one match hypothesis over another is computed using positive information (of prior features detected by the sonar) as well as negative information (of prior features that were expected but undetected by the sonar) and in this way matching can be done in a probabilistically rigorous manner. Negative information can be summarized as follows: *if one expects to detect features along a proposed robot trajectory and these features were not seen then the proposed trajectory is less likely.*

The inclusion of this extra information is motivated by the regular structure of the feature field and the inability of positive information metrics to estimate the relative position of the AUV along these lines. The incorporation of negative information in this way is, to our knowledge, a novel contribution and was motivated by information not captured by algorithms such as Joint Compatibility Branch and Bound (JCBB) [**?**].

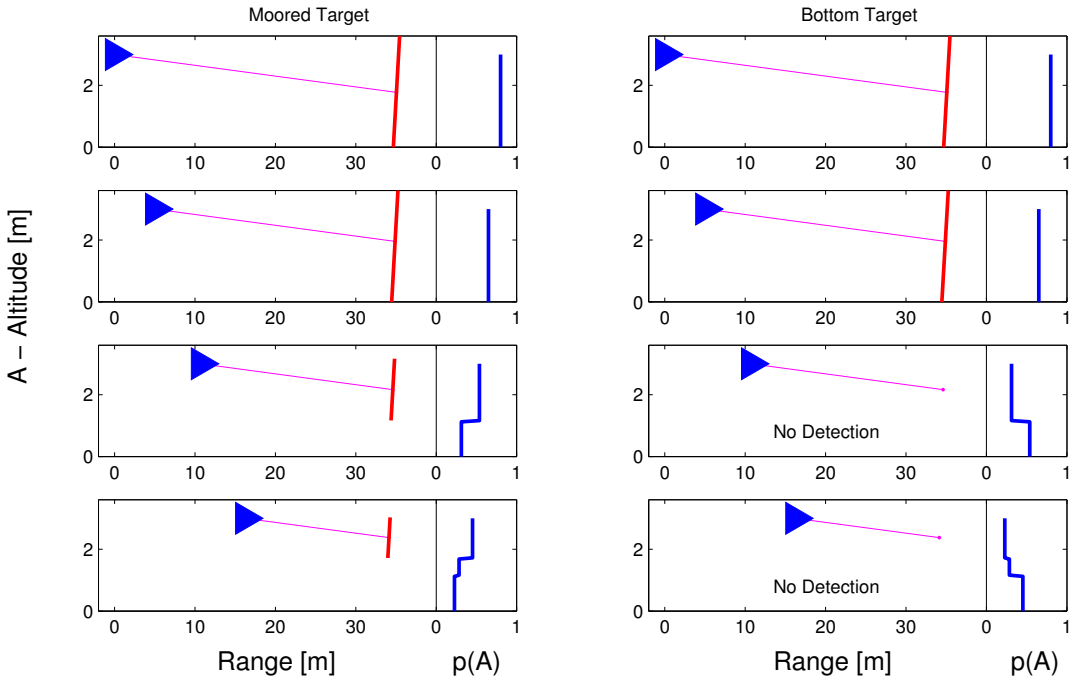

Fig. 9. *A side-view pictorial illustrating convergence of feature height estimates for a moored and a bottom feature — left and right columns respectively. On the left of each pictorial an AUV (blue triangle), at 2.5m altitude, approaches a feature (located at 35 m) and senses it (as indicated by the magenta line) with a horizontal sonar facing 5° down. The intensity of the feature detection infers a set of plausible off-beam heights (the red line indicates beam-width). Multiple observations allow the altitude PDF, $p(A)$, to converge to within a couple of meters. In the case of the bottom feature, no detections for the final two frames infer that the feature is below the main beam-width. In real experiments a feature could be detected as many as 30 times.*

A. Negative and Positive Scoring

In SLAM, multi-hypothesis comparison can typically be reduced to a scoring algorithm of the relative probabilities of candidate solutions. Here we propose an algorithm for multi-hypothesis scoring which uses both positive as well as negative information: we name it the Negative and Positive Scoring (NAPS) metric. An early version of this concept was introduced in [27].

