Toward Real-Time Visually Augmented Navigation for Autonomous Search and Inspection of Ship Hulls and Port Facilities

Ryan M. Eustice
Dept. of Naval Architecture & Marine Engineering
University of Michigan

Abstract

This paper reports on current research to automate the task of ship hull inspection and search using autonomous underwater vehicles (AUVs). We describe an automated feature-based navigation (FBN) and mapping framework that provides the AUV with precise in-situ hull-relative localization. Our vision-based perception approach eliminates the need for having to deploy additional navigation infrastructure, such as acoustic beacons (a traditional method for obtaining precise bounded-error navigation). We describe our mapping framework and show how we are now applying that framework to the task of automated ship-hull inspection using the HAUV testbed. The operational impact to the Navy of this technology development will be rapid, repeatable, automated 100% survey coverage for ship-hull inspection.

1 Introduction

Present day means for ship hull and port facility inspection require either putting divers in the water or piloting a remotely operated vehicle (ROV) over the area of interest — both of which are manpower intensive and generally cannot guarantee 100% survey coverage. The Navy would benefit from being able to automate this task, allowing for autonomous robotic inspection of its ships and port facilities for foreign objects such as limpet mines or improvised explosive devices (IEDs) on a routine round-the-clock basis. Automating this task, however, is challenging and compounded by the fact that areas around ships in berth are severely confined, cluttered, and complex sensing environments (e.g., acoustically, optically, magnetically). Current tethered robotic inspection systems present issues of snagging, maneuver degradation, and tether management, all of which make maneuvering around the ship at pier difficult. Moreover, current robotic inspection methods require human in-the-loop intervention for both sensory interpretation and control (piloting). Navigation feedback in these scenarios is typically performed using acoustic transponder-based time-of-flight ranging. This necessitates setup and calibration of the associated acoustic-beacon navigation infrastructure, and therefore vitiates our ability to rapidly and repeatably inspect multiple underwater structures.

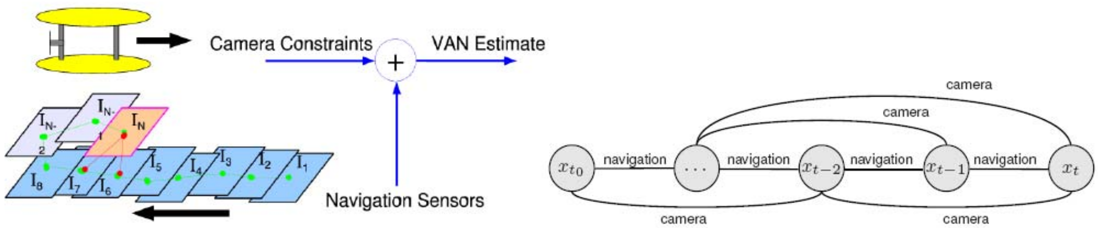
In light of this, there exists a need to automate this task through the use of untethered robotic vehicles. To do so with AUVs requires overcoming several present-day science and technology challenges inherent to the inspection task. For example, areas around ships in berth are severely confined, cluttered, and complex sensing environments (e.g., rudders, screws). This necessitates the need for advanced navigation and localization systems that can work in confined, magnetically noisy spaces. In addition, to ensure 100% survey coverage of ship hulls, pier structures, and pilings requires technological advances in our understanding of autonomous environmental perception and control. The underlying algorithm should facilitate in-situ sensor-reactive navigation and mapping in these environments while accommodating map-based learning through time via revisited exploration (a prerequisite for hull change detection). Moreover, the increased diversity of threat objects and associated potential for more false alarms due to a cluttered environment necessitates that fusion take place from multiple types of sensors for robustness and redundancy. In combination, all of these individual challenges/requirements, together, suggest that a feature-based navigation and mapping strategy approach would accommodate the needs of autonomous automated search and inspection by AUVs.

