SPECIAL ISSUE

AUV-enabled adaptive underwater surveying for optimal data collection

David P. Williams

Abstract A new adaptive strategy for performing data collection with a sonar-equipped autonomous underwater vehicle (AUV) is proposed. The approach is general in the sense that it is applicable to a wide range of underwater tasks that rely on subsequent processing of side-looking sonar imagery. By intelligently allocating resources and immediately reacting to the data collected in-mission, the proposed approach simultaneously maximizes the information content in the data and decreases overall survey time. These improvements are achieved by adapting the AUV route to prevent portions of the mission area from being either characterized by poor image quality or obscured by shadows caused by sand ripples. The peak correlation of consecutive sonar returns is used as a measure for image quality. To detect the presence of and estimate the orientation of sand ripples, a new innovative algorithm is developed. The components of the overall data-driven path-planning algorithm are purposely constructed to permit fast real-time execution with only minimal AUV onboard processing capabilities. Experimental results based on real sonar data collected at sea are used to demonstrate the promise of the proposed approach.

Keywords Autonomous underwater vehicles · Sonar data · Image quality · Ripple detection · Adaptive route planning

1 Introduction

The high-resolution imaging of underwater environments afforded by sonar has proven useful in a wide range of applications, including habitat mapping [1,2], seabed classification [3–5], mine detection [6,7], port protection [8], archaeology [9], and pipeline monitoring [10]. Thanks to breakthroughs in marine robot technology, the sonar data used to address these diverse tasks is invariably collected by an autonomous underwater vehicle (AUV).

Typically, however, the data collection process itself does not receive much attention. Instead, the aforementioned applications that exploit and rely on sonar data implicitly assume that the collected data is of sufficient quality to complete the desired task.

Unfortunately, data collected at sea with an AUV is rarely flawless [11]. For example, excessive undesired vehicle motion caused by currents can result in defocused imaging, navigation errors can lead to gaps in data coverage, and shadows caused by bathymetric features can obscure the seabed.

But recent advances in the onboard processing capabilities of AUVs now permit increasingly sophisticated ways to perform the data collection task. In particular, these advances make the in situ adaptation of AUV data-collection survey routes possible. This capability has the potential to ensure that the data collected at sea by an AUV meets the standards necessary for various subsequent objectives.

The promise of guaranteed data quality is intriguing because collecting data at sea is an extremely expensive enterprise that requires substantial assets as well as significant manpower (to handle everything from ship operation to AUV deployment and recovery). Moreover, ship transit and travel costs further increase the cost of performing sonar data collection.

With such costs, repeating an experiment at sea because some data is of poor quality is oftentimes simply not feasible. Instead, flawed data is usually quietly discarded and ignored. In certain applications, however, it is absolutely imperative that high-quality data is collected everywhere in the area of interest. In such cases, collecting additional data—and

D. P. Williams (✉)
NATO Undersea Research Centre, Viale San Bartolomeo 400,
19126 La Spezia (SP), Italy
e-mail: williams@nurc.nato.int

incurring the substantial costs associated with doing so—is unavoidable.

The development of intelligent methods to control AUV data collection to ensure high-quality sonar data would therefore be a considerable boon for a wide range of applications. Not only would these advances provide the opportunity to save resources, reduce manpower requirements, and reduce mission times at sea, but the various tasks that subsequently process the data would also benefit from the data being of increased quality.

In this work, we address this challenge. Specifically, we develop a strategy to collect optimal *side-looking sonar* data by adapting, in situ, the AUV survey route based on the environmental characteristics observed in the data itself, such that the data quality is sufficiently high to allow the completion of any desired task. By developing a general strategy rather than restricting ourselves to a single application, the algorithms presented here will be applicable to a wide range of tasks and will be useful for all researchers in the field of marine robotics who deal with sonar data.

Previous work related to adaptive underwater path planning has not focused on ensuring the *quality* of the data being collected, as we do here. Instead, the main body of work has been interested in finding selected regions of particular interest to survey or routes that satisfy certain objectives, rather than performing an exhaustive search of an area. For example, adaptive surveys have been used to locate oil spills [12] and reefs [13] by surveying only a subset of an area after the event of interest has been detected. Other work has adopted adaptive strategies to minimize errors due to temporal evolution of the ocean during the survey [14], or to reach a mission-specified destination while minimizing energy costs [15]. The work involving dense, exhaustive surveys of an area with an AUV has tended to focus on navigation correction and localization [16] or the fusion of multiple images into a mosaic of the seabed [17] for subsequent object detection tasks.

The remainder of this paper is organized as follows. In Sect. 2, background about sonar data collection with an AUV is discussed. Section 3 proposes a way to quantify the information content of a sonar image, based on image quality and shadows cast by sand ripples. A new, fast algorithm for the detection of sand ripples in sonar imagery is detailed in Sect. 4. The proposed strategy for performing adaptive data collection is outlined in Sect. 5. Experimental results based on real sonar data are shown in Sect. 6. Concluding remarks and directions for future work are provided in Sect. 7.

2 Side-looking sonar data collection

AUVs equipped with side-looking sonar—simple side-scan sonar or synthetic aperture sonar (SAS) [18]—are increasingly used to acquire high-resolution images of underwater environments.

Usually, such AUVs are equipped with two sonars, one on the port side and one on the starboard side. These sensors image in a direction orthogonal to the direction in which the AUV travels. Because of the geometry of the problem, a "dead zone"—from directly below the AUV out to a certain range on either side—between the two sonar swaths will lack sonar coverage [11].

The standard AUV survey plan used in practice [19–21] is a series of equidistant parallel tracks—often referred to as a "lawnmower pattern"—such that the sonar swaths of consecutive tracks interleave, resulting in sonar coverage for the entire area of interest. After the area has been surveyed in this primary orientation, a second survey is often performed in an orientation orthogonal to the first survey. This "cross-hatching" ensures that each seabed location is observed at least twice [22].

In benign environments, having two complementary views of the seabed is overkill for object detection purposes (having two views can still be desirable for classification objectives, in which the difficulty of identifying an object can be a strong function of aspect [23]). Obtaining orthogonal views becomes a necessity only in *complex* environments, so that the shadows cast by bathymetric features or objects will obscure different areas.

However, the inflexibility of this rigid pre-planned approach introduces major flaws in the standard sonar data collection procedure. Specifically, the inability to adapt the survey based on the data that is collected in situ during the mission—and the coverage actually achieved by the AUV—introduces severe inefficiencies into the process.

To see this, it must be noted that pre-planned survey routes are designed under the assumption that successful sonar coverage will be achieved up to a certain maximum range. If this maximum range is underestimated, the survey mission will take longer to complete and a considerable portion of the data collected will overlap, covering the same area unnecessarily. However, if the maximum range for which sonar coverage is expected is overestimated, the result will be gaps in data coverage. In order to successfully cover the entire mission area in this scenario, the AUV would need to be re-deployed to execute additional survey tracks.

