UC Davis
UC Davis Previously Published Works

Title
Teach-and-repeat path following for an autonomous underwater vehicle

Permalink
https://escholarship.org/uc/item/4s73v53p

Journal
Journal of Field Robotics, 35(5)

ISSN
1556-4959

Authors
King, Peter
Vardy, Andrew
Forrest, Alexander L

Publication Date
2018-08-01

DOI
10.1002/rob.21776

Peer reviewed

WILEY

Teach-and-repeat path following for an autonomous underwater vehicle

Peter King[1] (iD) | Andrew Vardy[2] | Alexander L. Forrest[1,3]

[1]Australian Maritime College, University of Tasmania, Australia

[2]Department of Computer Science, Department of Electrical and Computer Engineering, Memorial University of Newfoundland, St. John's, NL A1C 5S7, Canada

[3]Department of Civil and Environmental Engineering, University of California, Davis, Davis, California, 95616, USA

Correspondence
Peter King, Australian Maritime College, University of Tasmania, Australia.
Email: pdking@utas.edu.au

This paper presents a teach-and-repeat path-following method for an autonomous underwater vehicle (AUV) navigating long distances in environments where external navigation aides are denied. This method utilizes sonar images to construct a series of reference views along a path, stored as a topological map. The AUV can then renavigate along this path, either to return to the start location or to repeat the route. Utilizing unique assumptions about the sonar image-generation process, this system exhibits robust image-matching capabilities, providing observations to a discrete Bayesian filter that maintains an estimate of progress along the path. Image-matching also provides an estimate of offset from the path, allowing the AUV to correct its heading and effectively close the gap. Over a series of field trials, this system demonstrated online control of an AUV in the ocean environment of Holyrood Arm, Newfoundland and Labrador, Canada. The system was implemented on an International Submarine Engineering Ltd. Explorer AUV and performed multiple path completions over both a 1 and 5 km track. These trials illustrated an AUV operating in a fully autonomous mode, in which navigation was driven solely by sensor feedback and adaptive control. Path-following performance was as desired, with the AUV maintaining close offset to the path.

KEYWORDS
extreme environments, marine robotics, planning, underwater robotics

1 | INTRODUCTION

Autonomous underwater vehicles (AUVs) are free swimming robots that traverse some of the most remote and dynamic environments on Earth. These environments include beneath floating ice shelves (Jenkins et al., 2010) and under moving sea ice (Kaminski et al., 2010). In these scenarios, AUVs are required to traverse long distances away from a known safe launch-and-recover site. Where the environment is unknown and diverse, a path the AUV has taken previously may be one, or the only, safe passage to and from a particular site. In these instances, it is imperative that an AUV has the ability to retrace its steps and follow a path it has traversed before.

When robots have access to a consistent positioning aid, such as GPS, retracing a path is a simple matter of storing the positions along the path and following them in order. Due to the severe attenuation of high-frequency radio signals through water, AUVs do not have access to GPS signals when submerged and thus face a major challenge in maintaining an estimate of their location (Kinsey, Eustice, & Whitcomb, 2006). AUV localization solutions may be divided into three categories: inertial/dead reckoning, external aiding through acoustic beacons, and geophysical referencing (Paull, Saeedi, Seto, & Li, 2014). Inertial systems have an advantage in that they are self-contained,

but they suffer from cumulative errors as accelerometer biases are integrated into positions (McEwen & Thomas, 2003). Acoustic aiding provides a drift-free source of location, but a major financial and logistical cost is incurred to install the required infrastructure (Jakuba et al., 2008).

Geophysical referencing allows an AUV to position itself using sensory feedback. If a reference map is provided in advance of a mission, localization occurs by comparing local measurements against the map to limit the probable locations. When combined with knowledge of the vehicle motion, an estimate of position can be obtained; this is generally referred to as terrain relative navigation (TRN). Examples of TRN adapted for underwater vehicles are presented in Claus and Bachmayer (2015), Meduna, Rock, and McEwan (2008), and Rock, Hobson, and Houts (2014), with a general overview of systems provided by Chen, Li, Su, Chen, and Jiang (2015). Another approach is to simultaneously generate a map of the area and localize to it, known as simultaneous localization and mapping (SLAM). This is an active field of study, and many example systems exist (Mahon, Williams, Pizarro, & Johnson-Roberson, 2008; Paull et al., 2014; Tena Ruiz, de Raucourt, Petillot, & Lane, 2004).

An alternate geophysical approach is teach-and-repeat (TR) path-following, which does not require the estimation of position in the

global reference frame, but only with respect to previously collected data. This lack of positioning requirement allows for a less complex system than SLAM. TR enables an autonomous vehicle to refollow a path by relating its current sensory input to a stored sequence of sensory input from a previous traversal. TR localizes the agent relative only to the path, and it does not rely on a global localization of either the robot or locations along the path (Furgale & Barfoot, 2010; Matsumoto, Sakai, Inaba, & Inoue, 2000; Nguyen, Mann, Gosine, & Vardy, 2016). For long-range exploration missions, TR allows an AUV to venture into an unexplored area and return along the same path, regardless of its accumulated global position error.

Presented here is an adaption of TR for an AUV that builds upon similar work in the terrestrial robotics world, with adaptations to an underwater vehicle. The contribution of this work is the successful implementation of teach-and-repeat path-following on an autonomous underwater vehicle utilizing sonar as the primary imaging sensor. This implementation has been demonstrated with multiple successful field demonstrations of fully autonomous path-following in a true ocean environment over paths up to 5 km in length, with the vehicle under fully self-determined adaptive mission control. The success of this system relies on improvements made to image registration techniques, capitalizing on the unique aspects of the sonar imaging used in this work.

This paper is organized as follows: Section 2 provides an overview of the core image generation and matching methodology used for this work; Section 3 describes the implemented teach-and-repeat system; Section 4 presents results from field trials of the prototype system; and Section 5 is a discussion of the performance of the system and future directions for the work.

2 | METHODOLOGY

For a geophysical navigation system, how the environment is perceived is critical to its ability to localize and determine any required corrective actions. The primary development in adapting teach-and-repeat (TR) to an underwater vehicle was the use of images derived from available sonar data. This section provides an overview of techniques used for image generation and matching.

2.1 | Target platform

Development of this system was for Memorial University of Newfoundland's (MUN) Explorer AUV, shown in Figure 1. The MUN Explorer is a 4.5-m-long, 3000-m-depth rated vehicle manufactured by International Submarine Engineering (ISE) (ISE, 2016; Lewis et al., 2016). This survey class AUV can traverse distances over 100 km, and it navigates using a true-north sensing fiber optic gyroscope and a Doppler velocity log (DVL), capable of providing velocity relative to the seabed at altitudes less than 200 m. Additional operational specifications are given in Table 1.
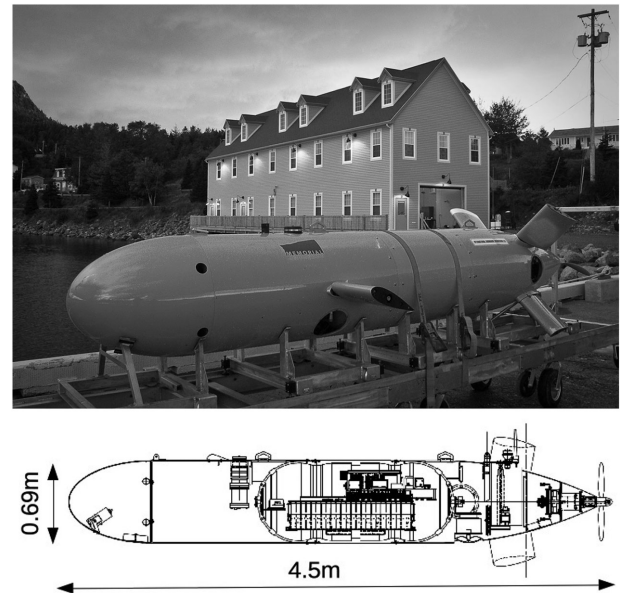

FIGURE 1 Memorial University of Newfoundland's Explorer AUV

2.2 | Vehicle integration

In addition to the AUV itself, our system requires two main computational components. The first requirement is a processing component that collects the raw sonar data, generates images, performs matching, and updates the navigation estimate. This was hosted on a dedicated payload computer connected to both the sonar and AUV control system through Ethernet. The second is a control interface to allow incoming requests from the payload computer to affect the movement of the AUV. The manufacturer, ISE, provided a set of short messages that would allow the processing computer to request control, provide a new target to navigate toward, and relinquish control.

At the start of the workflow, the AUV is given a mission script to follow, essentially a series of waypoints. This phase, referred to as the Discovery phase, is merely an intermediate step between the teach-and-repeat phases. Upon the TR system making a match and achieving a belief in its location above a threshold, *min_belief_localize*, a control request is sent to interrupt the current mission and await location targets; this is known as the Repeat phase. At any time the TR system can relinquish control and the AUV will return to the last executed step in the previously interrupted mission script. Figure 2 illustrates this workflow. It should be noted that in practice, returning to the previous executed mission may not be advisable, but it suited the test scenario.

2.3 | Image generation

For underwater vehicles, sonar is the primary means for sensory feedback from the environment, given the attenuation of optics in water (Al-Shamma'a, Shaw, & Saman, 2004). This work focuses specifically on sidescan sonar, for the following reasons: it is commonly found on survey class AUVs, it provides a value of reflected intensity that lends itself to the construction of a grayscale image, and it has an operational range that allows coverage over a large area while operating the AUV at safe distances from the seafloor.