2 Technical Approach

The technical objective of this work is to develop an optical/acoustic real-time FBN capability for explosive ordinance disposal (EOD) autonomous ship-hull inspection. FBN is a vital requirement for autonomous robotic ship-hull inspection. Current robotic inspection methods require human in-the-loop intervention for both sensory interpretation and control (piloting). Navigation feedback in these scenarios is typically performed using acoustic transponder-based time-of-flight ranging – which necessitates setup, calibration, and infrastructure – and thereby vitiates the Navy's ability to rapidly and repeatably inspect multiple underwater structures.

2.1 System Overview

Figure 1 depicts core elements of the overall FBN methodology, called visually augmented navigation (VAN). The VAN framework uses visual perception to augment the onboard dead-reckon navigation capabilities of the unmanned underwater vehicle (UUV). VAN uses a pose-graph simultaneous localization and mapping (SLAM) framework [1-6] to incorporate pairwise constraints from overlapping sensor imagery. These constraints form edges in the pose-graph and constrain the vehicle position estimate to bounded precision. This type of view-based approach is ideally suited to the ship-hull inspection task since the goal is to provide 100% survey coverage of the hull with minimal trajectory redundancy.



(a) VAN methodology. (b) View-based stochastic map.

Figure 1: The core mechanism behind VAN's bounded precision is the fusion of "zero-drift" camera measurements with dead-reckoned navigation data. Camera constraints are fused with onboard navigation sensor data in a view-based stochastic map framework where the model is comprised of a pose-graph. Nodes correspond to historical robot poses and the edges representing either Markov (navigation) or non-Markov (camera) constraints.

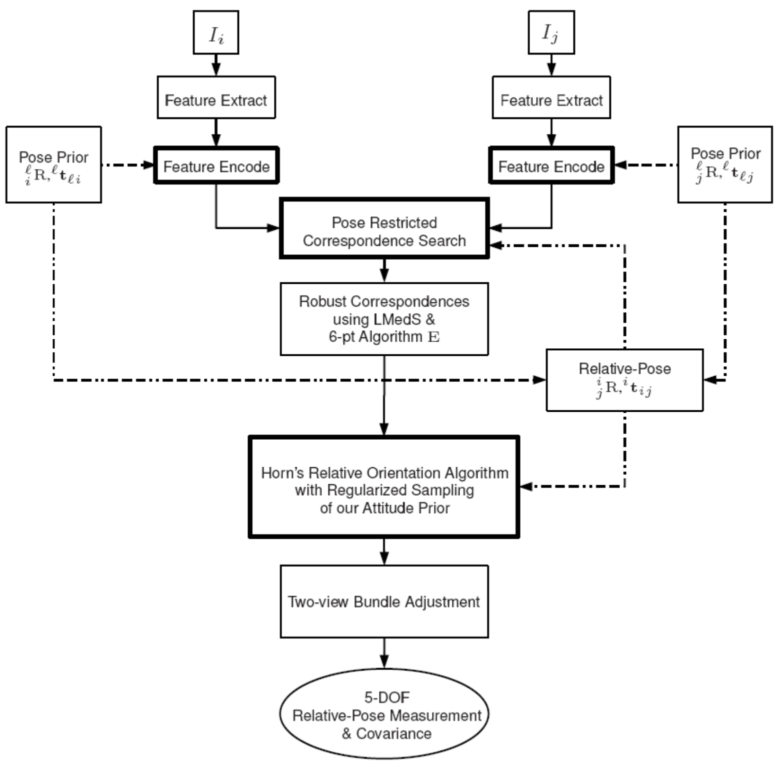


Figure 2: A block-diagram depicting VAN's systems-level approach to image registration. Dashed lines represent additional information provided by the VAN state estimate, while bold boxes represent systems-level extensions to a typical feature-based registration framework.