A second major flaw with the pre-planned surveying approach is that it prevents the efficient allocation of resources. Surveying the entire area of interest from two different orientations virtually guarantees that excessive attention will be devoted to some areas and insufficient resources will be applied to others. For areas characterized by flat, benign seabed, a single view is sufficient for acquiring all of the relevant information present. Therefore, surveying such an area twice (i.e., with the cross-hatching) doubles the mission time while providing very little new or additional

information. It would be better to instead execute additional surveys only on areas that warrant them, such as complex seabed containing anomalous features of interest.

These aforementioned inefficiencies result in longer mission times, which then translate directly into higher costs. Moreover, in time-sensitive military applications, protracted surveys can also delay subsequent operations and adversely impact future aspects of the mission.

The solution for overcoming these drawbacks of the standard surveying approach is to allow the AUV to adapt its route, in situ, based on the characteristics observed in the data being collected. In this work, we propose such a strategy based on the *information content* of the sonar imagery collected.

3 Information content of sonar imagery

The proposed survey strategy is based on adapting the route of an AUV according to the data that it collects in situ so that it can be exploited in the remainder of the mission. More specifically, the procedure will select survey routes to maximize the information contained in the data.

To accomplish this objective while also maintaining general applicability, some universal measure of information— that is relevant regardless of the eventual application for which the data is collected—must be defined and extracted from the sonar data.

We argue that the pertinent question entailed by this challenge is: Does the current data allow accurate, confident assessment of a given location of seabed? If the answer is affirmative, the data at the considered location can be deemed to contain information. Each pixel in an image can thus be classified in a binary manner based on this query.

More concretely, in this work we define the information content of a sonar image in terms of two quantities: image quality and shadows cast by sand ripples. Although extremely simple, this formulation will prove to be a powerful way to rapidly assess the information content in the sonar data. And importantly, both quantities can be evaluated quickly in near real-time onboard an AUV with limited processing capabilities. Other factors that can have a profound impact on the utility of the sonar data, such as a steeply sloped seabed, are not addressed in this work.

When an area of seabed is characterized by poor image quality or obscured by shadows, no information is provided by the data because the true nature of the location cannot be deduced. That is, an accurate assessment of what is present on the seabed is not possible.

To illustrate this point, an example SAS image is shown in Fig. 1. At long range (beyond approximately 112 m), the image is blurry because unwanted vehicle motion exceeded the allowable limits for successful imaging; at short range,
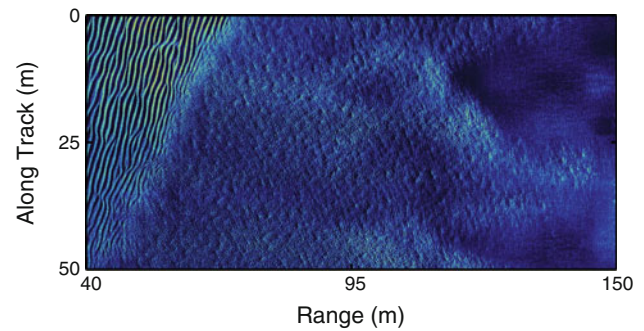
Fig. 1 Example SAS image exhibiting shadows due to sand ripples at short range and poor image quality at long range (beyond 112 m)

a significant portion of the seabed is obscured by shadows cast by a field of sand ripples.

3.1 Image quality

In our formulation, the first quantity on which the information content of sonar data depends is image quality. A sonar image with good image quality is one in which objects and seabed features appear more or less as they would when observed under the ideal operating conditions of the sonar system.

It is commonly assumed that the full extent of all sonar images are of good quality, but that is often not the case with real sonar data collected at sea [11]. In particular, image quality often degrades significantly at long range when the effects of multipath manifest [24]. Additionally, with SAS, the motion correction requirements for successful image formation are more stringent at long range [25]. As a result, the image quality of data at the longest ranges of the sonar system is often inadequate.

If it is assumed that good quality sonar data will be achieved up to some maximum range that is longer than is actually achieved, a substantial portion of the data that is collected will contain no useable information. This point highlights the importance of allowing the AUV survey route to adapt based on the image quality of the data collected.

The issue as to how to quantify the rather subjective concept of image quality must be addressed next.

Synthetic aperture sonar (SAS) provides an order-of-magnitude improvement in resolution over simple (real aperture) side-scan sonar data [18]. Because the higher-resolution data permits much more accurate assessment of underwater environments, challenging tasks are now often undertaken with SAS instead of side-scan sonar.

However, the ability to successfully form SAS images— and in turn, the achievable area coverage—depends on the motion of the sonar-equipped platform and on the environmental characteristics of the area (such as the seabed-type and the water properties) [26]. Excessive platform motion will limit the range for which successful image formation is

possible. Similarly, if the seabed is soft (e.g., muddy), less sonar energy will return to the receivers, also limiting the range to which a SAS image can be formed successfully.

SAS works by coherently summing received sonar signals, or ping-returns, of overlapping elements in the array. The (cross-)correlation between two received time-series signals, $s_i(t)$ and $s_j(t)$, is simply

$$\kappa(\tau) = \langle s_i(t)s_j^*(t+\tau) \rangle, \tag{1}$$

where τ is a time-shift. The normalized peak correlation between successive pings, termed the *coherence* [27], is then simply

$$\rho = \max_\tau \left| \frac{\kappa_{ij}(\tau)}{\sqrt{\kappa_{ii}(0)\kappa_{jj}(0)}} \right|. \tag{2}$$

This peak correlation between successive pings at a given range (i.e., distance from the sonar) provides a measure of the success of the SAS processing, as this quantity is directly proportional to the signal-to-noise ratio (SNR) [28],

$$\text{SNR} = \frac{\rho}{1-\rho}, \tag{3}$$

and in turn, the image quality. Granted, a high coherence does not always ensure good image quality, but in general, it is a reliable indicator that has been employed with success [27].

Although full SAS processing is a computationally intensive procedure, the coherence can be easily calculated onboard an AUV in a streaming manner as data is collected. Therefore, this work uses the coherence (or "ping-to-ping correlation"), ρ, calculated also as a function of range, as a measure of image quality.

Based on our extensive experience visually assessing the quality of sonar images, we deem an image at ranges where ρ is below the threshold $\tau_\rho = 2/3$ to be of poor image quality containing no information. (A threshold of $\tau_\rho = 3/4$ is used in [27].) To more rigorously justify this choice, we analyzed the performance of a detection algorithm [29] on a large data set of sonar images containing 480 man-made targets. In theory, the detection algorithm should detect all of these targets if the image quality is sufficient. The detection performance that resulted, as a function of the minimum coherence value of the targets, is shown in Fig. 2. From the figure, it can be seen that approximately 95% of the targets are successfully detected if the coherence at the target location is at least 2/3, which suggests that the threshold $\tau_\rho = 2/3$ is reasonable.

Thus, with a threshold selected, the (binary) image quality as a function of range can be predicted, and the survey route adapted accordingly, without forming the actual SAS image.