We define $NAPS_t$ as the ratio of the probability of a map matching hypothesis h_i when compared to a null hypothesis h_{null} both conditioned (at each timestep t) on the measurements $z_{1:t}$

$$NAPS(h_i) = \sum_{t=1}^{t} NAPS_t(h_i) \tag{14}$$

$$NAPS_t(h_i) = \ln\left(\frac{p(h_i|z_t)}{p(h_{null}|z_t)}\right) \tag{15}$$

We define a hypothesis h as the combination of an estimate of the graph structure of the SLAM problem x_h (the vehicle trajectory and all detected features) as well as all data association matches Ω of these features to map features in the prior map. The null hypothesis h_{null} is a special version of this hypothesis in which no data associations exist and in which it is proposed that each detected feature is a new feature independent of the map. We use it as normalization for maps of growing size.

Using Bayes' Rule gives

$$NAPS_t(h_i) = \ln\left(\frac{p(z_t|h_i)p(h_i)}{p(z_t|h_{null})p(h_{null})}\right) \tag{16}$$

We split $p(z|h)$ into two terms representing both negative and positive information

$$p(z_t|h) = \eta p(z_{t,pos}|h)p(z_{t,neg}|h) \tag{17}$$

Positive information is then defined, in the same way as for JCBB, as the likelihood of the measurements given the hypothesis

$$\begin{aligned} p(z_{t,pos}|h) &= \eta_{z,pos} e^{-\frac{1}{2}(x_h - z_t)^{\mathrm{T}}\Sigma^{-1}(x_h - z_t)} \\ &= \eta_{z,pos} e^{-D_h} \end{aligned}$$

where Σ represents the covariance, $\eta_{z,pos}$ is a normalization constant and D_h the Mahalanobis distance.

The term $p(h)$ represents a prior probability of a particular map hypothesis being created by the robot which we propose is directly related to the number of features N_f matched to the prior map is given by

$$p(h) = \eta_x e^{\lambda N_f} \tag{18}$$

where η_x is a normalization constant, λ is a free parameter. N_f is an integer between zero and the total number of features in the prior map. While this formulation does not take into account aspects such as a feature's measured visibility or other such specific terms, it does give us a measure of the confidence of a map match.

Combining these terms and canceling where possible gives the following expressions for NAPS and as well as more common positive-only scoring (POS) metrics:

$$NAPS_t(h) = -D_h + \lambda N_f + C_{h,neg} \tag{19}$$

$$POS_t(h) = -D_h + \lambda N_f \tag{20}$$

This specifically indicates the contribution of Negative Information $C_{h,neg}$ that we believe has been neglected in typical multi-hypothesis scoring algorithms. POS algorithms (such as JCBB) implicitly

assume $C_{h,neg} = 0$ and do not account for it in scoring the hypotheses. Most approaches assume very-high λ: essentially selecting the hypotheses that match the most total features and then ordering those by Mahalanobis distance — as in the case of JCBB. A good overview of such algorithms is presented in [28], [29].

B. Evaluating Negative Information

We define Negative Information as:

$$\begin{aligned} C_{h,neg} &= \ln\left(\frac{p(z_{t,neg}|h)}{p(z_{t,neg}|h_{null})}\right) \\ &= \ln(p(z_{t,neg}|h)) - \ln(p(z_{t,neg}|h_{null})) \end{aligned} \tag{21}$$

As each hypothesis NAPS score will eventually be compared to one another, the second term need not be calculated.

For a particular hypothesis, consider an entire vehicle trajectory and the sonar footprint that it traced out (such as in Figure 10). Also consider a prior map feature which is located within this footprint but was not detected. We wish to measure the number of times that this feature ought to have been detected, given that trajectory. NI is formed as the product of the probability of each undetected feature given the hypothesized vehicle trajectory

$$\begin{aligned} p(z_{t,neg}|h) &= p(z_{t,neg,f_1} \cap \ldots \cap z_{t,neg,f_{n_u}}|h) \\ &= \prod_{f \in N_u} p(z_{t,neg,f}|h) \\ &= \prod_{f \in N_u} (1 - p(z_{t,pos,f}|h)) \end{aligned} \tag{22}$$

where s_t is whole area sensed during measurement z_t, thus:

$$p(z_{t,pos,f}|h) = \int_{p(f) \cap p(s_t)} v_f p(f) p(s_t) dA \tag{23}$$

where v_f is the visibility of feature f and $p(f)$ is the prior probability of that feature.

In words, the probability of not detecting each conditionally-independent feature is the product of one minus the probability of detecting each feature, integrated across the intersection of the PDF of each feature and the PDF of the scanned sensor area. This formulation is subject to the following assumptions: 1) the sensor occlusion model is well-defined and accurate, 2) all features are static, 3) feature detections are independent, and 4) feature visibility can be accurately modeled. This calculation, often intractable due

to complicated integration limits, theoretically defines the probability of a set of negative measurements $z_{t,neg}$ given sensed area s_t.