Camera measurements are used as constraints in VAN's pose-graph framework. Recursive inference is use to determine the global poses consistent with the camera measurements and navigation prior. VAN's goal is to be a real-time filtering algorithm, focused primarily on navigation and capable of scaling to large environments (image sequences consisting of thousands of key frames). To efficiently and robustly do so, VAN takes advantage of the complementary aspects of inertial sensing within the vision processing pipeline (as depicted in Figure 2).

VAN uses an extended information filter (EIF) (i.e., the dual of an extended Kalman filter (EKF)) for sensor fusion. The EIF representation results in an information matrix (i.e., inverse of the covariance matrix) that is exactly sparse without any approximation and allows for highly efficient, large-area scalable FBN [4-5]. For example, Figure 3 depicts the resulting information matrix associated with registering 866 images and fusing them with navigation data from a lawn-mower trajectory ROV survey of the RMS Titanic. The off-diagonal elements in the information matrix correspond to cross-track camera measurements while the block-tridiagonal structure naturally arises from the first-order Markov process model. The wreck was surveyed amidships to stern and then amidships to bow, requiring a total of 344 minutes to survey. During the course of the survey, the ROV traveled a total survey path length of 3.4 km. For this data set the VAN technique required only 39 minutes worth of processing time to update and maintain the global SLAM estimate, which is approximately a 9x speed-up over real-time.

$$\boldsymbol{\xi}_t = \begin{bmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_{866} \\ \mathbf{x}_r \end{bmatrix}$$

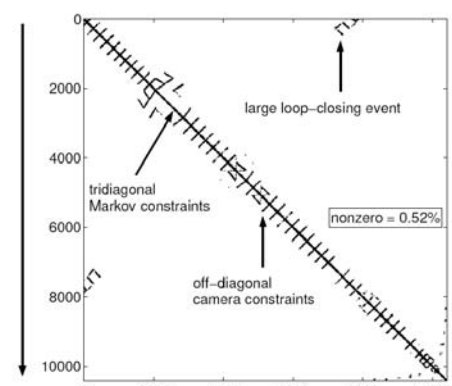
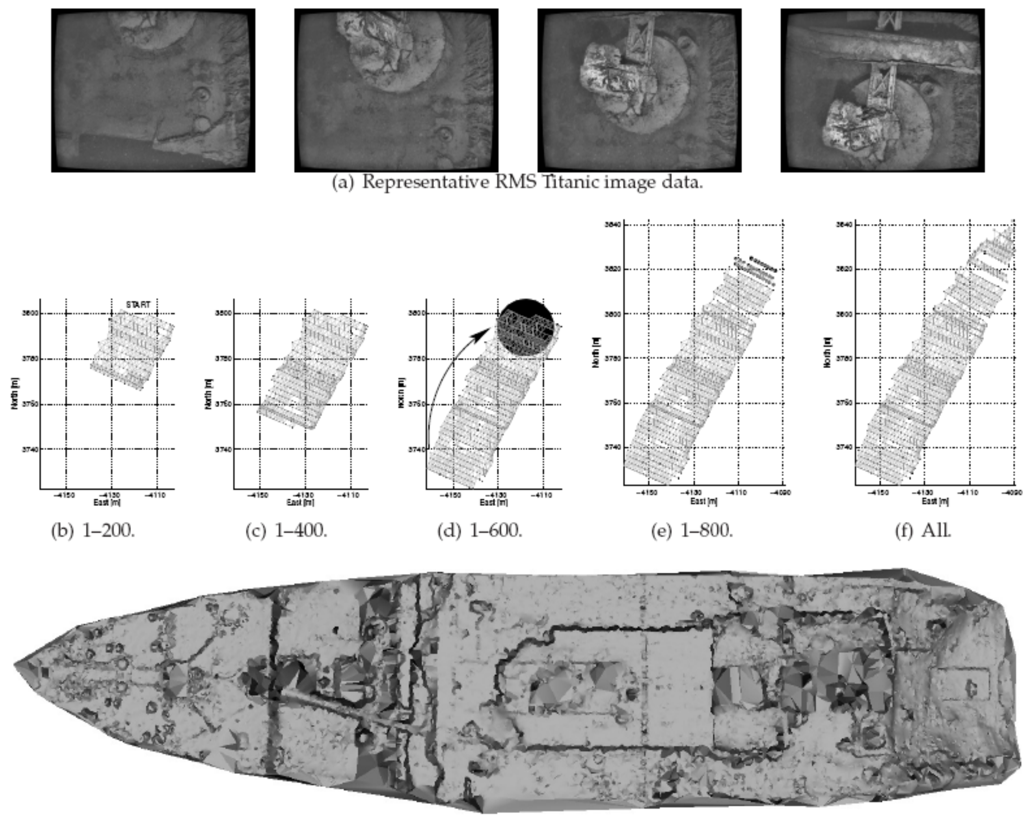