TABLE 1 System specification of memorial university's explorer AUV

System	Components	Specification
Navigation	DGPS, iXBlue Fiber Optic INS, RDI Doppler Velocity Log (300 kHz)	Positional accuracy to 0.1% of distance traveled when submerged and relative seafloor velocity provided by DVL. Heading accuracy to 0.01°
Payload	Edgetech 2200 m, SeaBird CTD	400 kHz sidescan sonar. 200 m maximum range per side, 75 m for this work at 7m altitude. Processed resolution to 0.2m at 1.5m/s survey speed.
Processing	Payload Computer	2.1 GHz Intel i7, 4GB RAM, 128GB SSD
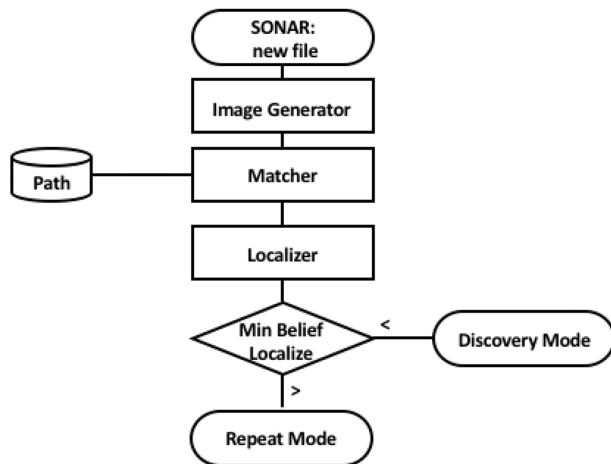
FIGURE 2 Flowchart of AUV decision to remain in Discovery phase, or switch to Repeat phase. Flow is from top to bottom, with final decision based on localization belief being either above, >, or below, <, a threshold

A sidescan sonar projects acoustic energy down and to each side of the AUV in a pattern that is narrow in the along track dimension and wide across track. This is illustrated in Figure 3. Reflected intensity decreases as the beam travels away from the AUV, increases when encountering a strong reflector, such as the illustrated rock, and may reduce to nil in the shadow zone behind obstructions, as can be seen on the side of the rock furthest from the AUV. The intensity of the measured reflected signal is dictated by the reflectivity of the seabed material, angle of incidence to the bottom or objects, occlusion, and the normal loss of sound intensity as it moves away from the source (Pinto, Ferreira, Matos, & Cruz, 2009). In addition, there will be an area of poor coverage directly below the AUV, known as the nadir. The sonar used in this work had a maximum operating range of 75 m from each side of the AUV, for a total coverage width of 150 m at a fixed altitude of 7 m.

The AUV measures its true heading and speed relative to the seabed to construct a two-dimensional projection of the sonar intensity by georeferencing sequential sonar pings onto a common north-up image grid of fixed resolution. These images assume a localized flat bottom with intensities corrected for attenuation such that relative intensity variations in the image are related to artifacts of the seafloor and not to propagation of the sound energy. Figure 4 provides examples of generated images. A more detailed description of the image-generation process used in this work is provided in King, Vardy, and Anstey (2012). A
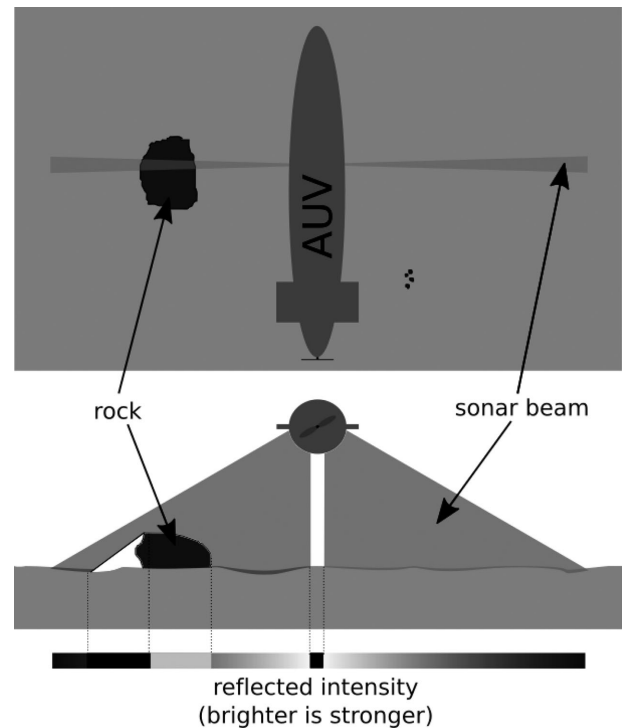
FIGURE 3 Top: Overhead view showing coverage area of a single sonar ping. The ping encounters a rock, which acts as a strong reflector. Middle: Rear view showing the occlusion of the sonar beam by the rock. Bottom: Relative intensity over the horizontal dimension. Intensity decreases as we move away from the AUV, becomes stronger as it reflects off the rock, then turns to shadow in the region occluded by the rock

distinguishing feature of sonar images is the blank nadir region, which is masked out in this work to ignore poor image quality directly below the vehicle.

It should be noted that in this work a fixed image size of 1,000 pings was used. This value was selected ad hoc, resulting in an image that was nearly square and continually performed well in testing. This is an area where further investigation could yield optimizations in both localization and efficiency.

2.4 | Image matching

For navigation, we must compare images from the repeat phase to those from the teach phase to determine how likely they are to be in
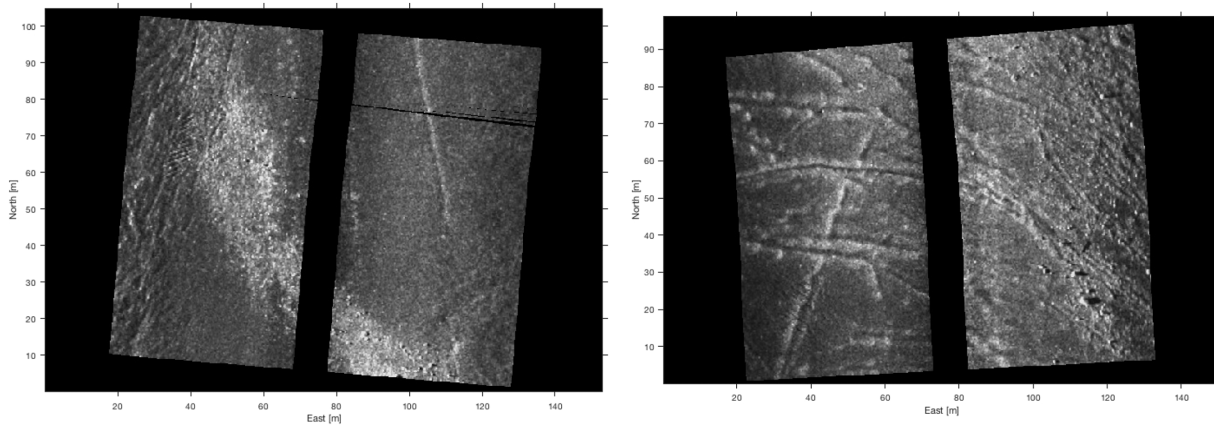
FIGURE 4 Examples of sonar images. The boundaries of the sonar data (including the nadir) are masked to ignore any deformations, hence the black regions. Left: A flat sandy area showing a patch of gravel. Right: A region with visible scour marks

the same location, and, should they match, how the AUV's estimated position and orientation differ. Image-matching techniques utilized in this work rely on feature extraction and matching and include SIFT (Lowe, 2004), SURF (Bay, Ess, Tuytelaars, & Gool, 2008), and FREAK (Alahi, Ortiz, & Vandergheynst, 2012), with implementations taken from the Open Source Computer Vision Library (OpenCV) (Itseez, 2015). Preliminary work on the selection and characterization of the chosen feature matching techniques as they relate to sonar images are described further in King, Anstey, and Vardy (2013) and Vandrish, Vardy, Walker, and Dobre (2011).

At its core, feature extraction is the process of locating areas of interest in an image, referred to as keypoints. These keypoints are expected to be consistently discernible from repeated views of the same image subject. Keypoints are described by the size of their included pixel neighborhood, as well as an orientation, derived by the dominant gradient component of the feature. Figure 5 illustrates a single sonar image and the detected set of keypoints drawn with relative size and direction indicators.

Image matching is the determination of the likelihood that two images represent an overlapping view of the seabed. If we assume that keypoints are fixed artifacts of a particular area of the seabed, then two overlapping views may contain the same keypoints. Thus, if two images are compared based on keypoints, we can infer a likelihood that they represent a common view of an area of seabed.

2.4.1 | Match filtering

The image match filtering process described here is a key contribution of this work. Its effectiveness stems from the unique characteristics of images produced from sonar data. The process is to compare the set of keypoints in one image to those of another image to form a set of candidate match pairs. In this work, an exhaustive matcher was used in which each keypoint in the first set is compared against all keypoints in the second set to select the closest match. This initial brute force pairing of best matches produces a large pool of pairs containing many likely false matches. The goal of the filtering process is to reduce this set to those pairs that are likely to be true matches.
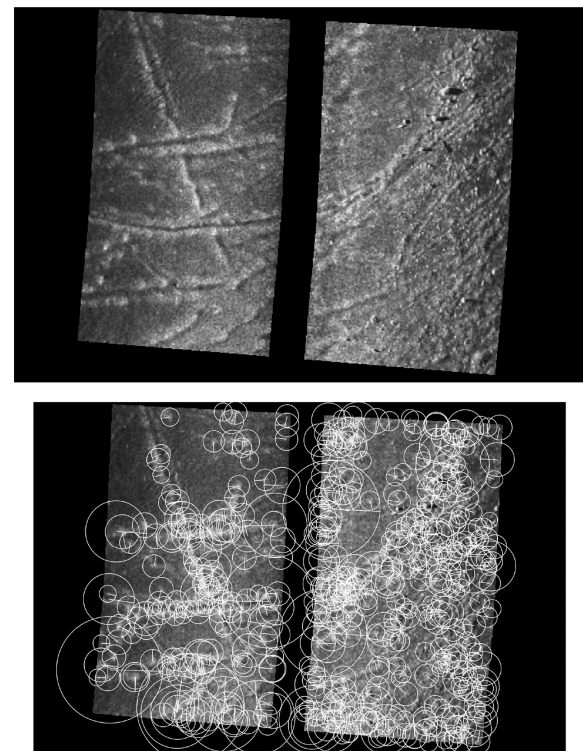

FIGURE 5 Top: Sonar image. Bottom: Same image with overlaid keypoint indicators, where circle size represents the size of the keypoint, and the inscribed line represents the orientation, or major gradient direction

To extract the true matches from all possible matches, we avail of two invariants:

1. Given the AUV's ability to directly measure its true heading, the generated images share a common north-up orientation.