We approximate this measure using a grid-based approach in discrete cells around each composite measurement. The approach used an estimate of the total number of scans taken within a particular cell s_c. It is specified that to prevent spurious measurements from becoming incorrect features, there exists some number of positive detections n_d associated to the same particular position (as discussed in Section IV-B). This allows computation of the $p(z_{t,neg}|h)$ term for each feature f via the following binomial distribution:

$$p(z_{t,neg,f}|h) = \sum_{j=0}^{n_d-1} \binom{s_c}{j} v_f^{(j)} (1 - v_f)^{(s_c - j)} \tag{24}$$

These contributions are combined across all features predicted to lie in a mapped region via Equation 23 to give $p(z_{t,neg}|h)$.

As the location of a feature is not exactly known, we assume a Gaussian distribution and use a Gaussian weighting $w_{c,f}$ of feature f in cell c within a 95% confidence ellipse (for computational reasons) which is summed to give an estimate of the number of scans of feature f using

$$l_f = \sum_{c \in ellipse} w_{c,f} s_c \tag{25}$$

This estimate is then used to calculate both the Negative Information Contribution and the on-line estimate of feature visibility, via:

$$v_f = \frac{d_f}{l_f} \tag{26}$$

where d_f is simply the number of times feature f is detected (on that specific pass), according to data association in the current hypothesis.

In terms of implementation, this approach requires the simple summation across each sonar scan. For the small number of features detected in this type of mission and the small number of hypotheses, this extensive summation is tractable for our system in real time. The result of the metric is a positive value which scores a particular hypothesis more likely when its observations do not contradict the prior map.

In particular, combining Negative Information with the other (positive-only) metrics in Equation 19 allowed us to disambiguate similar locations along a row of otherwise indistinguishable features, as indicated in Figure 10.

While the AUV operated in the field this metric is evaluated for each hypothesis. The vehicle controls itself off of the most likely hypothesis, giving heading, speed and depth commands to the low level vehicle controller so as to travel to a set of preprogrammed way-points in the field. When the map match

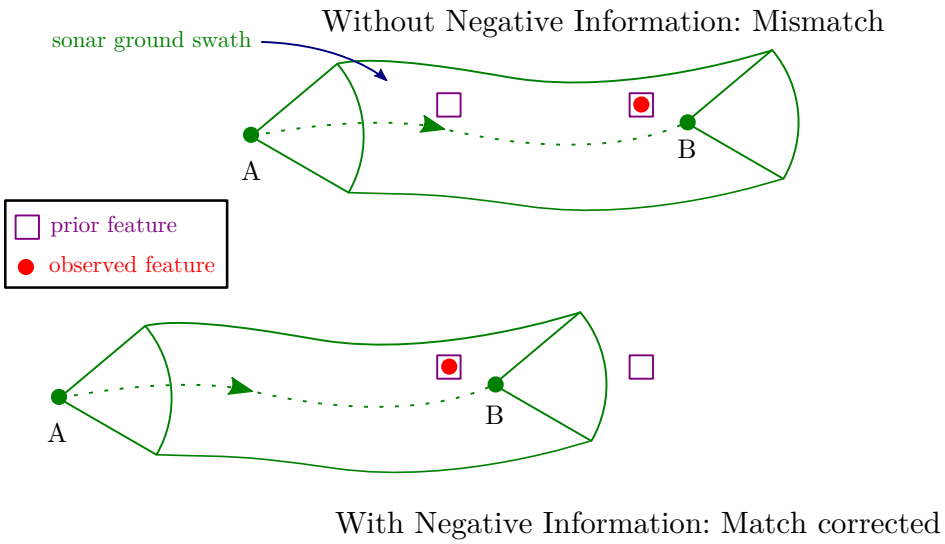

Fig. 10. *Illustration of the effect of Negative and Positive Scoring (NAPS). Consider the AUV trajectory from A to B with the sonar sensor footprint enclosed in green. If the AUV observes the red feature, how do we match it's trajectory to the prior map (purple squares)? Using JCBB the observed feature will be matched equally well to either prior feature. However using negative information, NAPS indicates that the match in the lower figure is more likely. The upper figure is less likely because we would have expected to have observed both features — but only observed one.*

metric for a particular hypothesis exceeds a threshold, it is decided that the AUV is matched to the prior map and switches to a final target capture mode.