3.2 Sand-ripple shadows

Because of the unique geometry defined by the AUV's sonar and the seafloor, a general object or structure that rises above
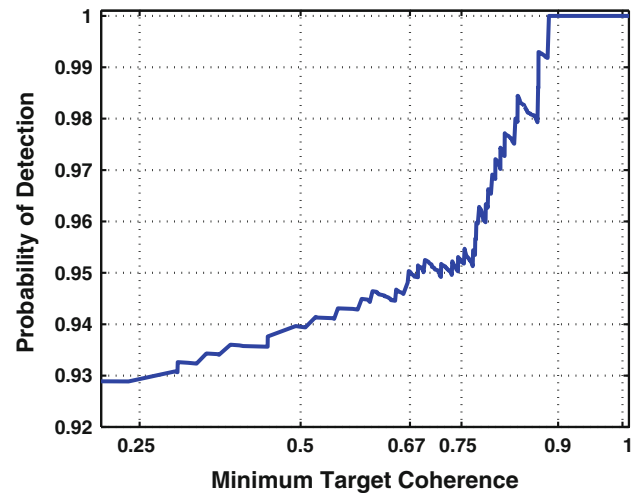
Fig. 2 For one sonar data set, the probability of detecting a man-made target when considering only those targets with a coherence above some minimum threshold

the seafloor will necessarily cast a shadow (except in pathological cases that can be ignored here).

Whereas the image quality is mainly a factor when processing side-scan sonar into SAS, the second quantity on which we defined information content, shadows, is relevant even for side-scan sonar data.

In this work, we focus specifically on the shadows cast by sand ripples because such shadows are largely preventable by surveying at particular orientations. An object such as a boulder will cast a shadow—albeit over different areas—regardless of the survey orientation. Moreover, sand ripples are usually large-scale phenomena that can impact a substantial portion of a set of collected data, a situation for which the execution of additional survey tracks would be justified and worthwhile.

Seabed characterized by sand ripples exhibits a strong orientation dependence; when the sand ripple "waves" propagate in the along-track direction of the vehicle (which is assumed to be orthogonal to the look-direction of the sonar), the proportion of seabed that is obscured by shadows is minimized. An example of this phenomenon in real SAS data is shown in Fig. 3 [unless otherwise noted, for the remainder of this paper, we define the orientation of a sand ripple as the orientation at which a given crest or trough of the ripple field is directed (i.e., orthogonal to the "wave" propagation direction), as measured from the across-track direction of a sonar survey].

As can be noted in Fig. 3, the amount of seabed obscured by shadows cast by ripples is strongly influenced by the relative survey orientation. This anecdotal evidence suggests that, according to our definition of information content, data will be much more informative and valuable when areas characterized by sand ripples are surveyed at certain orientations.
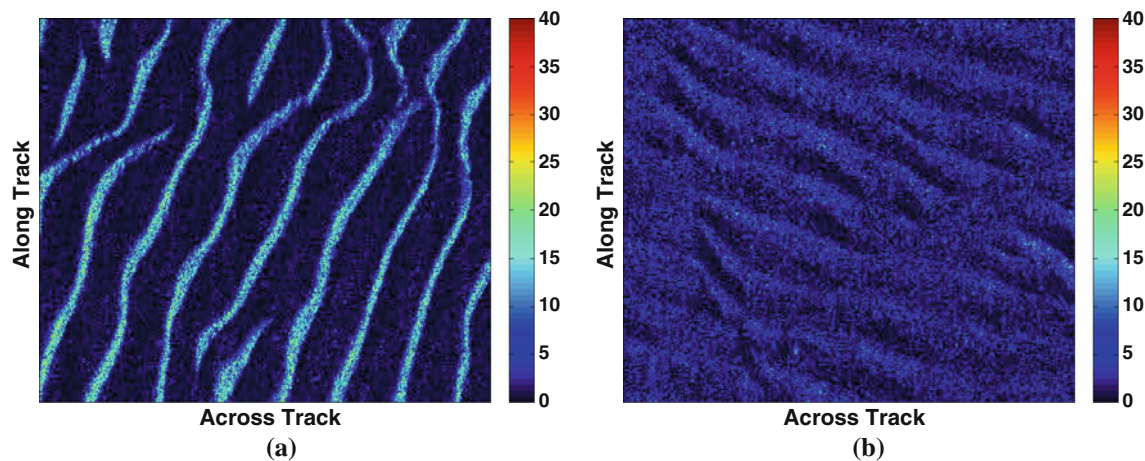
Fig. 3 Sonar imagery of the same area of seabed characterized by ripples when the along-track direction of the vehicle is (**a**) nearly parallel to the ripple orientation ($\theta \approx 65°$), and (**b**) nearly perpendicular to the ripple orientation ($\theta \approx -25°$)

To better illustrate this fact more rigorously, we conducted experiments using simulated sonar data generated by the method in [30]. Specifically, patches of seabed characterized by sand ripples with different amplitudes and different wavelengths were generated. We then simulated the sonar data that would result if the data was collected at various survey orientations and also at different ranges, when the sonar was at an altitude of 13 m. For each case, the fraction of the sonar data that would be obscured by shadows caused by the ripples was calculated. The mean results of these experiments, averaged over 100 random trials (i.e., seabed patch realizations), are shown in Fig. 4.

As can be seen from the figure, the proportion of the seabed obscured by shadows caused by sand ripples is minimized when the along-track direction of the vehicle is nearly perpendicular to the ripple orientation (i.e., when the sonar look-direction is *along* a ripple wave's crest or trough).

These experimental results provide further evidence that basing the information content of sonar data on shadows caused by ripples is justified. At certain survey orientations, a substantial portion of the resulting sonar data may be obscured by shadows, and therefore not useful for subsequent tasks.

It should be noted that any data processing required to determine how best to adapt the survey route must also be executed in near real-time. With this caveat in mind, we developed a new fast algorithm that both detects the presence of and estimates the orientation of sand ripples. The ripple detection algorithm in [31], relying on numerous wavelet transforms, is too computationally demanding and slow for this purpose and cannot be executed in a streaming manner the way our proposed approach can. The algorithm proposed in [32] is more accurately a ripple *suppression* algorithm that artificially removes the ripple shadows from the image. Since this neither adds information to the image nor establishes the orientation of the ripples, the approach is not appropriate for the adaptive surveying approach.

Since ripples are relatively large structures (compared to the resolution of the imagery), even lower-resolution side-scan data is sufficient for ripple detection and estimation. Therefore, the computationally expensive SAS image formation can be avoided in this context as well. Although SAS could achieve higher detection precision, side-scan data is adequate for detecting ripples and estimating their orientation for purposes of survey route adaptation.