2. Given the AUV's ability to measure its height above the seafloor, the generated images are projected onto a common flat plane, thus they share the same scale.

Therefore, common keypoints across images should be similar in both size and orientation, quite different from what is experienced in optical images of natural terrestrial scenes. These invariants hold up in

areas of relatively flat terrain but have been untested in areas with a sloping or more dynamic terrain.

The initial set of matched pairs are subjected to a filtering stage to disregard mismatched keypoint pairs. Allowing for some error in the AUV's orientation and ability to scale generated images, keypoint matches that differ in size or orientation beyond some nominal threshold are considered nonmatches, and they are discarded. Algorithm 1 includes the filtering steps. This threshold is a tuned parameter and, in these results, was set based on prior offline testing, described further in King et al. (2012).

Following this initial filtering, we investigate the remaining pairs, again exploiting our invariant that images are in a common north-up orientation that imposes the constraint that should the images match, the resulting geometric transformation between them must be solely translation, with negligible rotation or scaling. For a given pair of images with multiple feature matches, true-matches will represent a common feature translated from one image to another by a common offset. Conversely, the feature offset in a false-match will be unpredictably distributed.

Figure 6 illustrates this process, where two views, indicated by the dashed boxes, overlap over a set of features. The resultant images have the features located in differing locations, based on the orientation of the view. When compared, we can see how the features must translate from one image to another and that good feature matches have offsets consistent in both direction and displacement.

2.4.2 | Match consensus

To determine a consensus within the raw matches, given our already filtered set of matches and constraints on orientation, it is possible to employ a brute-force consensus algorithm that determines the largest possible inlier set of matches satisfying the invariant. Again, the ability to rely on a common scale and orientation greatly improves match performance. This differs from the more common use of methods based on the random sample consensus algorithm (RANSAC), which provides a suitable consensus within a group, but not necessarily the largest or optimal consensus (Fischler & Bolles, 1981).

This complete matching algorithm, provided in Algorithm 1, includes the steps for determining consensus. Every possible group of matched pairs, with similar orientation and lateral displacement within a specified threshold, is considered. The largest of these groups determines the estimate of geometric translation, with matches pairs outside this group rejected as outliers. Figure 7 illustrates the steps in image match determination for both the matching and nonmatching cases. Should no inlier group exist, or should it be smaller than a set threshold, no match is assumed to have occurred. In the event of two inlier groups having equal size, an event not seen in all testing to date, the system would only consider the first largest set encountered.

From the resultant set of matches, an average of the translation is taken as the difference in pose from the first image to the second. The size, or absence, of a dominant inlier set relates to the likelihood that

---

Algorithm 1 Matching algorithm

1: **function** MATCH(refImageSet, image)
2:      **for all** $refImg$ in $refImageSet$ **do**
3:          $matches \leftarrow$ IMAGEMATCHER($image, refImage$)   ▷ result of the image matching
4:          $goodMatches \leftarrow empty$
5:          **for all** m in $matches$ **do**   ▷ Filter keypoint matches
6:              **if** ABS($m.train.size - m.query.size$) $< sizeThreshold$ **then**
7:                  **if** ABS($m.train.angle - m.query.angle$)$< angleThreshold$ **then**
8:                      APPEND($goodMatches,m$)
9:                  **end if**
10:             **end if**
11:         **end for**
12:         **for all** m_i in $goodMatches$ **do**   ▷ Determine consensus for inlier keypoint matches
13:             $inliers_i \leftarrow empty$
14:             $projection \leftarrow (m_i.train.xy - m_i.query.xy)$
15:             **for all** m_j in $goodMatches$ **do**
16:                 **if** $m_i! = m_j$ **then**
17:                     $error \leftarrow m_j.train.xy - m_j.query.xy - projection$
18:                     **if** $error < minReprojection$ **then**
19:                         APPEND($inliers_i,m_j$)
20:                     **end if**
21:                 **end if**
22:             **end for**
23:         **end for**
24:         APPEND($inlier\_sets, inliers_i$)
25:     **end for**
26:     **return** $GetLargest(inlier\_sets)$
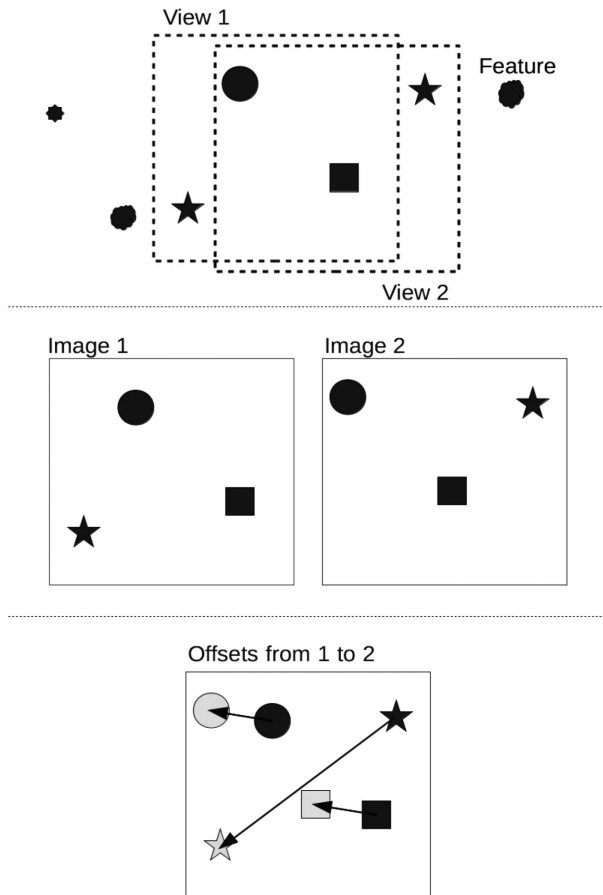27: **end function**

FIGURE 6 Top: Overlapping views of various features. Middle: Resultant images of features. Bottom: Match result between images; arrows show offset between matching keypoints. Short arrows are in agreement in both displacement and direction, longer arrow represents an outlier

the measurement occurs at a particular view in the path, described further in the next section.

3 | TEACH-AND-REPEAT PATH FOLLOWING

Teach-and-repeat path following allows a vehicle to navigate along a previously traversed path. The path, as described by Matsumoto et al. (2000), may be a sequence of view images, each with information describing its relation to the next image in the sequence. During training, images are added if they represent a significant change in view. During the repeat phase, image matching is employed to compare the current view against the estimated subsequent training image. If the subsequent image has a higher correspondence, the estimated location along the route is incremented. Image matching is used to determine the correspondence between images and the azimuthal offset between them. The offset is used to direct the robot toward the center of the path, allowing correction of horizontal offsets from the prescribed route (Matsumoto, Inaba, & Inoue, 1996).

An example of this type of system presented by Furgale and Barfoot (2010) models the path as a topologically connected set of submaps,
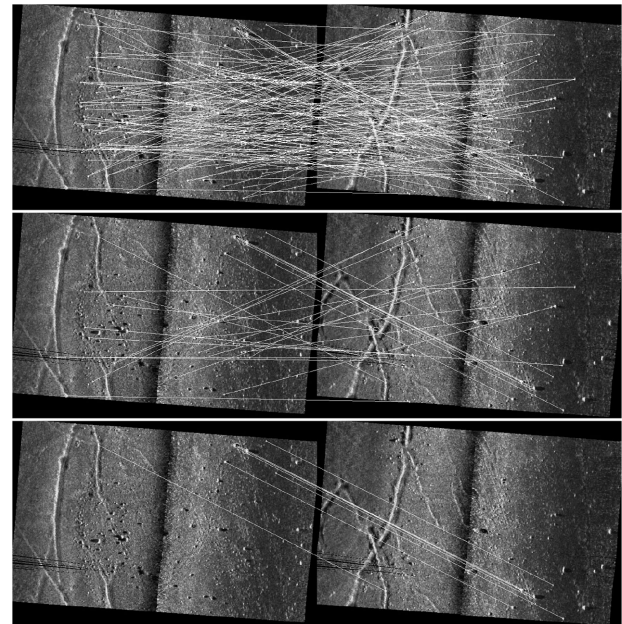

FIGURE 7 Top: Sonar images with raw keypoint matches. Middle: Remaining matches after filtering on size and orientation. Bottom: Only matches from the largest consensus group
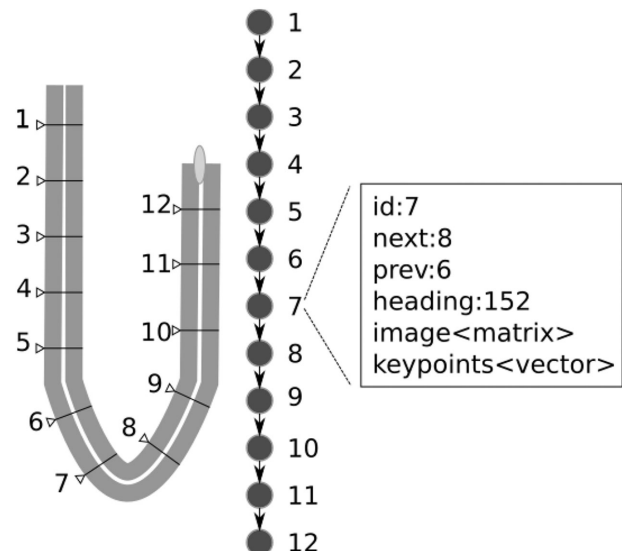

FIGURE 8 Representation of views as locations along a path: Left: Sonar images as there are collected by the AUV. Middle: The vector of nodes, where location is a single dimension index. Right: Data stored with each node, including ID of previous and next node, and average heading along the node

where only the submaps are described as locally Euclidean. Localization is based on a combination of feature detection and knowledge of the robot's motion. Several trials of the system are described, with the longest a 32 km route over rough outdoor terrain (Furgale & Barfoot, 2010). The motivation of their work is to support extra-planetary exploration by robots, a field not unlike that of exploratory AUVs.