C. Servoing to the Feature of Interest

Having decided upon a confident matching to the prior map, the vehicle then homes to the estimated location of the feature of interest. When it approaches this location, the FOI should be observed in the sonar imagery. The mission controller then transitions to direct control using the local sonar detections and a PID control on heading — which we call sonar servoing. It opens a pair of tines with a tip separation of approximately 1m and drives onto the mooring line of the FOI. Detecting the abrupt change in velocity due to the line, the tine controller will clasp around the FOI ending the mission.

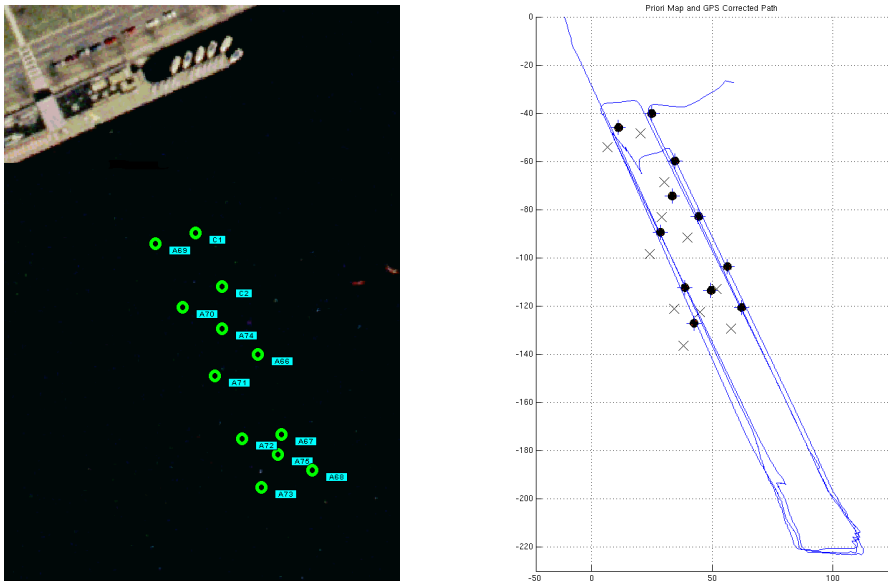

Fig. 11. *Feature layout and vehicle path for an initial experiment (November, 2006), demonstrating the basic capabilities of feature detection, local map construction, and global map matching running in real-time onboard the Ranger UUV. Terminal homing to the feature of interest was added in subsequent experiments. Units are in meters*

VI. Field Experiments

The system has undergone extensive testing and evolution over a five year period. Starting in November, 2006 we have conducted approximately 14 sea trials, each lasting 2 to 3 weeks. Our experiments began in fairly benign environments, using highly reflective moored objected as features, and progressed to more challenging conditions such as natural sea bottom features and strong currents. After each trial we have refined and improved the system. In the following we summarize the progress of the development of the algorithms and the vehicle platform.

Typically the ingress point and direction to the field were varied for each mission while the choice of feature of interest was taken at random just before placing the AUV in the water. After reaching the field, the vehicle typically traveled along the rows of features indicated in Figure 2. This was so as to keep the number of map match hypotheses low (to about 4–5). The typical mission duration was 15–25 minutes, although the mission planner could be programmed to repeat the mission if the AUV failed to find the feature field. During a successful mission the AUV spent in the region of 4–5 minutes actually in the feature field. In all these missions a 45° sonar sensor was used. The typical water depth was 15m.

Detailed comparison of mission parameters is difficult because the effect of the vehicle's control

decisions is that different paths (and observations) follow. For this reason, this section describes the progression of our core map matching algorithm.

1) Charles River and Boston Harbor, 2006 and 2007: A series of initial experiments were carried out to establish the basic operation of the system in easier conditions. An overview of those missions is shown in Figure 11.

2) St. Andrews Bay, Florida, June 2007: The NAPS and JCBB criteria were alternately used over 18 trials on a field of strongly reflective moored features. The JCBB implementation uses a threshold on the Mahalanobis distance for multiple pair matching and chooses the most compatible pairs. The results of this live test and selected other tests are summarized in Table I.

In addition to the frequency of success we have also presented the unbiased significance of these trials which we calculated as

$$S = \sqrt{\frac{n_s(n - n_s)}{(n-1)n^2}} \tag{27}$$

where n is the number of trials and n_s is the number of successful trials. A lower values indicates a more confident result. We believe that the results from this trial demonstrate that the NAPS outperforms the simpler JCBB matching criteria for our application.