Because of its import on information content and the adaptive survey strategy, our new ripple detection algorithm is described in detail in the next section.
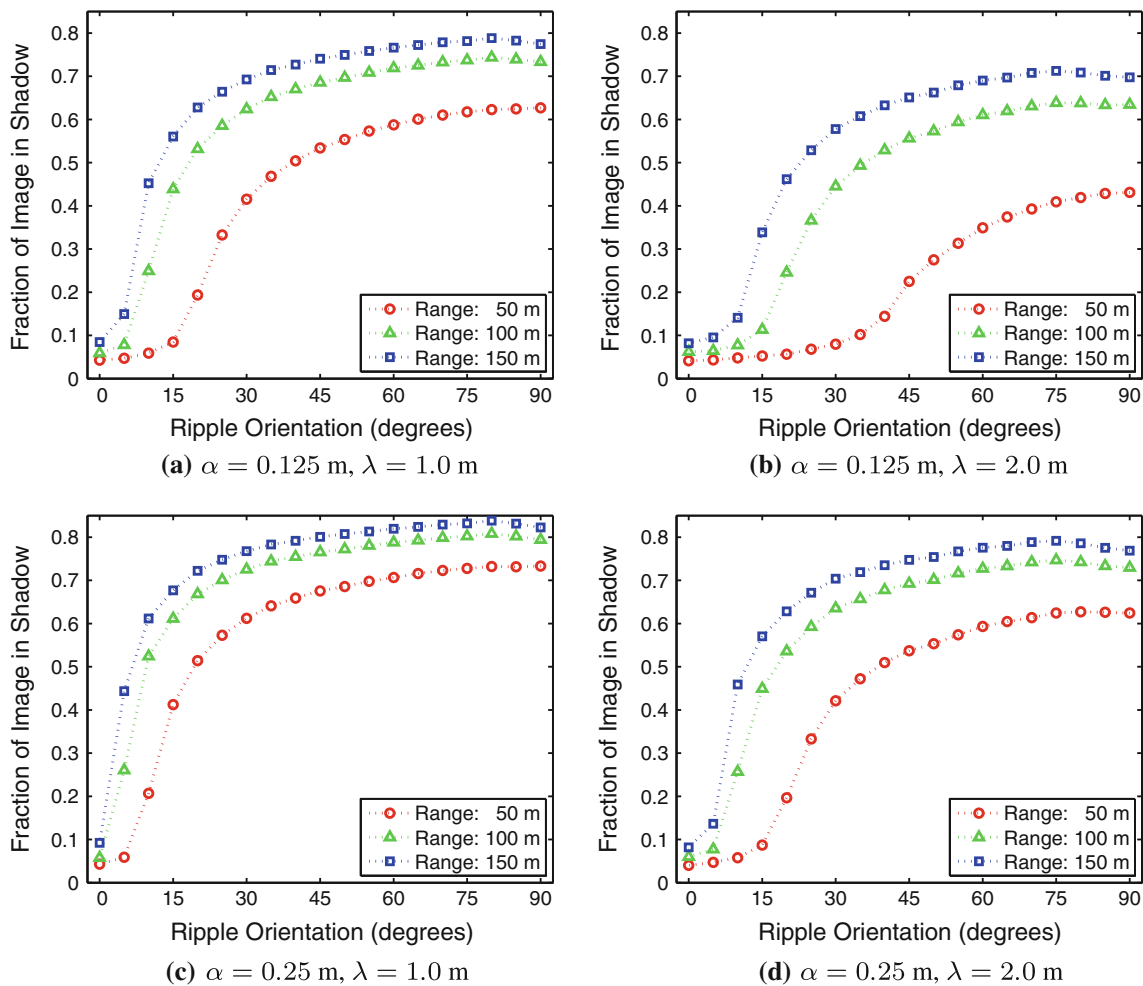
4 Ripple detection algorithm

4.1 Overview

The proposed ripple detection algorithm is composed of several components. Before becoming mired in the necessary details, we first briefly provide an overview of the motivations and general philosophy used to create the algorithm.

The overall objective is to create a robust algorithm that will reliably detect sand ripples. However, the algorithm must also be fast to permit real-time streaming detection onboard an AUV equipped with limited processing capabilities. Since no human intervention is allowed, the algorithm must be flexible enough to sense and adapt to changing environmental conditions from the data collected in situ.

To satisfy these requirements, extensive domain-specific knowledge we possess about the problem is exploited. Emphasis is also placed on tailoring the algorithm, where possible, to the fundamental underlying physics and geometry of the application. The result is a hybrid of rigor and flexibility.

(a) $\alpha = 0.125$ m, $\lambda = 1.0$ m

(b) $\alpha = 0.125$ m, $\lambda = 2.0$ m

(c) $\alpha = 0.25$ m, $\lambda = 1.0$ m

(d) $\alpha = 0.25$ m, $\lambda = 2.0$ m

Fig. 4 For an image characterized by sand ripples, fraction of image that would be obscured by shadows, as a function of range and ripple orientation, for different ripple amplitudes, α, and wavelengths, λ.

The ripple orientation is the angle at which one ripple wave's crest or trough is directed, with 0° in the across-track direction

The following sections will describe each step of the detection algorithm that resulted from these considerations. Because of the visual nature of the application, the algorithm description will be supplemented by figures illustrating the effects of each step on a sonar image containing sand ripples, shown in Fig. 5a.

However, it should be noted that the algorithm does not require a fully formed image as the input (i.e., streaming detection using only a few rows of an "image" is possible). But for illustrative purposes, the stages of the algorithm will be shown applied to a full image.

4.2 Integral image

An integral image [33] is an image representation that allows for very fast computation of rectangular, Haar-like features at any scale or location in constant time. In subsequent stages of the algorithm, such rectangular features will be invaluable for assessing certain distinguishing characteristics of ripples.