In this work, the path is constructed by continuously generating sonar images, consisting of a fixed number of individual pings, of the seafloor and storing them as a connected series of images, with
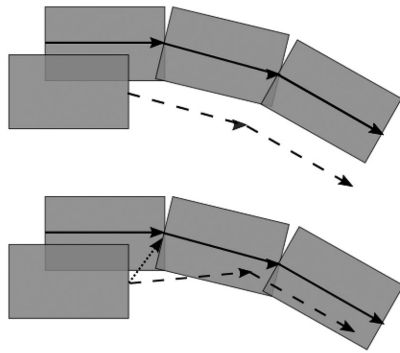
--- AUV Track   — Reference Track   ⋯⋯ Match vector

FIGURE 9 Top: AUV following path using the stored node-to-node vectors, but with an offset. Bottom: Corrective vector from image matching added to path vector so that the AUV will reduce its offset
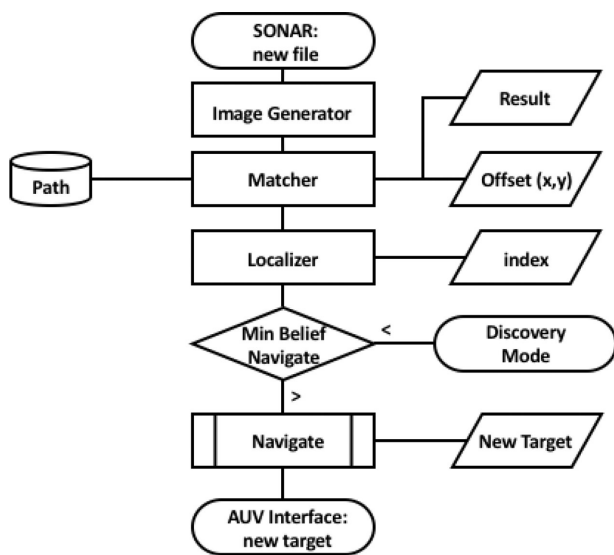

FIGURE 11 Flowchart of navigation determination

FIGURE 10 Flowchart of how a new sonar file is processed and the exchange of data between the major components

information of how each subsequent view connects to the next; we refer to this series of views as the reference path, and each individual view a node. As in Furgale and Barfoot (2010), the constraints are that the images are locally consistent with reasonable transformations between them. In the context of this work, the transformation is minimal as each view connects to the previous one such that they are essentially seamless. Unlike the work of Matsumoto et al. (2000), every completed image is added to the path, given they are each distinctive, and the AUV's internal navigation system provides the relative position of each image to the previously added.

Localization along the reference path is one-dimensional, as the AUV need only know at which node in the set it is most likely at. The arrangement of the nodes is topological, not globally topographic, such as in Matsumoto et al. (1996) and Matsumoto et al. (2000) where it is referred to as Visual Sequenced Route Representation (VSSR). We are simply concerned with the relative relation of views to each other and not the precise placement relative to a coordinate system, Figure 8 illustrates this concept.
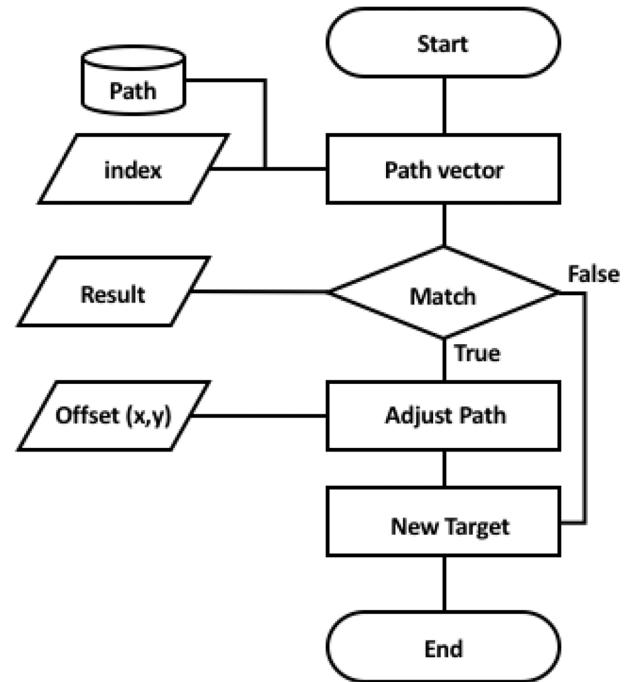
In the repeat stage, the AUV determines its present location relative to the reference path so that navigation actions may be determined to maintain its track along the path to completion. The likelihood value provided by the image matching gives a quantitative measure of which nodes the AUV's current view matches. But, this is insufficient as there is the possibility that (a) the AUV matches to multiple locations, (b) the AUV makes no match and thus gains no insight into its location, or (c) the AUV matches to a single incorrect location.

To maintain a robust estimate of the AUV's location along the route, a discrete probability filter, or Markov localization filter, is employed, similar to the along-route localization employed in Zhang and Kleeman (2009). The Markov localization filter maintains an array of probabilities, representing each node in the path, where the probability is the belief that the AUV is at that node. This array represents the probability distribution of the AUVs location along the route. At any given time, the array location of the distribution peak is taken as the current location estimate. If this peak is above a specified threshold, the vehicle is considered localized at the corresponding path node.

This distribution is updated in two steps: a prediction step, accounting for the AUV's motion; and an observation step, accounting for the current sensor view. This filter suits this problem due to its inherent discrete nature and the condition that the AUV may face phases of global uncertainty due to a difficult data association problem in which matches yield multiple hypotheses, as stated in Thrun, Burgard, and Fox (2005).

Generation of a new sonar image triggers the update cycle. At this point, the prediction and observation updates are made to the belief array. For each node, the predicted probability of the AUV being located at a particular node is based on the sum of probabilities that
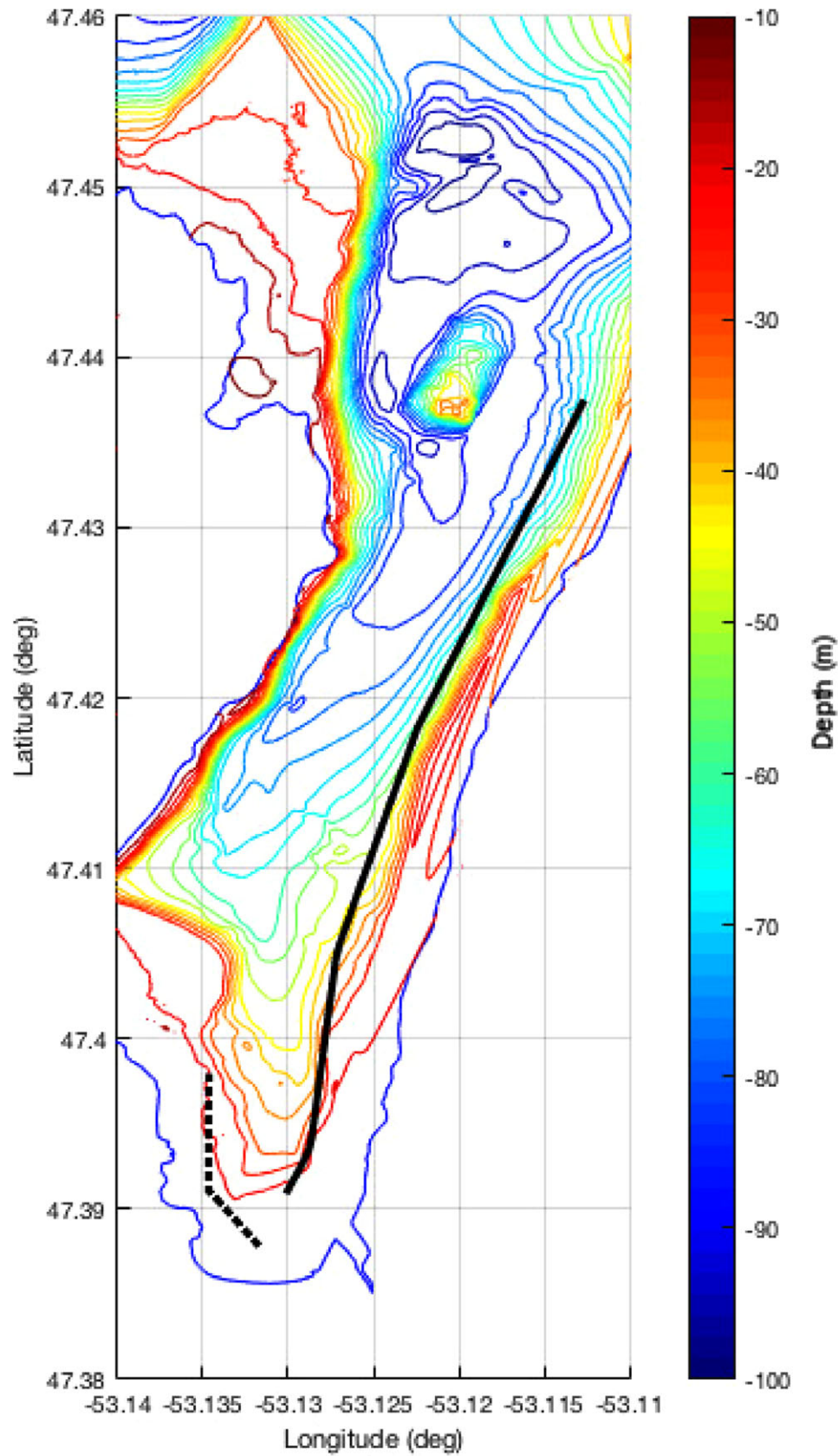
FIGURE 12 Reference paths for field trials shown in black. Shorter path used in 2013 trials and repeated in 2014. Longer path used in 2014 trials only. Location is Holyrood Arm, Newfoundland and Labrador, Canada