3) Narragansett Bay, Rhode Island, June 2008: Using the data from June 2007, significant improvements to our sonar processing algorithms allowed for improved detection of man-made and natural bottom features. This included the addition of an adaptive noise floor model discussed in Section III-A and a reimplementation in integer logic for increased efficiency. In addition we began to test the control to and attachment to the feature. The field for these tests consisted of various man-made and naturally occurring objects on the sea bottom as well as moored features. The bay had a significant tidal current comparable to the 0.5m/s velocity of the vehicle, which gave us substantial dead-reckoning errors.

In the 9 runs, we attached to the feature once and had two mechanical failures. In both cases the tine mechanism broke upon hitting the target mooring line. Thus the overall success rate of the sonar navigation system was 33%. After these tests the current model mentioned in Section II-B was developed.

4) Stillwater Test Pond, March 2009: In this man-made test pond, we performed tests specifically on final control to the FOI and on the tine attachment mechanism. Only one feature was used and there were no currents, thus this experiment is not comparable to the others in terms of perceptive complexity. The pond was approximately 100m wide at the location of the tests. The vehicle successfully latched onto the line in each of the 17 trials.

5) Gulf of Mexico, near Panama City, Florida, June 2009: The entire system was tested on a field of 12 bottom objects and 3 moored objects over a two week period. These experiments tested an improved

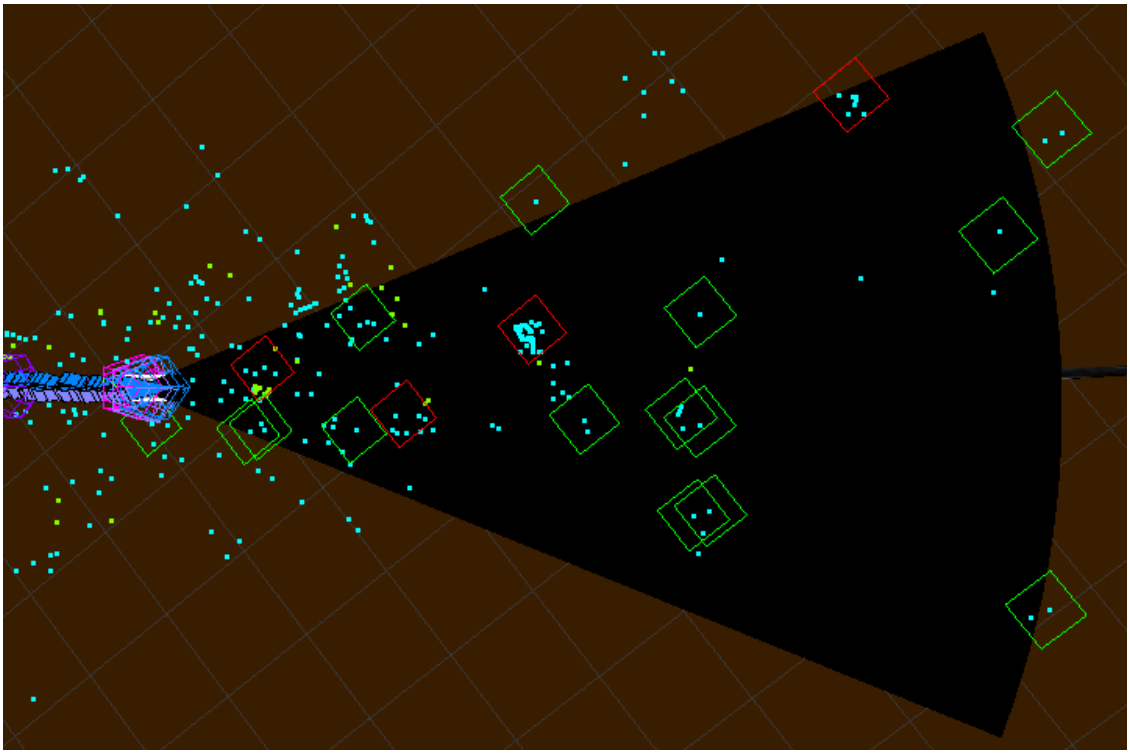

Fig. 12. *Screenshot of the data viewer for an experiment performed in Monterey, CA. Cyan dots are raw sonar detections, which are clustered as nursery features green squares. When justified, clustered nursery features are promoted to become FBN features, shown as red squares, and matched against the prior map.*

model for current estimation along with minor adjustments to the feature modeling. The current during this period was estimated as being 0.2m/s (using GPS surfaces). We had 17 successful feature attachments in 26 runs.