The use of an integral-image representation is also a key that makes *streaming* detection possible.
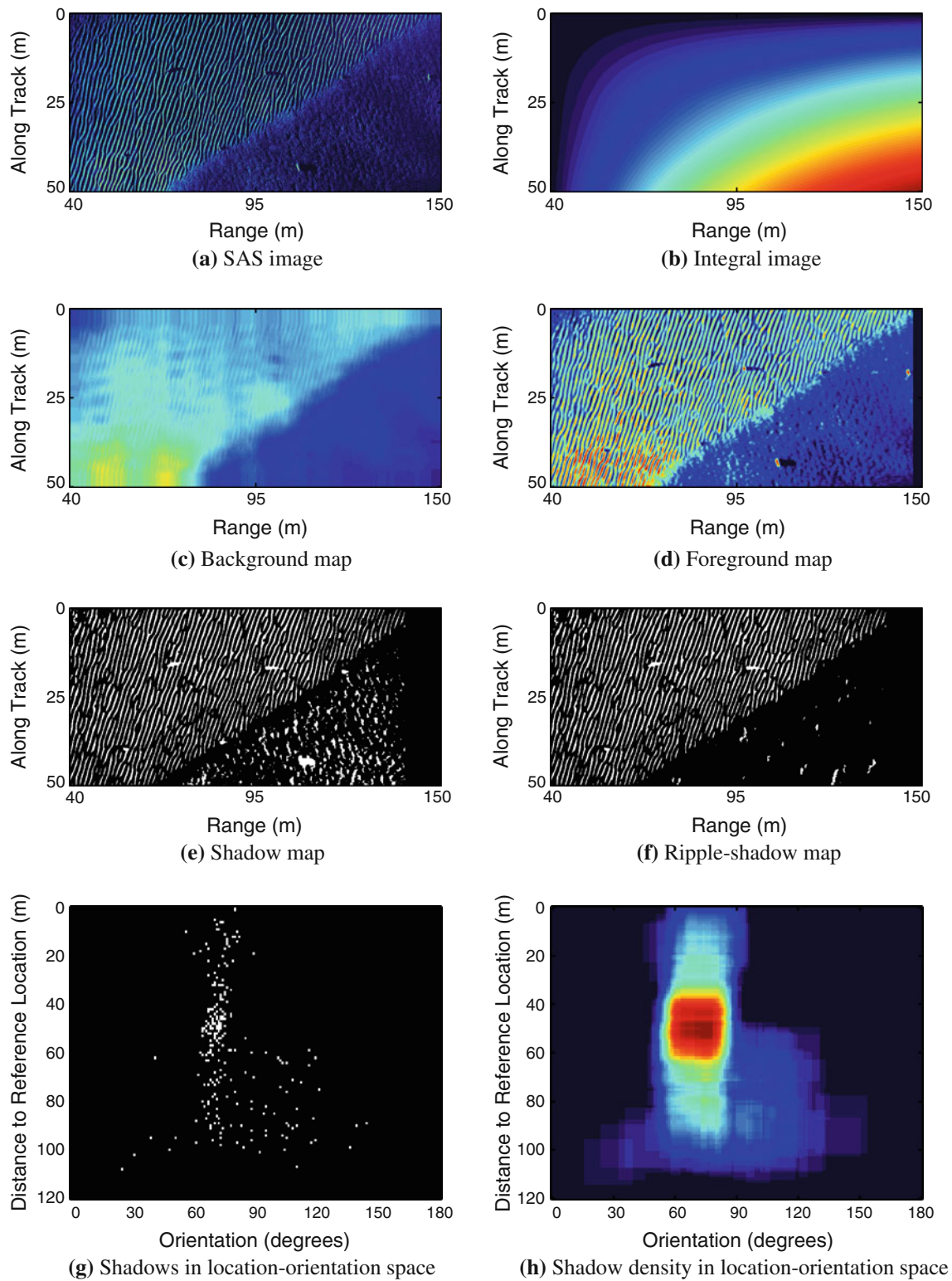
Thus, rather than operating on the pixel-based system of the sonar image, we immediately transform to an integral-image system. (In fact, one need not ever form the full sonar image onboard the AUV since the integral-image representation contains equivalent information.) Starting from an original sonar image, A, the corresponding integral image, I, is constructed as follows.

The value at a location (x, y) in the integral image corresponds to the sum of the pixels above and to the left of (x, y), inclusive, in the original image, A. That is,

$$I(x, y) = \sum_{x' \leq x, y' \leq y} A(x', y').$$ (4)

The integral image is quickly generated using the recursive relation

$$I(x, y) = I(x - 1, y) + z(x, y),$$ (5)

(a) SAS image

(b) Integral image

(c) Background map

(d) Foreground map

(e) Shadow map

(f) Ripple-shadow map

(g) Shadows in location-orientation space

(h) Shadow density in location-orientation space

Fig. 5 Ripple detection algorithm. The original sonar image (**a**) is converted into its equivalent integral-image representation (**b**), which is then used to estimate a background map (**c**) and a foreground map (**d**). The result of the background map and foreground map comparison produces a shadow map (**e**). A location and orientation feature are then extracted for each discrete shadow (**g**), from which density estimation is performed (**h**) to determine the presence and orientation of sand ripple shadows, resulting in a final ripple-shadow map (**f**). The principal orientation of the ripples in this image was estimated to be $\theta_r = 73°$, with the full set of ripple orientations spanning $\Theta_r = [57°, 86°]$

where $z(x, y)$ is the cumulative sum of pixels in a row of the original image,

$$z(x, y) = z(x, y - 1) + A(x, y). \tag{6}$$

The integral-image approach is particularly well-suited for tasks with streaming sonar data because the construction of an integral image exactly mimics the manner in which the data is collected, namely in a row-wise fashion. For a given along-track position of the AUV, the sonar returns corresponding to a set of ranges—which comprise the columns of the sonar "row-image"—are used in the calculation.

The integral-image representation corresponding to the sonar image in Fig. 5a is shown in Fig. 5b.

Once the integral image is constructed, the ripple detection algorithm never again requires use of the original sonar image. Subsequent stages of the detection algorithm instead repeatedly exploit the integral image.

4.3 Background estimation

The first use of the integral image, I, is in the estimation of the sonar-image background map, B. The purpose of the background map is to establish the reverberation level of the seabed in order to subsequently determine locations of shadows in the image.

The reverberation level is a strong function of the seabed composition, so assuming a pre-defined threshold to determine what constitutes background levels is not reliable. For instance, the reverberation level of a soft muddy seabed will be lower than that of a seabed of hard-packed sand.

In fact, natural seabed variations can cause the reverberation level to vary substantially at a given site or even within the same image. In Fig. 5a, it can be seen that significant variation exists within the region of flat seabed; the presence of sand ripples complicates matters even further.

A data-driven approach that defines the background level according to some global average over the sonar image would therefore fail catastrophically on such an image. Such a method would also require possession of a full image, which would preclude the possibility of streaming, real-time execution. Instead, we argue that the background estimation should be performed in the image *locally*.

Thus, in this work, a robust estimation of the background is obtained by using the local characteristics of the seabed in the measured data. The inherent flexibility afforded by tailoring the estimation to the data itself improves the method's robustness for successfully dealing with different environments. The adaptive estimation also eliminates the possibility of training data mismatch [34,35], since no training data is used.

The background estimation is calculated using a split window that we call an "Austrian flag" template since it is composed of three equally-sized horizontal bands of which the top and bottom bands (red in the flag) contribute to the background value, while the middle band (white in the flag) does not. Each band, or rectangle, is of the same size, $b_x = 2.5$ m in the along-track direction and $b_y = 5.0$ m in the range direction. These sizes are chosen so that a sufficiently large area of seabed is used in the background calculation to produce relatively smooth estimates.

The background score at location (x, y), $B(x, y)$, is then the mean pixel value in the two rectangles with the "flag" centered around (x, y). Thanks to the integral-image representation, the calculation of the two rectangles that contribute to the background value at a given location can be computed quickly, with a total of only eight array references from the integral image.

Specifically, the background score at (x, y) is calculated as

$$
\begin{aligned}
B(x, y) = a_b^{-1} \times \Bigg[& I\left(x - \frac{3\delta x}{2}, y - \frac{\delta y}{2}\right) \\
& - I\left(x - \frac{3\delta x}{2}, y + \frac{\delta y}{2}\right) \\
& - I\left(x - \frac{\delta x}{2}, y - \frac{\delta y}{2}\right) + I\left(x - \frac{\delta x}{2}, y + \frac{\delta y}{2}\right) \\
& + I\left(x + \frac{\delta x}{2}, y - \frac{\delta y}{2}\right) - I\left(x + \frac{\delta x}{2}, y + \frac{\delta y}{2}\right) \\
& - I\left(x + \frac{3\delta x}{2}, y - \frac{\delta y}{2}\right) + I\left(x + \frac{3\delta x}{2}, y + \frac{\delta y}{2}\right) \Bigg],
\end{aligned} \tag{7}
$$

where δx and δy are the numbers of pixels that correspond to b_x and b_y, respectively, and a_b is the total number of pixels involved in the sums, so that the result is the mean pixel value from the two rectangular areas.