Figure 3: VAN uses a sparse EIF representation for inference. Shown is the information matrix for a ROV survey of the wreck of the RMS Titanic. In all there are 867 nodes where each state, $\mathbf{x}_i$, is a 12-vector consisting of 6-pose (i.e., Cartesian position and Euler attitude) and 6-kinematic components (i.e., linear and angular body-frame velocities). The resulting information matrix is a 10,404 x 10,404 matrix with only 0.52% nonzero elements. From [4-5].



(a) Representative RMS Titanic image data.

(b) 1–200. (c) 1–400. (d) 1–600. (e) 1–800. (f) All.

(g) Plan view of the Delaunay triangulated surface.

Figure 4: Results from a 2004 ROV survey of the RMS Titanic previously published by the Author in [4-5]. The mapped area encompasses over 3100 m^2 and over 3.4 km of traveled path length. (a) Four frame image sequence from the RMS Titanic data set depicting typical monocular image overlap and representative 3D structure content. (b)–(f) Time-evolution of the RMS Titanic pose constraint network. Camera links are shown in gray and 3-σ covariance bounds are depicted in black. Time progression is from left to right: images 1–200, 1–400, 1–600, 1–800, all. Note the large loop-closing event that occurs in (d) when the vehicle returns to the bow of the ship (depicted by the black arrow) after having traveled from the stern with the camera off. (g) The resulting Delaunay triangulated surface derived from the SLAM poses and image correspondences.

2.2 Relevance to Sonar-Based Ship Hull Inspection

While the proposed real-time VAN implementation is founded upon optical visual perception, the augmented navigation framework it comprises (Figure 5) is largely modality independent. The stages of:

- map building and maintenance,
- large-scale estimation,
- automatic link hypothesis,
- fusion of relative-pose and navigation constraints,

are all independent of *how* the pose constraints are derived. To replace the underwater image registration module with a DIDSON or BlueView forward-look sonar, for example, would require a different physics-based modeling of the sensor and the associated pose-constraints that can be extracted, but otherwise the rest of the VAN framework holds. Therefore, understanding the theoretical and technical issues behind developing a robust real-time optical VAN system will also lead to a better understanding of how to perform large-area autonomous search and inspection using other modalities such as sonar.