the AUV has moved to this node from all possible previous nodes, stated as

$$\bar{p}(n_{i,t}) = \sum_j p(n_{i,t}|u_t, n_{j,t-1})p(n_{j,t-1}) \tag{1}$$

$$= p(u_t = stay)p(n_i = n_j) + p(u_t = move)p(n_{i-1} = n_j),$$

where $\bar{p}(n_{i,t})$ is the estimated probability of being at node i at time step t. $p(n_{i,t}|u_t, n_{j,t-1})$ is the probability of moving to node i from node j, given some action u. In this system, there are only two possibilities of arriving at n_i: already being at n_i in the previous time step and not moving; and moving to n_i from the prior node in the sequence, n_{i-1}. For the linear-type paths attempted in this work and the fine control of the AUV, we assume the AUV can only remain at a node or move forward. For this work, we repeat the path at the same speed as was taught, thus we maintain a high probability of reaching the next node in one update cycle. For completeness, the probability of moving is as follows:

$$p(u = stay) = 1 - p(u = move). \tag{2}$$

The likelihood of an observation at a node results from the image matching and the current localization estimate outcome of the prediction step. As stated in the previous section, this likelihood, or measurement value, relates to the consensus of feature match pairs:

$$p(n_{i,t}) = p(z_t|n_{i,t})\bar{p}(n_{i,t}) \tag{3}$$

$$= \left(S + \frac{c_i}{\sum_N c_i}\right)\bar{p}(n_{i,t}),$$

TABLE 2 Test parameters

Parameter	Value	Description
grid_resolution	0.2	size in meters of each image pixel
size_threshold	10	the maximum allowable difference in feature size, given in pixels
angle_threshold	10	the maximum allowable difference in feature orientation, given in degrees
uniform_measuremer	0.2	seed value to avoid 0 belief conditions
min_belief_to_localize	0.8	minimum peak belief value to begin repeat phase
min_belief_to_navigate	0.3	minimum peak belief value to produce a navigation correction

TABLE 3 2013 trial results

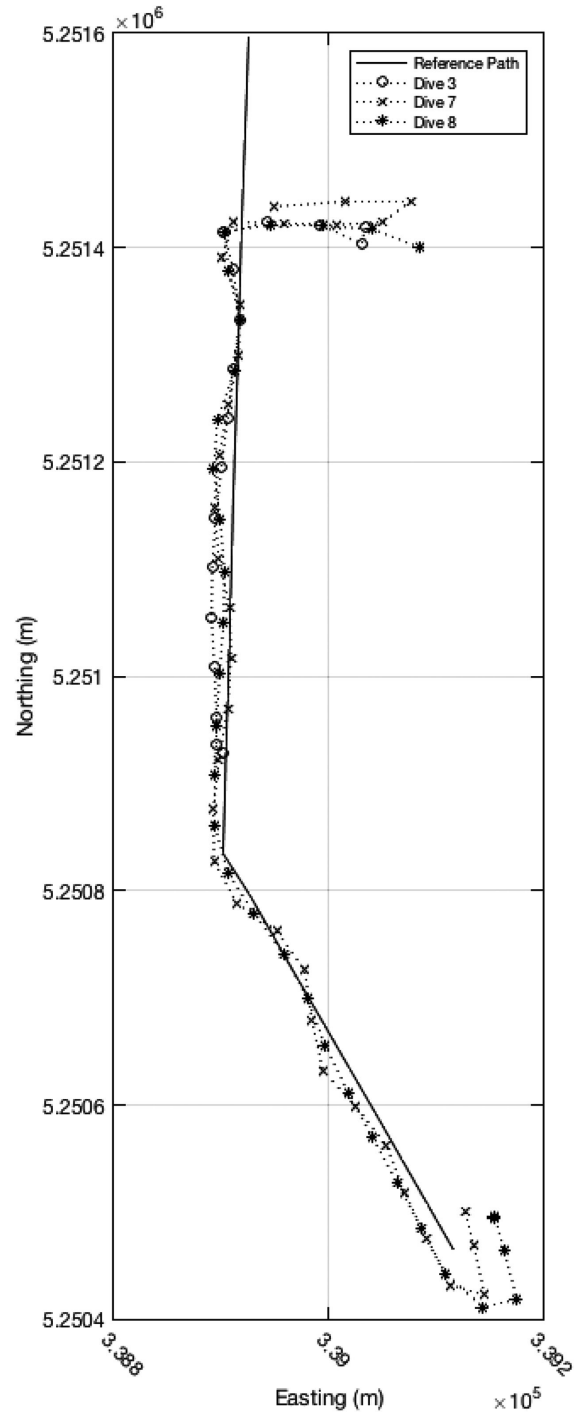| Dive	Local-izations	Mean Error (stdev)	% Dist. Traveled	Result
3	9	7.94 m (5.3)	1.24	Successful until premature abort due to incorrect out-of-bounds parameter
7	14	8.09 m (7.2)	0.55	Successful completion
8	17	5.71 m (4.3)	0.41	Successful completion

FIGURE 13 AUV track for three successful attempts at following the reference path

where $p(z_t|n_{i,t})$ is the probability that the measurement comes from matching to node i, S is a small seed value given to all nodes to avoid a zero belief condition, c_i is the number of matches in the largest inlier set when matching against the ith image in the path, and N is the total number of nodes in the path. In essence, the measurement is proportional to the number of good matches of one node and inversely proportional to the matches to all other nodes. Thus the likelihood measure grows as there are more good matches, but it is diluted as there are more total matches in other nodes.

The Markov filter algorithm is given in Algorithm 2.

Algorithm 2 Markov Localization

1: **function** LOCALIZE($p(n), c(n)$)
2:     **for** i in $p(n), c(n)$ **do**
3:         $\overline{p}(n_i) \leftarrow p(u = stay)p(n_{i-1}) + p(u = move)p(n_i)$ ▷ prediction
4:         $p(n_i) \leftarrow \left(S + \frac{c(n_i)}{\sum c(n)}\right) \overline{p}(n_i)$ ▷ observation
5:     **end for**
6:     $p(n) \leftarrow \frac{p(n)}{\sum p(n)}$
7:     **return** $p(n)$
8: **end function**

From each filtered match, an offset between images is provided, in pixels, for both the x and y axis. The mean offset is then calculated for all remaining matches and converted to a physical offset in both the north and east axis, based on the image resolution:

$$offset_{east} = offset_x \times grid\_resolution, \tag{4}$$

$$offset_{north} = offset_y \times grid\_resolution,$$

where *grid_resolution* is expressed as meters per pixel. The translation is the distance between the AUV's current position and its location when the reference view was captured. To maintain a consistent retracing of the path, the offset, with respect to the centerline of the path, should be reduced.

When a successful localization is made, a vector addition is performed between the offset vector and the path vector connecting the current estimated node to the next. There is no separation between each view—the end point of one image is adjacent to the start point of the next image in the reference path—thus, when localized, the AUV is assumed to be offset from the end point of the current estimated node, or at the start of the next node in the path. If we add the offset vector to the vector across the next node in the series, we will perform an action both to close the offset and to traverse the path.

When no localization is made, the navigation is based on the stored node-to-node vectors in the series from our predicted location.
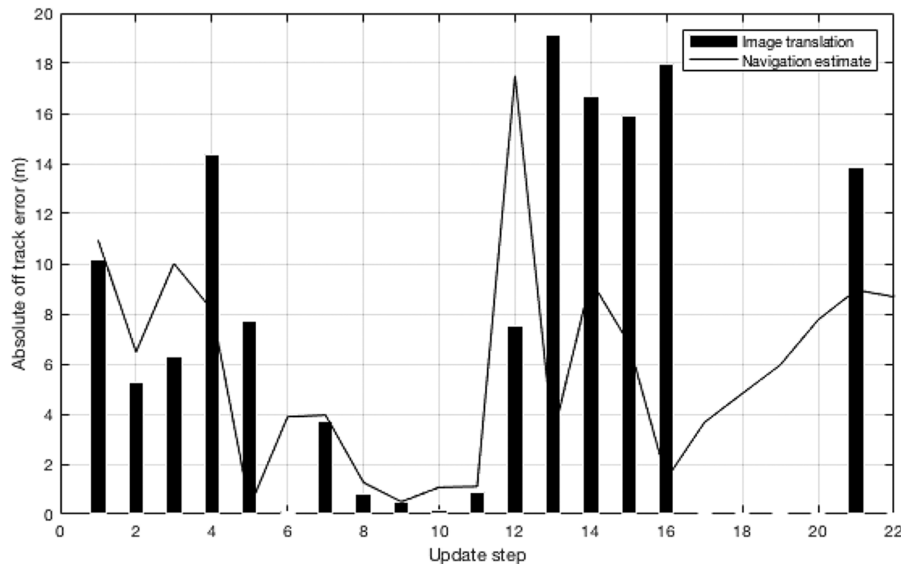