6) Gulf of Mexico, July 2010: The final set of experiments with the Ranger AUV were carried out. In this test, we observed much higher currents. These currents varied significantly from day to day but were estimated to be greater than the vehicle velocity (greater than 0.5m/s) on certain days. In this conditions, the vehicle could not make any headway against the current when it found itself down-current from the feature field.

Presented in Table I are two different results for this experiment. One result gives the overall percent success when including all of the 42 runs carried out: 31%. Filtering the runs to the 18 runs in which the AUV was able to enter the field (as defined by at least a single feature detection in the sonar) produced a success percentage of 72%. We believe that this value is more in fitting with the performance of

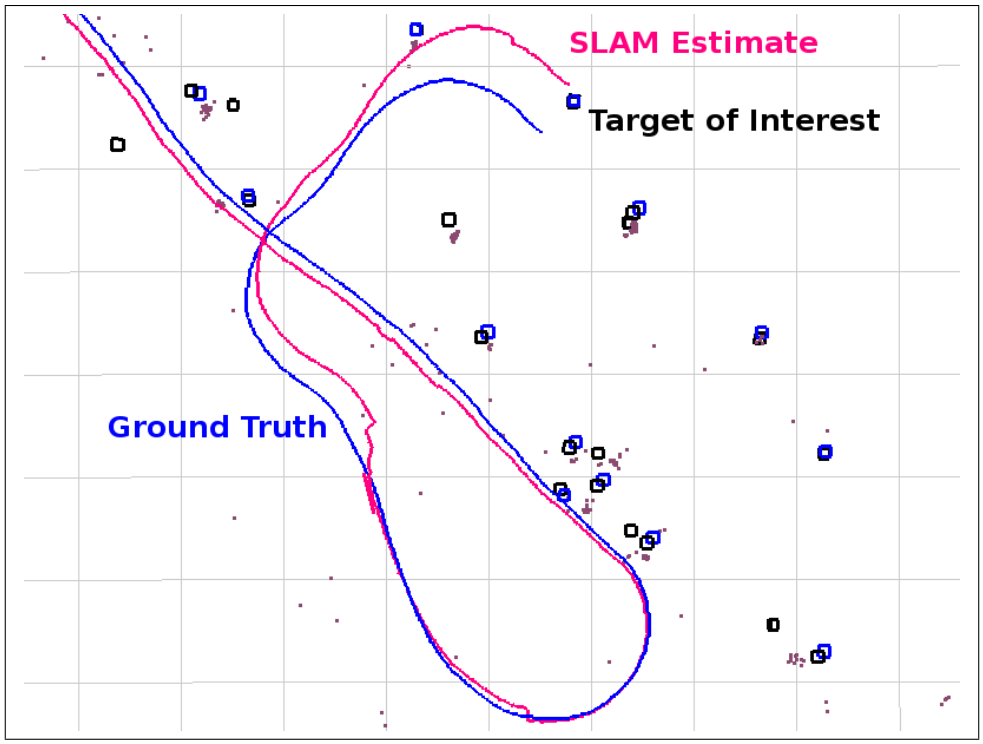

Fig. 13. *A top-down overview of a successful mission using the REMUS 100 vehicle. The vehicle approached from the north-west and extracted feature points (purple dots). Using these points and the prior map (blue squares), the SLAM map (black squares) and the vehicle trajectory estimate (magenta line) was formed. Having matched against the map the vehicle homed to the feature of interest. The abrupt position changes are the result of new updates from the smoothing algorithm. The scale of the grid is 10m. It is important to note that the DVL-INS-enabled AUV would have failed to reacquire the FOI without using sonar as the map itself was only accurate to 5m (blue line).*

the SLAM algorithm and comparable to the previous year's results. Nonetheless, this demonstrates the limitation of this particular vehicle platform as well as ocean current estimation without a direct sensor.