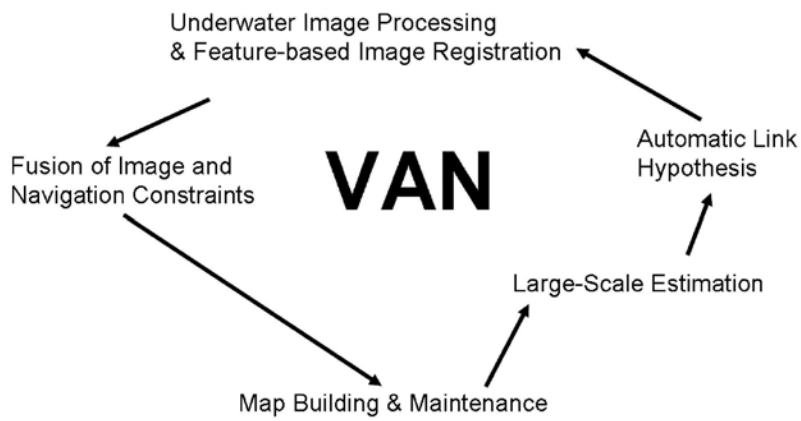
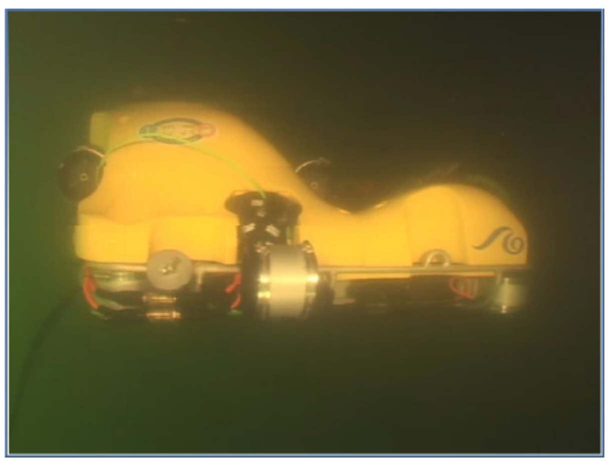


Figure 5: A depiction of the different areas involved in the visually augmented navigation (VAN) framework.

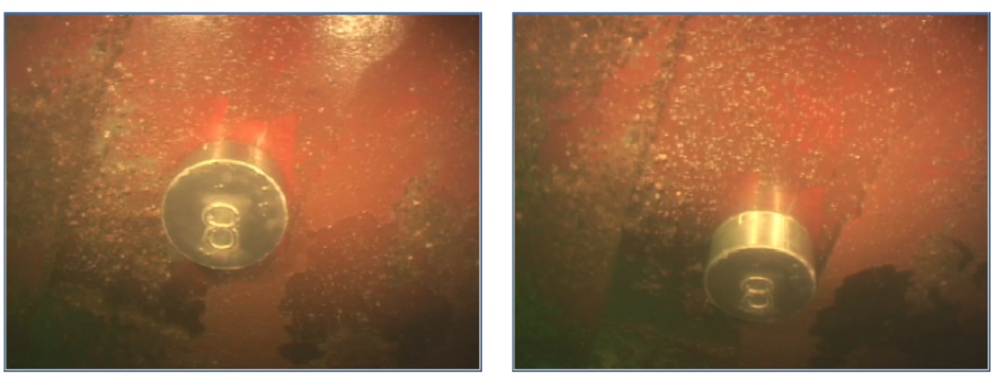
We are currently in Year 1 of a three year project to develop a real-time feature-based navigation system. Years 1 and 2 are focused on developing the overall mapping framework using vision as the main perceptual sensor. Year 3 of the project will investigate transitioning the VAN framework to sonar-based perception, which will yield a larger standoff range sensing capability in turbid water. The transition to sonar perception will require developing an appropriate sonar registration engine so that overlapping sonar images can be registered to extract pose-constraints similar to the optical registration engine of Figure 2. This work will explore the adaptation of feature extraction and description techniques developed within the computer vision community to the physics constraints of sonar imaging.

3 Real-Time VAN Testing and Development with the HAUV

We are currently collaborating with MIT and Bluefin Robotics to prototype and test our VAN algorithms on real-world ship hull inspection data using the Hovering-AUV (HAUV) (Figure 6) [7]. The vehicle is designed around a Doppler-based hull-relative navigation strategy using a 1200 kHz DVL mounted on a tilt actuator to measure vehicle velocities with respect to the ship hull for positioning. Open-water navigation to the hull is achieved using the DVL tilted toward the seafloor in bottom-lock mode aided by GPS surface fixes. Hull sensing is achieved using a Dual frequency IDentification SONar (DIDSON); this sonar modality was chosen for its ability to see through turbid water with high resolution. As an operation requirement, the DIDSON requires a grazing angle of 15°–20°, therefore, it also is mounted on a tilt actuator so that this particular graze angle can be maintained. Current work to date has demonstrated an ability to work on non-complex areas of the hull (i.e., flat areas of the hull like the sides and bottom and not complex regions like the rudder or screws), using a boustrophedon survey strategy with sonar-based hull mosaics produced in an offline post-processing step.