FIGURE 14 AUV off-track error for Dive 8. Bars indicate update steps, which included a successful match, with height indicating the measured distance between teach-and-repeat views. The line is the off-track error as measured by the AUV's onboard navigation system

TABLE 4 2014-a trial results

| Dive	Localizations	Mean Error (stdev)	% Dist. Traveled	Result
30	13	8.11 m (3.7)	0.43	Successful completion
37	14	7.30 m (6.5)	0.57	Successful completion
40	15	5.49 m (5.5)	0.43	Successful completion
41	13	11.52 m (6.2)	0.84	Successful completion
42	12	10.28 m (7.3)	0.80	Successful completion
67	5	14.63 m (17.0)	1.25	Successful, larger offset on 2nd leg due to lack of localizations, stored vectors allowed to track parallel
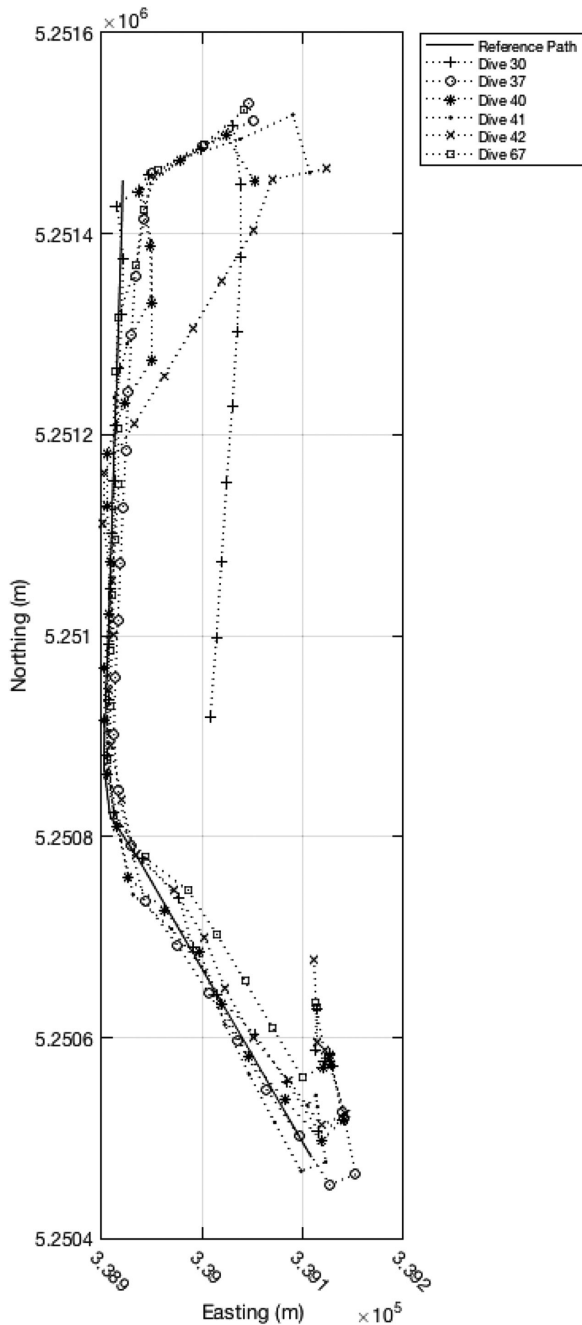
FIGURE 15 AUV track for follow-up attempts at short reference path

Closure of the path offset occurs when a localization is made and a corrective vector is added to the path vector as shown in Figure 9.

Figure 10 illustrates the processing of a new sonar file and the exchange of data between the major components. Figure 11 expands the calculation of a new navigation target.

4 | FIELD TRIALS

Tests of the complete teach-and-repeat system were conducted in May, 2013 and November, 2014, in Holyrood Arm, Newfoundland and Labrador, Canada. This is a sheltered 5 km inlet with water depths

ranging from 20 to 80 m. Prior surveys had been conducted in this area allowing empirical configuration of the sidescan parameters based on repeated processing of the previously collected sonar data.

Tests were performed on two predefined reference paths: a short path for initial testing in 2013, repeated in 2014; and a longer path only attempted in 2014. In each scenario, the reference mission plan was traversed in teach mode and the resulting reference path then utilized for multiple repeat attempts. Figure 12 show the reference paths used in testing. On each repeat attempt, the AUV conducted a prescribed mission that guaranteed it to both cross the reference path and then pass alongside while generating sonar images in an attempt to make an initial match and localization to the reference set—the discovery phase. If a strong match was made to provide the initial localization, the TR system requested an interrupt to the ongoing mission, entered the repeat phase, and attempted to follow the reference path to completion. The trials provided an opportunity to test all aspects of the TR system: image-matching, localization, navigation, and autonomous AUV control.

The shorter path consisted of a 400 m straight line section running north-south followed by a 600 m line to the south-east. The reference path held 26 individual image views. Discovery attempts for this path approached from the east, crossing near orthogonally. Given the time constraints on field testing, this approach was thought to be the most difficult and thus the worse-case scenario.

In the following field season, a repeat of the shorter path was conducted to ensure that continued development of the system and the vehicle had not affected the performance of the TR system. This development included bug fixes, refactoring for improved efficiency, and the inclusion of additional feature detection algorithms. In these tests, the same reference path route as the previous year was utilized, but with a shortened north-south line due to deployed fishing gear, giving 20 individual image views. A longer route was also used to further show the performance and utility of the system for extended operations. In this test, a south to north route following the eastern coastline was selected. This reference path extended approximately 5 km, with 61 views.

4.1 | Operating parameters

For the reference path and all attempted repeat runs, the vehicle operated at a speed of 1 m/s and a constant altitude of 7 m. These values had been used on previous AUV surveys and were known to produce high-quality sonar images. Images were constructed by combining 1001 individual sonar pings, a value that generally produced good results in both image generation and image matching in offline trials (Vandrish, Vardy, & King, 2012). This ping count also ensures a sufficient update period to allow for all necessary processing to occur. In the case of 1001 pings, processing takes 2.2 s, matching 0.004 s per pair, with tile generation occurring every 44 s. Complete test results for image generation and matching times are discussed further in King et al. (2012).

The core parameters related to matching and localization were initially set and remained unchanged throughout all attempts. Prior offline tests using precollected data from the same region informed the
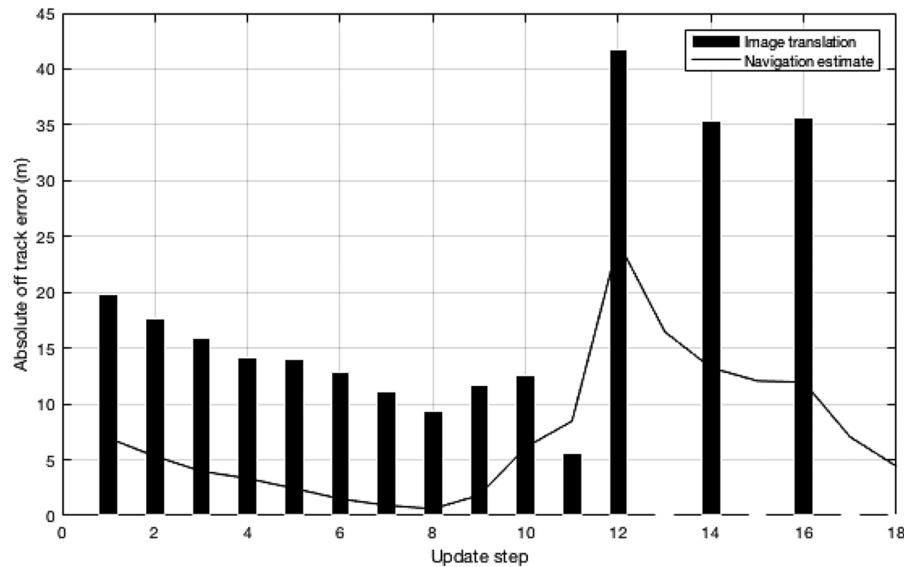
FIGURE 16 AUV off-track error for Dive 37. Bars indicate update steps, which included a successful match, with height indicating the measured distance between teach-and-repeat views. The line is the off-track error as measured by the AUV's onboard navigation system

TABLE 5 2014-B Trial Results

Dive	Localizations	Mean Error (stdev)	% Dist. Traveled	Result
48	24	5.20 m (4.4)	0.09	Successful, matches made throughout track
75	10	19.23 m (7.0)	0.38	Successful, matches made throughout track

selection of each parameter through trial and error. Table 2 provides the core parameters used throughout testing.

In 2013, only SURF feature extraction was employed. For the 2014 attempts, the system was modified to allow multiple feature extractors. In these instances, the extractors were SURF, SIFT, and FREAK. The final match pair sets from each extractor were combined prior to match filtering.

4.2 | 2013 trials

To test the complete teach-and-repeat system, a new module was added to the AUV control system to allow mission interruption and control. The integration of this system and its subsequent testing was a major priority of 2013 and thus only a single day of testing was allocated for TR field tests. Following an initial teach mission to acquire the reference path, a set of three repeat missions were conducted. Table 3 is a summary of each attempt, referenced by its dive number. The localization column is a count of how many successful observations occurred; a successful observation is one in which a match led to a correction providing the number of matches that led to a localization update. The mean error is based on the navigation estimate, seeded by the GPS at either end of the trial, and the % Dist. Traveled is the percent error versus the distance the AUV traveled during the repeat attempt. This metric is often used as an overall performance metric for localization systems; from Table 1 we see the best-case performance being 0.1%, though practical performance often approaches 0.5%. The Result column was the overall outcome of the test.

Overall the system performed as expected with the AUV able to discover the reference path and make sufficient corrective actions to maintain its track toward completion. The first test, dive-3, aborted early due to boundary parameters set in the mission interruption system. These boundaries are meant to prevent the AUV external control system from taking the vehicle outside a predetermined safe area. In this instance, the parameters were incorrectly set and resulted in an AUV fault condition and mission termination.