The region immediately surrounding a location (i.e., the "middle band") is not used in the background calculation because those pixels are more representative of the "foreground."

The resulting background map from using the integral image in Fig. 5b is shown in Fig. 5c.

4.4 Foreground estimation

Next, the integral image is used again to construct a foreground map. The purpose of the foreground map is to subsequently aid in locating shadows in the image.

The foreground score at (x, y), $F(x, y)$, is taken to be the mean pixel value over a rectangle centered around (x, y). The size of the rectangle used in the calculation is $f_x = 0.5$ m in the along-track direction and $f_y = 1.5$ m in the range-direction. These sizes are chosen so shadows (cast by sand ripples) that are large enough to obscure objects of potential interest—e.g., mines—would be detected by the algorithm.

The foreground score is again calculated quickly thanks to the integral-image representation,

$$F(x, y) = a_f^{-1} \times \left[I\left(x - \frac{\delta x}{2}, y\right) \right.$$
$$- I\left(x - \frac{\delta x}{2}, y + \delta y\right)$$
$$\left. - I\left(x + \frac{\delta x}{2}, y\right) + I\left(x + \frac{\delta x}{2}, y + \delta y\right) \right], \quad (8)$$

where δx and δy are the numbers of pixels that correspond to f_x and f_y, respectively, and a_f is the total number of pixels involved in the sum, so that the result is the mean pixel value from the rectangular area.

The resulting foreground map from using the integral image in Fig. 5b is shown in Fig. 5d.

4.5 Shadow determination

After a background map B and a foreground map F are constructed, a shadow map S indicating the regions of the image that contain shadows is determined.

The substantial variations that exist in sonar images and at different sites mean that defining shadows simply as locations with low pixel values would be unreliable. Instead, any pixel for which the foreground map value is sufficiently lower than the corresponding background map value is declared to be a shadow that will receive further investigation.

Specifically, if $F(x, y) < \beta B(x, y)$, then the pixel (x, y) is deemed to be part of a shadow. The scaling factor $\beta = 0.5$ is used in this work to make the requirement for shadows more stringent. (Essentially, this *defines* that a shadow exists when the pixel value is less than half of the average value of the surrounding background.) This use of contextual information makes the shadow detection process more robust in order to succeed across different environments.

The resulting shadow map from using the background and foreground maps in Fig. 5c and Fig. 5d, respectively, is shown in Fig. 5e.

4.6 Ripple-shadow determination

4.6.1 Overview

The principled ripple detection algorithm developed here is tailored both to the fundamental characteristics that *define* sand ripples and the underlying physics that creates them.

Sand ripples are formed when currents or waves flow in a particular direction [36–38]. Because of this underlying mechanism by which they are created, ripples tend to exist in large *fields*, rather than as individual instantiations. But a ripple field, by definition, must contain a certain high density of ripples.

Although the ripples will be characterized by a dominant orientation (dictated by the direction of flow), considerable variability will also exist. That is, ripples will not manifest at a single fixed orientation. Rather, they will span a range of orientations about some dominant one. Moreover, the amplitudes and periods of the ripples can also vary. Therefore, in order to be able to detect such variability, a method that is inherently flexible is required.

Additionally, the relative geometry between the sonar and the mound of sand that defines the ripple means that the shadows that are cast by the ripples will be elongated along the direction of the ripple crest (or trough) [39].

These three fundamental physical characteristics of ripples, which can be observed in Fig. 5a, will be exploited in the ensuing stages of the ripple-detection algorithm. Specifically, the proposed ripple detection algorithm searches for a high density of elongated shadows that are oriented similarly.

To perform this search efficiently, each discrete shadow region from the shadow map (cf. Fig. 5e) is mapped into a new feature space in which it is easy to detect such characteristics. A novel, unconventional use of the integral-image formulation for purposes of density estimation is then employed.

4.6.2 Feature extraction

To begin, the centroid and orientation is calculated for each discrete shadow (i.e., each connected region, or "blob") in the shadow map. The orientation (in degrees), ϕ, of a shadow is obtained by fitting an ellipse to the shadow area, and then correcting for unequal pixel sizes. The centroid is converted to a distance feature (in meters), χ, that is simply the distance of the centroid to a fixed reference location in the mission (for the case of a single image here, the "first" pixel of the image: the pixel at the nearest range at the initial along-track position).

This feature extraction process effectively maps each shadow to a specific location, (χ, ϕ), in a new feature space in $\mathbb{R}^2$ (in order to treat this space as an image later, these features are then discretized to $1°$ in the orientation-dimension and 1 m in the location-dimension).

Before proceeding though, certain shadows are removed from further consideration. If $\phi < d_\phi/2$ or $\phi > 180° - d_\phi/2$, an indication that the shadow is nearly horizontal, the shadow is ignored. (In this work, $d_\phi = 30°$.) This choice is made because it is known a priori, from the geometry of the sonar and seabed, that the majority of shadows at those prohibited near-horizontal orientations are likely caused by objects like rocks (as opposed to ripples). Additionally, this choice eliminates any inconsistency that would arise from not respecting the "wrap-around" nature of orientations in the upcoming integral-image density estimation.

The result of this feature extraction process that transforms the shadows in sonar-image space in Fig. 5e to location-orientation space is shown in Fig. 5g.

4.6.3 Density estimation

By treating this new feature space as a two-dimensional "image," A', the corresponding integral image, I', can be computed readily. The integral-image formulation can then be exploited to quickly perform density estimation in this space. Specifically, the mean number of discrete shadows in a rectangular area about a given location is used as the shadow density estimate.

The size of the rectangle used in the calculation is $d_\chi = 25$ m in the location-dimension and $d_\phi = 30°$ in the orientation-dimension. This size is chosen because it respects the physical proximity and orientation variability of the underlying mechanisms that help define ripple fields.

The density estimate at $D(\chi, \phi)$ is taken to be the mean pixel value (though irrespective of orientation) over the rectangle centered around (χ, ϕ), calculated quickly as

$$D(\chi, \phi) = d_\chi^{-1} \times \left[I'\left(\chi - \frac{d_\chi}{2}, \phi - \frac{d_\phi}{2}\right) \right.$$
$$- I'\left(\chi - \frac{d_\chi}{2}, \phi + \frac{d_\phi}{2}\right)$$
$$\left. - I'\left(\chi + \frac{d_\chi}{2}, \phi - \frac{d_\phi}{2}\right) + I'\left(\chi + \frac{d_\chi}{2}, \phi + \frac{d_\phi}{2}\right) \right]. \quad (9)$$

Although a span of orientations are included in the density estimate, the mean is calculated over only the location-dimension because we do not wish to distinguish among ripples at similar orientations (since minor orientation variations naturally exist in a general ripple field).

The resulting shadow density estimate, calculated from Fig. 5g, is shown in Fig. 5h.