(a) HAUV-B ship hull inspection vehicle at MIT.



(b) Representative ship-hull imagery.

Figure 6: The HAUV and representative imagery from a 2007 field experiment where it performed inspection of a flat-bottom barge. The imagery is from a handheld diver video camera, but is representative of the type of imagery and hull texture that VAN will use for feature extraction and image registration. (Imagery courtesy J. Leonard)

To participate in hull-search experiments with the HAUV testbed, we have developed a strap-on camera bottle that can be easily mounted to the vehicle. The VAN hardware consists of a 12-bit Prosilica GigE camera and a 150 W remote light. This system can be run shore side from a laptop computer over the HAUV's fiber-optic tether. Real-world ship hull inspection data sets for post-processing algorithmic development will be collected at AUVFest'08 in this manner. Years 2 and 3 of the project will demonstrate real-time optical FBN ship-hull inspection using the HAUV with localization output being generated by our real-time VAN implementation. Year 3 will also begin transition of optical FBN algorithms to HAUV onboard sonar perception for inspection tasks.

4 Conclusion

Robotic perception and its coupling to autonomous navigation are key technical challenges in the quest for intelligent, robust, long-term, robotic autonomy. At the forefront of this research within the robotics community are the challenges of real-time visual perception and unconstrained 6-DOF motion – both of these aspects are inherent to the autonomous ship-hull inspection problem. Success in this arena will come in the form of in-situ perception-based navigation – eliminating the need for any external navigation infrastructure. The operational impact for the Navy is that this technology will allow for rapid automated inspection deployment with guaranteed 100% survey coverage; this will in turn reduce the need for having to put divers into the water to perform dangerous underwater search and ship hull inspection missions.

Acknowledgements

This work is funded by the Office of Naval Research through a grant from the Young Investigator Program and by Code 321OE (Ocean Engineering & Marine Systems) under award number N000140710791.

Bibliography

1. Eustice, R. M.; Singh, H. & Leonard, J. Exactly sparse delayed-state filters Proc. IEEE Intl. Conf. Robot. Auto., 2005, 2417-2424

2. Eustice, R. M.; Pizarro, O. & Singh, H. Visually augmented navigation in an unstructured environment using a delayed state history Proc. IEEE Intl. Conf. Robot. Auto., 2004, 1, 25-3

3. Eustice, R. M.; Singh, H.; Leonard, J. J.; Walter, M. R. & Ballard, R. Visually navigating the RMS Titanic with SLAM information filters Proc. Robotics: Science & Systems, MIT Press, 2005, 57-64

4. Eustice, R. M.; Singh, H. & Leonard, J. J. Exactly sparse delayed-state filters for view-based SLAM IEEE Trans. Robot., 2006, 22, 1100-1114

5. Eustice, R. M.; Singh, H.; Leonard, J. J. & Walter, M. R. Visually mapping the RMS Titanic: conservative covariance estimates for SLAM information filters Intl. J. Robotics Research, 2006, 25, 1223-1242

6. Eustice, R. M.; Pizarro, O. & Singh, H. Visually augmented navigation for autonomous underwater vehicles IEEE J. Oceanic Eng., 2007, In Print

7. Vaganay, J.; Elkins, M.; Willcox, S.; Hover, F.; Damus, R.; Desset, S.; Morash, J. & Polidoro, V. Ship hull inspection by hull-relative navigation and control Proc. IEEE/MTS OCEANS Conf. Exhib., 2005, 761-766