In each trial attempt, the AUV calculated suitable headings to close in on and maintain track on the reference path. Figure 13 shows the resulting tracks of each attempt. The plotted positions are the vehicle's onboard estimate of position. As there was no external tracking, the accuracy was limited by GPS fixes at either end for both the reference path collection and the repeat attempt. Figure 14 displays the off-track error from both the AUV's onboard navigation estimate and the measured image translation of any successful matchers over each update step of Dive 8. We observe that image matches generally align with changes in the off-track error, as expected.

We see a mean off-track error of 7.04 m, with a standard deviation of 5.8 m. What is evident from Dive 8 is that the error does not increase steadily as we traverse the path, such as would occur for an inertial navigation system.

4.3 | 2014 trials

Following the initial trials, further development continued on the TR system. This included implementing the navigation subsystem's ability to correct itself by setting a target waypoint rather than a heading. To test the relative performance to the previous years, the same mission track was used for the reference path generation, but with a shorter north-south line, due to an obstruction, i.e., deployed fishing gear. A
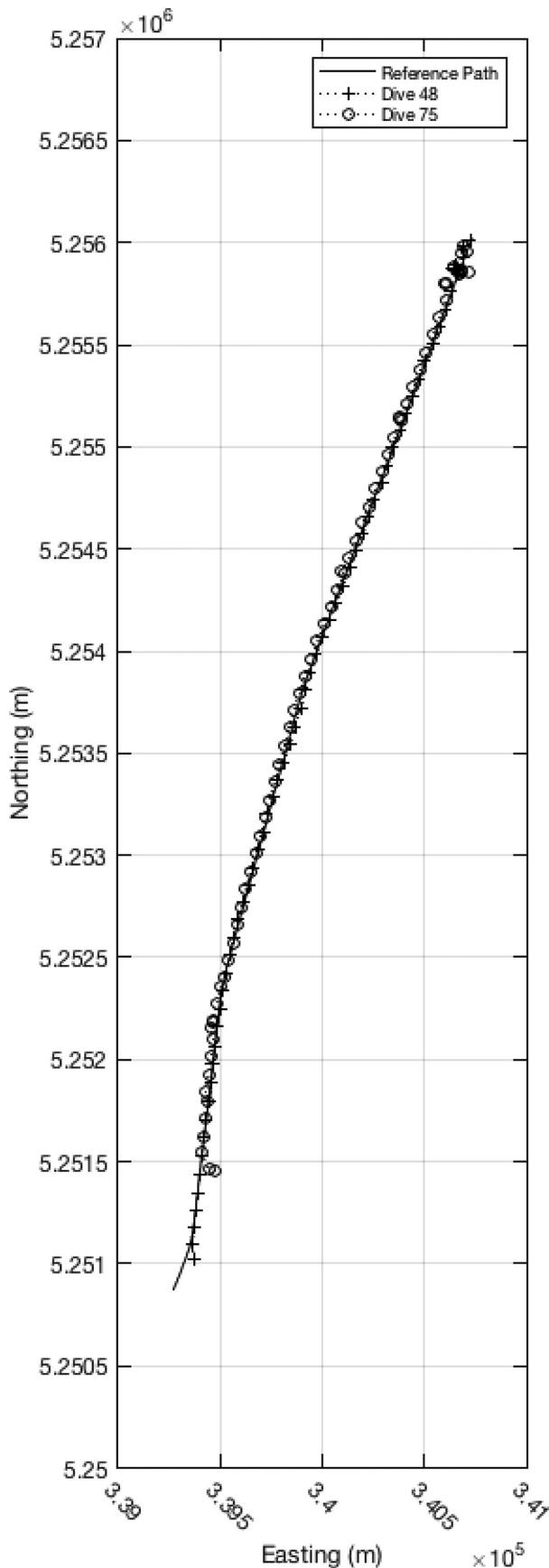
FIGURE 17 AUV tracks for long reference path

total of six repeat attempts were performed over this track. Table 4 summarizes the results of these attempts.

In each instance, the AUV was able to maintain a track along the reference path, even in instances of sparse matching, such as Dive 67. As for why there were so few matches on Dive 67, we are uncertain. Figure 15 illustrates the AUV tracks, and Figure 16 shows the off-track error and the measured image translation for each successful image match.

In addition to the shorter path, a longer path was attempted to gauge the longer-term performance, as the ultimate utility of this system would be for long-term exploration and path-following, where it would be unsuitable to rely on the internal navigation estimate alone. Again, waypoint targeting was used to drive the navigation along the path. Table 5 summarizes the results of the two repeat tracks.

As seen in Figure 17, the first attempt, Dive 48, is a great example of the system at work, with matches made throughout the track. Figure 18 is the off-track error for Dive 48 and the image translation for each successful match. Again, even in the longer path, we see the error is not monotonically increasing, but somewhat bounded and reacting to image matches.

The second attempt, Dive 75, brought to light an important issue with the waypoint targeting system. As the AUV proceeds along the path, new waypoint targets are generated in series, with the intent being that as the vehicle achieves a waypoint, the next subsequent target is issued unless a localization is made, in which case a new target is issued immediately and all subsequent waypoints are updated. In the final trial, as the AUV completed a waypoint, there was a delay before the next waypoint was issued. During this time, the AUV enacted its internal behavior for reaching a waypoint, which is to circle the point. As is evident in the path, there were instances where the vehicle began to circle waypoints until the new waypoint was issued. Upon the new waypoint being issued, the AUV then continued along the track. Figure 19 provides a closer look at this behavior, where we see large spikes in the heading setpoint that then reconvergence to the path heading.

It is important to note that on the longer attempts, we maintained an overall percent error of 0.09% and 0.38%. Both values are below that of the practical expected performance of a high-quality INS system, with Dive 48 beating even the best-case estimate of 0.01%. As this system continually corrects over the path, we would expect a bounded error and a decreasing error rate as the path length is extended.

5 | CONCLUSIONS

An autonomous teach-and-repeat path-following system has been presented for a survey grade AUV. This system is an adaption of similar work in the realm of terrestrial robotics. The core of this system is a sonar image generation and matching capability, which relies on well-known feature extraction techniques to compare a current view of the seabed to a set of previously collected views. Utilizing a quality indication of the resultant matches and Markov localization,
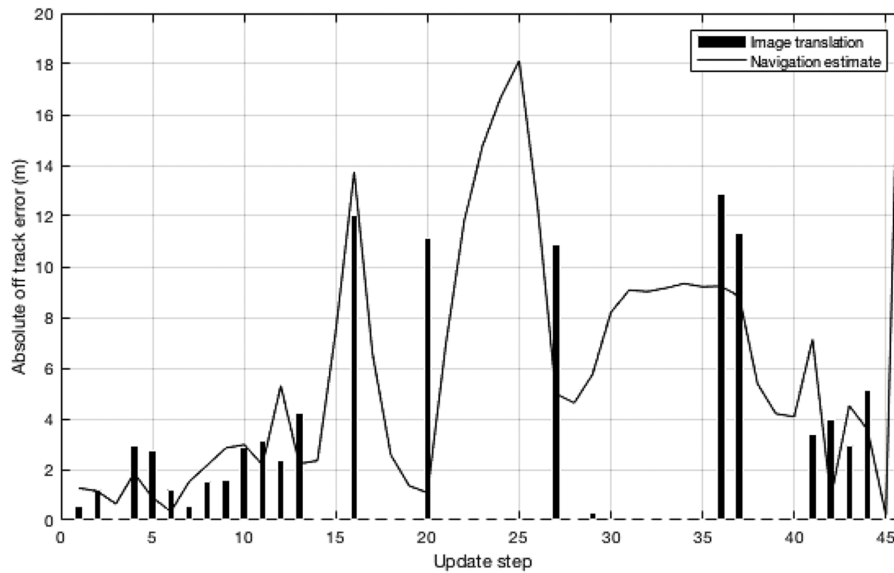
FIGURE 18 AUV off-track error for Dive 48. Bars indicate update steps, which included a successful match, with height indicating the measured distance between teach and repeat views. The line is the off-track error as measured by the AUV's onboard navigation system
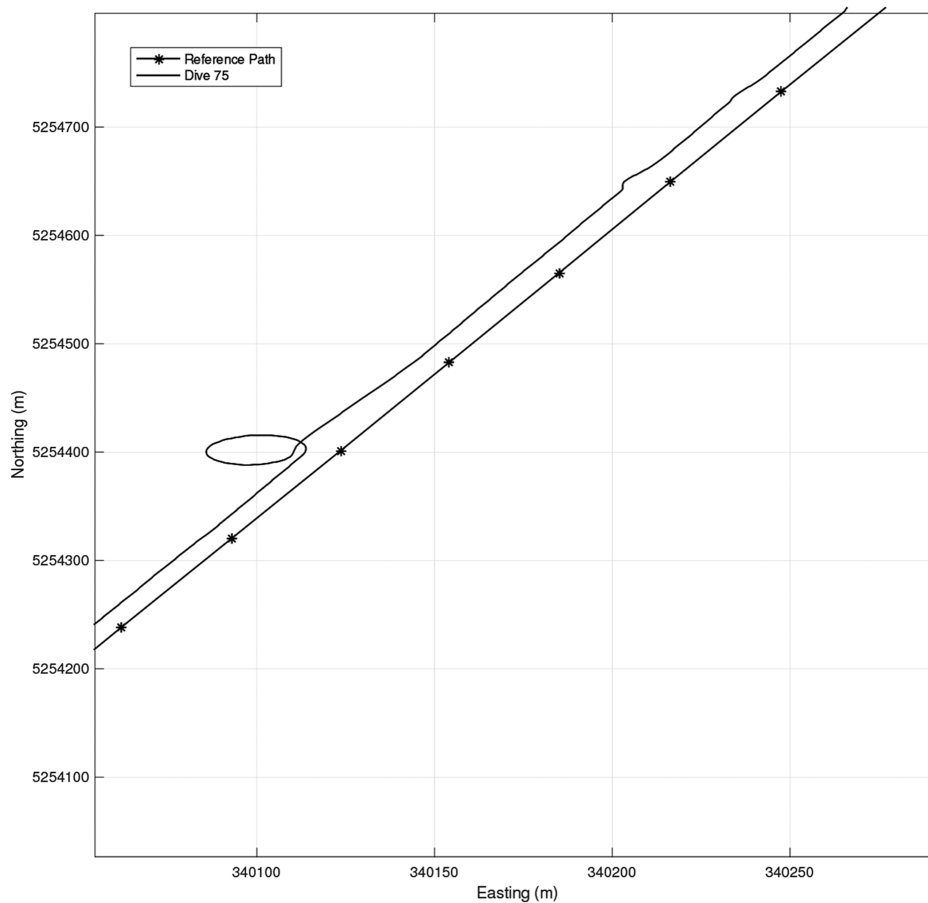

FIGURE 19 AUV waypoint circling behavior

a best match is selected. This best match provides a vector indicating the bearing of the AUV to the path, which, when combined with the known vector from the matched view to the subsequent view in the stored path, provides a corrective heading to align the AUV to the path.