4.6.4 Classification

If the maximum density is above a predefined threshold, τ_d, sand ripples are declared to be present in the original sonar image. The threshold τ_d effectively *defines* the minimum density of shadows, oriented similarly, needed to constitute a ripple field. (In this work, $\tau_d = 0.5$, which dictates that at least 12 shadows at similar orientations within a given distance of each other are needed to permit a declaration of a ripple field.)

The principal (center) orientation of the ripple field is estimated to be the orientation at which the density achieves a maximum. Additionally, the ripple field is deemed to span a set of orientations, with the minimum and maximum orientations of that span set to the orientations at which the density drops to half the maximum density. That is, the span

is defined adaptively based on the evidence supported by the data itself.

For the image related to Fig. 5h, the principal ripple orientation was declared to be $\theta_r = 73°$, spanning the orientations $\Theta_r = [57°, 86°]$. Comparison with the original sonar image in Fig. 5a indicates the ripple orientations were estimated accurately.

If the presence of sand ripples has been detected, those shadows for which the orientation is within the estimated span of ripple orientations, Θ_r, are classified as shadows caused by ripples. The final ripple-shadow map is shown in Fig. 5f, from which it can be observed that the ripple shadows have been detected successfully, albeit with some false alarms.

5 Proposed adaptive data-collection algorithm

5.1 Strategy

The proposed adaptive data-collection algorithm is composed of two distinct surveys. The first survey is executed in an adaptive manner based on the image quality of the data being collected in situ, until quality data has been collected over the entire mission area. The data quality is measured in terms of the coherence (i.e., ping-to-ping peak correlation). Ripple detection is also performed on this data to establish the presence and orientation of any ripples that may exist.

The objective of the second survey pass is to reduce the amount of mission area obscured by ripple shadows. Therefore, the second survey is executed only if sand ripples have been detected during the first survey, and only on those portions of the mission area characterized by sand ripples. This selective re-survey, which depends heavily on the data collected during the first survey, is conducted in an orientation orthogonal to that at which the ripples were estimated to exist.

Additionally, after the first survey pass, a confident estimate of the maximum range for which good quality data can be achieved at the site will be possessed. The tracks of the second survey will also exploit these reliable data-driven estimates of the sonar coverage.

5.2 Implementation: first survey pass

In the previous section, the strategy for the adaptive data collection was outlined. Now this strategy must be translated into an actual plan that can be implemented on the vehicle and executed quickly (so that the next track can be selected as soon as the current track has been completed). The description is overly detailed in parts for the benefit of other researchers wishing to re-implement the algorithm.

5.2.1 Preliminaries

For ease of discussion, assume the mission area that we wish to survey is a parallelogram. If no a priori preference about the initial survey orientation is held, proceed as follows.

Determine which pair of sides of the mission area parallelogram is longer. The vehicle tracks in the mission will run parallel to these sides in order to minimize transit time between tracks. Denote the distance between this pair of longer sides as the mission area width, w.

Next discretize the mission area into $N = w/\delta$ parallel swaths at intervals of δ (e.g., $\delta = 10$ m); the line segments demarcating the swaths will be the locations of the possible tracks. Compute the geographical coordinates of the pair of endpoints of each of the N hypothetical track-centerlines. Store these N pairs of coordinates and associate a unique track identification number, $t_i \in \{1, 2, \ldots, N\}$, with each.

Similarly, assign a unique swath identification number, $s_i \in \{1, 2, \ldots, N - 1\}$, to each δ-sized swath of seabed. Let a binary indicator, $\alpha(s_i)$, denote whether or not swath s_i has been successfully covered to a sufficient level of quality. (Initially, $\alpha(s_i) = 0$ for all s_i.)

Let $\mathcal{T} = \{t_i\}_{i=1}^{N}$ indicate the set of all possible tracks, and $\mathcal{S} = \{s_i\}_{i=1}^{N-1}$ indicate the set of all swaths of seabed.

The set of tracks, $\mathcal{T}$, will be the universe of possible tracks from which the AUV will be allowed to adaptively choose during the survey. (Since there is no human supervision during the onboard decision-making process, this discretization provides a layer of safety that prevents unexpected commands from being made.)

To begin the survey, an initial track near one edge of the mission area—i.e., either the first or last track in the track universe—will be selected. The starting point (and hence direction, also) of the initial track will be chosen to be the endpoint (of the two tracks considered) closest to the vehicle's location, thereby minimizing transit time. All subsequent tracks will be parallel to this first track.[1]

Associated with each track t_i is a set of swaths, $S(t_i)$, that is expected to be covered when the track is executed. This set of swaths is based on geographical considerations and the minimum and maximum ranges of the sonar. Let $k_{\min}$ and $k_{\max}$ be the minimum and maximum range of the sonar, respectively (e.g., $k_{\min} = 40$ m and $k_{\max} = 150$ m).

The coverage quality achieved for each swath, based on the coherence, will be maintained during the course of a track. The determination of whether sonar data of sufficient quality has been collected for a given swath of seabed is calculated as follows.

5.2.2 Quality coverage calculation

Suppose a track, t_i, is being executed. During the course of the track, the AUV's onboard processor will compute the coherence (i.e., peak correlation value between consecutive pings), $\rho(r, u)$, at range r on each sonar side $u \in \{port\ starboard\}$. The range and sonar side associated with a coherence value, when coupled with the track information, will uniquely determine the swath in which each coherence value falls. That is, there exists a mapping from each (r, u) to some particular s_i.

For each swath, a running mean of the coherence values within the swath will be maintained during a given track. By storing only the running mean of the coherence values and the number of values used to produce that mean, the computer storage requirements are independent of track length, requiring space for only two values per swath (or equivalently, range bin). In addition, the mean coherence value calculation will be immediately available at the end of the track, thereby minimizing computation at the critical end-of-track time. Therefore, the AUV can immediately maneuver to begin the subsequent track without delay.

The update equation for the running coherence mean is simply

$$\overline{\rho}(r, u) = \frac{M\overline{\rho}'(r, u) + \rho(r, u)}{M + 1} \tag{10}$$

where the previous coherence mean, $\overline{\rho}'(r, u)$, was calculated from M coherence values, and $\rho(r, u)$ is the new input value.

Once the AUV has completed its track, the mean coherence value of each range bin (of each sonar side) will be immediately available. If this mean value is above a set threshold for image-quality, τ_ρ, the corresponding swath coverage variable, $\alpha(s_i)$, will be set to unity. (Recall that, for a given track, each range-sonar side pair is associated with a particular seabed swath in the mission area.)

Once the coverage variables have been updated, a new track must be selected. The procedure for doing so is as follows.

5.2.3 Track selection

Upon completing the execution of a track, the utility of running each track in the track universe will be calculated. This utility quantity will be comprised of two terms, the benefit expressing the improved coverage expected to be achieved by the hypothetical track, and the cost (in terms of transit distance) of running the track.

The benefit expressing the expected improvement in coverage is calculated as follows.

The historical success rate, ν, of the vehicle in collecting quality sonar data as a function of range r and sonar side u will be maintained during the mission. This success rate will be simply a running tabulation of the proportion of tracks at

---

[1] The adherence to traversing *parallel* tracks is partly because the collected raw data is subsequently processed into imagery, for which such data is preferable.

each range bin that resulted in quality data during the mission, and will be updated after each track is executed.