7) Gulf of Mexico, May 2011: An additional series of tests was performed with the system running in real-time onboard a REMUS AUV (without using its DVL). The motivation for porting the software to the REMUS was to be able to perform tests in a wider range of currents, and to compare trajectories estimated by the vehicle against the trajectory estimated by the REMUS's integrated Doppler-inertial navigation system. While further fields tests are required to fully evaluate the system's performance, we were able to achieve a number of successful target reacquisitions, thereby demonstrating that the FBN approach is applicable to multiple types of AUVs. The results for a typical successful reacquisition

Match Criteria	No. of Runs	Successes	Frequency	Significance
Bright Features - June 2007				
NAPS	9	6	67%	17%
JCBB	9	3	33%	17%
NAPS - JCBB (difference)			33%	24%
NAPS & Multi-hypothesis	18	14	78%	10%
Normal Features - June 2008				
NAPS Multi-hypothesis	9	3	33%	17%
Normal Features, Low Currents - June 2009				
NAPS Multi-hypothesis	26	17	65%	10%
Normal Features, High Currents - June 2010				
NAPS Multi-hypothesis	42	13	31%	7%
As above, having reached the field - June 2010				
NAPS Multi-hypothesis	18	13	72%	11%

TABLE I

SELECTED RESULTS FOR HOMING EXPERIMENTS IN DIFFERENT CONDITIONS, WITH AND WITHOUT USE OF NAPS.

mission with this vehicle are shown in Figure 13.

VII. CONCLUSIONS AND DISCUSSION

This paper has developed a feature reacquisition system for small low-cost AUVs, based on forward looking sonar-based SLAM and demonstrated its performance in challenging and realistic field conditions. Our results indicate that when the AUV correctly matches to the prior feature map, it is regularly able to re-visit a designated feature of interest. This was demonstrated by re-attachment using a pair of mechanical tines.

The main failure mode of the algorithm is failing to enter the feature field, due to disturbances that exceed the vehicle's control authority. For small to moderate ocean currents we developed an on-line current estimation procedure which allows the vehicle to avoid being driven off course during the initial

vehicle dive. However, it should be noted that failure to successfully enter the field is not in itself a system failure as the vehicle could surface for a GPS fix before trying again. Nonetheless, room exists to improve the current estimation procedure described in Section II-B.

Unsurprisingly, in currents of more than 50-70% of the vehicle's velocity, successful performance was limited. This presented an engineering limitation for this technology. As mentioned in previous sections, our most recent work as been focused in transferring this project to an AUV with a greater velocity.

Given the generic appearance of individual sonar features, robust loop closure is difficult in this domain. However with larger field of view sonars and denser feature fields, it may be possible to implement dense feature matching where more than a single feature is observed. This approach could well provide a single distinct loop closure constraint.

While more research is necessary to understand the many variables that can effect the system performance, such as the density and complexity of environmental features, the project has shown the viability of the FBN concept for feature reacquisition with low-cost vehicles.

Acknowledgments: This work was funded by the Office of Naval Research under grants N00014-05-1-0244 and N00014-11-10119, which we gratefully acknowledge. We would also like to thank our collaborators at iRobot, Blueview, Naval Postgraduate School, SeeByte Ltd. and Naval Surface Warfare Center Panama City for their assistance in this project, especially Doug Horner, Dan Kucik, Ed Matson, Bryan Schulz, Paul Carroll, and Scott Reed.

REFERENCES

[1] D. B. Kilfoyle and A. B. Baggeroer, "The current state-of-the-art in underwater acoustic telemetry," *IEEE J. Ocean Engineering*, vol. 25, no. 1, pp. 4–27, 2000.

[2] 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, 1998, pp. 45–54.

[3] J. J. Leonard, A. A. Bennett, C. M. Smith, and H. J. S. Feder, "Autonomous underwater vehicle navigation," MIT, Tech. Rep. Marine Robotics Laboratory Technical Memorandum 98-1, 1998.

[4] H. Durrant-Whyte and T. Bailey, "Simultaneous localisation and mapping (SLAM): Part I," *Robotics & Automation Magazine*, vol. 13, no. 2, pp. 99 –110, June 2006.

[5] B. Schulz, R. Hughes, R. Matson, R. Moody, and B. Hobson, "The development of a free-swimming UUV for mine neutralization," in *Proceedings of the IEEE/MTS OCEANS Conference and Exhibition*. IEEE, 2005, pp. 1443–1447.

[6] J. Long, J. Schumacher, N. Livingston, and M. Kemp, "Four flippers or two? tetrapodal swimming with an aquatic robot," *Bioinspiration & Biomimetics*, vol. 1, p. 20, 2006.

[7] J. C. Kinsey, R. M. Eustice, and L. L. Whitcomb, "A survey of underwater vehicle navigation: Recent advances and new challenges," in *IFAC Conference of Manoeuvering and Control of Marine Craft*, Lisbon, Portugal, Sept. 2006, invited paper.