This method was tested in two sets of trials in 2013 and 2014, in which two scenarios, namely a 1.2 and a 5 km path, were collected in a teaching phase to become the reference path, and multiple repeat attempts were made. Overall the system performed well, with the system able to first discover the path, obtain control of the AUV, and

produce sufficient corrective actions to move the AUV along the path to completion. Although no external position tracking was available due to the lack of a functioning acoustic tracking system, the recorded heading commands of the system indicate that the AUV repeatedly returned to a near-zero corrective state where alignment with the path was seen. Due to the dynamics of the AUV, there were consistent deviations from the path, which were subsequently corrected. As image matches were seen throughout the traversal, the AUV maintained proximity to the path within the bounds of the sonar image footprint.

The success of this system in the field is due to the robustness of the image generation and matching. The use of images that share a common rotation and scale allows poor matches to be quickly and reliably discarded. Overall, very few false positives are present, and thus successful matches provide a strong indication of location and offset. The success of image matching during all phases of the repeat traversal, including initial approaches that were orthogonal to the taught path, was due to the image-based approach taken. By aligning all images to a common orientation, utilizing feature detection not restricted to strong shadow landmarks, and matching images as a whole, strong matching performance even in the face of stark orientation changes was possible.

5.1 | Future considerations

Given the success of the trials, it is expected that development of teach-and-repeat path-following will continue. One area in which this method will find application is long-range under-ice applications, where an AUV may traverse long distances in an unknown environment and wish to return to the original site, for example an open-water lead. For this to be possible, the effects of ice-motion on navigation and the sonar data must be studied and understood.

The implementation in this work entailed several simplifications and assumptions to reduce the overall complexity. These assumptions may limit the extendability of this system and will need to be investigated moving forward. Specifically, these assumption include the AUV continuously moving along the track, a relatively flat seabed, and a discrete probability filter covering the entire space. In future work, a more realistic motion model would be required, taking input from the AUV's own motion sensors. For the navigation filter, as distances grow, a more efficient particle filter would exhibit decreased computational demands. There is also the issue of unbounded growth in the reference image set. As longer paths are attempted, a mechanism to only search the most likely region of this image set may be required.

To better understand the benefit of this system with regard to the ability to maintain path-following with bounded error, longer attempts should be made. As seen in the percent error versus distance traveled, as path lengths grow, the TR system should begin to outperform even the best commercial INS systems given that the error is bounded. Illustration of this through longer trials is critical to prove the worth of implementation.

Finally, the interaction with the AUV's control system should be addressed. This includes a proactive approach to ensure that waypoint completion does not result in completion behavior—circling—around the waypoint. This is specific to the current implementation on the Explorer AUV and it may not affect subsequent vehicles, although it does highlight the potential issues when interconnecting complex vehicle control systems.

ORCID

Peter King http://orcid.org/0000-0001-9436-0936

REFERENCES

Al-Shamma'a, A. I., Shaw, A., & Saman, S. (2004). Propagation of electromagnetic waves at MHz frequencies through seawater. *IEEE Transactions on Antennas and Propagation*, 52(11).

Alahi, A., Ortiz, R., & Vandergheynst, P. (2012). FREAK: Fast retina keypoint. In *Proceedings of Computer Vision and Pattern Recognition*.

Bay, H., Ess, A., Tuytelaars, T., Van Gool, L. (2008). SURF: Speeded-up robust features, Computer Vision and Image Understanding, 110(3), 346–359.

Chen, P., Li, Y., Su, Y., Chen, X., & Jiang, Y. (2015). Review of AUV underwater terra in matching navigation. *Journal of Navigation*, 68(06), 1155–1172.

Claus, B., & Bachmayer, R. (2015). Terrain-aided navigation for an underwater glider. *Journal of Field Robotics*, 32(7), 935–951.

Fischler, M., & Bolles, R. (1981). Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6).

Furgale, P., & Barfoot, T. (2010). Visual teach and repeat for long-range rover autonomy. *Journal of Field Robotics*, 1(27).

ISE. (2016). International submarine engineering. http://www.ise.bc.ca/explorer.html.

Itseez. (2015). Open source computer vision library. https://github.com/itseez/opencv.

Jakuba, M., Roman, C., Singh, H., Murphy, C., Kunz, C., Willis, C., … Sohn, R. (2008). Long-baseline acoustic navigation for under-ice autonomous underwater vehicle operation. *Journal of Field Robotics*, 25(11–12), 861–879.

Jenkins, A., Dutrieux, P., Jacobs, S. S., McPhail, S. D., Perrett, J. R., Webb, A. T., & White, D. (2010). Observations beneath Pine Island Glacier in West Antarctica and implications for its retreat. *Nature Geoscience*, 3(7), 468–472.

Kaminski, C., Crees, T., Ferguson, J., Forrest, A., Williams, J., Hopkin, D., & Heard, G. (2010). 12 days under ice: an historic AUV deployment in the Canadian high Arctic. In *2010 IEEE/OES Autonomous Underwater Vehicles*. pp. 1–11.

King, P., Anstey, B., & Vardy, A. (2013). Comparison of feature detection techniques for AUV navigation along a trained route. In *Proceedings of IEEE-Oceans*. San Diego.

King, P., Vardy, A., & Anstey, B. (2012). Real-time image generation and registration framework for AUV route following. In *Proceedings of IEEE-AUV*. Southampton, UK.

Kinsey, J., Eustice, R., & Whitcomb, L. (2006). Underwater vehicle navigation: recent advances and new challenges. In *IFAC Conference on Manoeuvring and Control of Marine Craft*. Lisbon, Portugal. In press.

Lewis, R., Bose, N., Lewis, S., King, P., Walker, D., Devillers, R., … Vardy, A. (2016). MERLIN—a decade of large AUV experience at Memorial University of Newfoundland. In *2016 IEEE/OES Autonomous Underwater Vehicles (AUV)*, (pp. 222–229).

Lowe, D. (2004). Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2).

Mahon, I., Williams, S. B., Pizarro, O., & Johnson-Roberson, M. (2008). Efficient view-based SLAM using visual loop closures. *IEEE Transactions on Robotics*, 24(5), 1002–1014.

Matsumoto, Y., Inaba, M., & Inoue, H. (1996). Visual navigation using view-sequenced route representation. In *Proceedings of International Conference on Robotics and Automation*, Minneapolis, MN.

Matsumoto, Y., Sakai, K., Inaba, M., & Inoue, H. (2000). View-based approach to robot navigation. In *Proceedings of IEEE Conference on Intelligent Robots and Systems*.

McEwen, R., & Thomas, H. (2003). Performance of an AUV navigation system at Arctic latitudes. In *OCEANS 2003. Proceedings*, Vol. 2, pp. 642–653.

Meduna, D., Rock, S., & McEwan, R. (2008). Low-cost terrain relative navigation for long-range AUVs. In *OCEANS 2008*.

Nguyen, T., Mann, G., Gosine, R., & Vardy, A. (2016). Appearance-based visual-teach-and-repeat navigation technique for micro aerial vehicle. *Journal of Intelligent Robotic Systems*, 84, 217–240.

Paull, L., Saeedi, S., Seto, M., & Li, H. (2014). AUV navigation and localization: A review. *IEEE Journal of Oceanic Engineering*, 39(1).

Pinto, M., Ferreira, B., Matos, A., & Cruz, N. (2009). Using side scan sonar to relative navigation. In *OCEANS 2009*, (pp. 1–9).

Rock, S., Hobson, B., & Houts, S. (2014). Return-to-site of an AUV using terrain-relative navigation: Field trials. In *Proceedings of IEEE-AUV*. Oxford, MS.

Tena Ruiz, I., de Raucourt, S., Petillot, Y., & Lane, D. M. (2004). Concurrent mapping and localization using sidescan sonar. *IEEE Journal of Oceanic Engineering*, 29(2), 442–456.

Thrun, S., Burgard, W., & Fox, D. (2005). *Probabilistic robotics*. Cambridge, MA: MIT Press.

Vandrish, P., Vardy, A., & King, P. (2012). Towards AUV route following using qualitative navigation. In *Proceedings of Computer and Robot Vision*. Toronto, ON.

Vandrish, P., Vardy, A., Walker, D., & Dobre, O. (2011). Side-scan sonar image registration for AUV navigation. In *Proceedings of IEEE-Oceans*.

Zhang, A., & Kleeman, L. (2009). Robust appearance based visual route following for navigation in large-scale outdoor environments. *International Journal of Robotics Research*, 28(3).