[8] D. Meduna, "Terrain relative navigation for sensor-limited systems with application to underwater vehicles," Ph.D. dissertation, Stanford University, 2011.

[9] B. Cushman-Roisin and J. Beckers, *Introduction to Geophysical Fluid Dynamics*. Academic Press, 2011.

[10] R. Thompson and W. Zehner, "Frequency-steered acoustic beam forming system and process," 1999, US Patent 5,923,617.

[11] 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*, 2005, pp. 452–459.

[12] S. Williams, G. Dissanayake, and H. Durrant-Whyte, "Towards terrain-aided navigation for underwater robotics," *Advanced Robotics*, vol. 15, pp. 533–550, 2001.

[13] H. Johannsson, M. Kaess, B. Englot, F. Hover, and J. Leonard, "Imaging sonar-aided navigation for autonomous underwater harbor surveillance," in *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*, Taipei, Taiwan, Oct. 2010.

[14] I. Tena Ruiz, S. de Raucourt, Y. Petillot, and D. Lane, "Concurrent mapping and localization using sidescan sonar," *Journal of Oceanic Engineering*, vol. 29, no. 2, pp. 442–456, Apr. 2004.

[15] N. Fairfield, A. G. Kantor, and W. D., "Real-time SLAM with octree evidence grids for exploration in underwater tunnels," *J. of Field Robotics*, 2007.

[16] Y. Petillot, I. T. Ruiz, and D. M. Lane, "Underwater vehicle obstacle avoidance and path planning using a multi-beam forward looking sonar," *Journal of Oceanic Engineering*, vol. 26, pp. 240–251, Apr. 2001.

[17] D. E. Clark and J. Bell, "Bayesian multiple target tracking in forward scan sonar images using the PHD filter," *IEE Radar, Sonar and Navigation*, vol. 152, pp. 327–334, 2005.

[18] M. G. Dissanayake, P. Newman, S. Clark, H. Durrant-Whyte, and M. Csorba, "A solution to the simultaneous localization and map building (SLAM) problem," *IEEE Trans. Robotics*, vol. 17, no. 3, pp. 229–241, July 2001.

[19] P. Newman, J. J. Leonard, and R. Rikoski, "Towards constant-time SLAM on an autonomous underwater vehicle using synthetic aperture sonar," in *Intl. J. of Robotics Research*, 2003.

[20] M. Montemerlo, S. Thrun, D. Roller, and B. Wegbreit, "FastSLAM 2.0: An improved particle filtering algorithm for simultaneous localization and mapping that provably converges," in *Intl. Joint Conf. on Artificial Intelligence*. Morgan Kaufmann Publishers Inc., 2003, pp. 1151–1156.

[21] G. Grisetti, C. Stachniss, and W. Burgard, "Improved techniques for grid mapping with Rao-Blackwellized particle filters," *IEEE Trans. Robotics*, vol. 23, pp. 34–46, 2007.

[22] S. Julier and J. Uhlmann, "A counter example to the theory of simultaneous localization and map building," in *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, vol. 4, 2001, pp. 4238–4243.

[23] F. Dellaert, "Square Root SAM: Simultaneous location and mapping via square root information smoothing," in *Robotics: Science and Systems (RSS)*, Cambridge, MA, 2005.

[24] F. Dellaert and M. Kaess, "Square Root SAM: Simultaneous localization and mapping via square root information smoothing," *Intl. J. of Robotics Research*, vol. 25, no. 12, pp. 1181–1203, Dec. 2006.

[25] J. Folkesson, J. Leedekerken, R. Williams, and J. Leonard, "Feature tracking for underwater navigation using sonar," in *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*, San Diego, CA, Oct. 2007.

[26] M. Kaess, A. Ranganathan, and F. Dellaert, "iSAM: Incremental smoothing and mapping," *IEEE Trans. Robotics*, vol. 24, no. 6, pp. 1365–1378, Dec. 2008.

[27] J. Folkesson and J. Leonard, "Autonomy through SLAM for an underwater robot," in *Proc. of the Intl. Symp. of Robotics Research (ISRR)*, 2009.

[28] J. Neira, J. Tardos, and J. Castellanos, "Linear time vehicle relocation in SLAM," in *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, Sept. 2003, pp. 427–433.

[29] L. Paz, J. Guivant, J. Tardós, and J. Neira, "Data association in *O(n)* for divide and conquer SLAM," in *Robotics: Science and Systems,RSS*, Atlanta, GA, USA, June 2007.