Initially, the prior expressing our a priori beliefs about the sonar's performance is $\nu(r, u) = 1$ for ranges $r \in [k_{\min}, k_{\max}]$ and sonar side $u \in \{\text{port starboard}\}$. After each track run, $\nu(r, u)$ will be updated according to whether the given area of seabed corresponding to a particular range and sonar side was interrogated with sufficient quality, as

$$\nu(r, u) = \frac{1 + \sum_{i=1}^{T} Q_i(r, u)}{1 + T}, \qquad (11)$$

where $Q_i(r, u)$ is a binary variable indicating whether or not the i-th of T tracks in the present mission resulted in quality coverage at range r from sonar side u (i.e., whether $\overline{\rho}(r, u) \geq \tau_\rho$ during the i-th track).

As will be described below, this quantity will be used to influence the expectation we have about successfully covering any given swath of seabed during a future track. For example, if during the mission, the port sonar has been collecting quality data up to only, say, 100 m rather than the anticipated 150 m, expectations for collecting quality data out to 150 m in the remainder of the mission will be tempered accordingly. That is, the anticipated coverage is adapted in situ based on the data quality collected during the mission thus far.

The benefit of running each particular track, $t_i \in \mathcal{T}$, as the next track is then taken to be the weighted sum of the expected quality coverage

$$B(t_i) = \sum_{r} \sum_{u} \nu(r, u)(1 - \alpha(r, u)), \qquad (12)$$

where $\alpha(r, u)$ is the binary coverage quantity for the swath of seabed at range r on sonar side u, computed with respect to the track centerline of track t_i. It can readily be seen that those swaths of seabed for which quality coverage has already been achieved (i.e., $\alpha = 1$) will contribute nothing to the benefit.

The cost of running each particular track, $t_i \in \mathcal{T}$, will be a function of the transit distance required to reach the new track from the vehicle's present location (at the end of the most-recently executed track). Let $d(t_o, t_i)$ be the (absolute, or unsigned) distance between the most-recently executed track, t_o, and a potential track t_i. Let ω denote twice the turning radius of the vehicle. The cost of executing track t_i, having just executed track t_o, is defined to be

$$C(t_i | t_o) = \left(\frac{d(t_i, t_o)}{\omega} \right)^2 - \frac{2d(t_i, t_o)}{\omega} + 1, \qquad (13)$$

which is the equation of a parabola with a vertex at $(d, C) = (\omega, 0)$, indicating that the transit cost is a minimum (namely, zero) when two tracks are separated by twice the turning radius of the vehicle.

This cost is then weighted by a scaling factor, γ, that reduces the transit penalty as the mission progresses,

$$\gamma = \exp \left\{ \frac{-\sum_{i=1}^{N-1} \alpha(s_i)}{N-1} \right\}. \qquad (14)$$

That is, as more and more of the $N - 1$ seabed swaths are covered to sufficient quality, γ decreases, thereby lessening the impact of the transit penalty. Therefore, this scaling factor acts as a "soft" constraint such that in the beginning of a mission, the vehicle will avoid making long transits to a distant track that would produce only marginally better coverage than a nearby track. Instead, this factor encourages the vehicle to follow a traditional "lawn-mower pattern" that progressively moves from one edge of the mission area to the opposite edge in an orderly manner. When much of the mission area has been covered, the influence of this factor will be smaller, giving the vehicle greater freedom to make longer transits to cover the remaining uncovered swaths of seabed rapidly.

One final stipulation is included for determining the next track to select. Only those potential tracks for which the benefit, B, is positive are considered. This constraint ensures that every track selected contributes toward covering the mission area. (Otherwise, a case could arise in which a track that made no progress in covering new swaths of seabed but possessed a low transit cost had the highest overall benefit of any track.)

The overall utility of running a new track, t_i, having just executed track t_o, is then simply the difference between the expected benefit and the weighted cost,

$$U(t_i | t_o) = \left[B(t_i) - \gamma C(t_i | t_o) \right] \mathbf{1}_{B(t_i) > 0}, \qquad (15)$$

where the indicator function $\mathbf{1}_{B(t_i) > 0}$ is included to ensure only tracks with positive benefits are considered.

The track t_i (that improves the coverage) for which the utility is maximized is then selected as the next track to be executed. The track is forced to begin at the endpoint that is closer to the current location of the vehicle. (Invariably, this means that consecutive tracks will be run in opposite directions, as in a standard survey.)

This entire process is repeated until $\alpha(s_i) = 1 \, \forall s_i$ (i.e., all coverage variables are unity), indicating that quality sonar data has been collected for the entire mission area.

However, as a safety precaution, a second condition that can cause the mission to terminate is related to the number of tracks executed. If the number of tracks executed reaches a predefined maximum number, τ_t, the mission is terminated. This condition is included to guard against the case in which the battery life of the vehicle is insufficient to cover a large mission area.

5.2.4 Ripple detection

While the AUV is collecting data, the ripple detection algorithm from Sect. 4 is executed in a streaming manner, this enabled by the integral-image representation employed. The ripple detection algorithm will determine whether sand ripples are present in the track, and if so, the principal orientation and span of orientations at which the ripples are present.

To combine the ripple detection results from multiple tracks, we collect the span, Θ_r, of orientations that the ripples were deemed to exist at (discretized to 1° precision), across all tracks. The mode of this set of ripple orientations is then taken to be the principal orientation, θ_r, of the ripples in the mission area. For a reasonably sized mission area, the distribution of the ripple orientations should be unimodal since sand ripples are created mainly by currents and waves [36–38].[2]

The result of the ripple detection algorithm effectively performs binary classification, segmenting the mission area according to whether each seabed location is characterized by sand ripple shadows. Let the binary information map, J, at a location (x, y) be equal to unity if the location is *not* characterized by shadows cast by a sand ripple; otherwise, $J(x, y) = 0$. This information map is exploited in the second survey pass, which will seek to ensure the information at all locations is unity.

5.3 Implementation: second survey pass

By construction, after the first survey pass, good quality data will have been obtained everywhere in the mission area. The objective of the second survey pass is to further increase the information content of the data by exploiting the results of the ripple detection algorithm on the data collected during the first survey pass. Specifically, the second survey pass seeks to further refine the mission data by reducing the area that is obscured by shadows cast by sand ripples.

If no ripples have been detected in the first survey pass—and therefore $J(x, y) = 1 \, \forall (x, y)$—the second survey is not executed. If ripples have been detected, at a principal orientation θ_r, a second survey is conducted orthogonal to this orientation. Furthermore, this second survey is performed only over a subset of the mission area, namely in regions characterized by sand ripples.

5.3.1 Objective function

The approach we employ for the second survey pass is adapted from [40], where here the focus is placed on

surveying ripple areas rather than on maximizing object detection probabilities.

The objective function we seek to optimize in the second survey is defined to be the utility of performing a track; this utility is simply the difference of the benefit and the cost of the track. The benefit of the track is defined to be the increase in the information content gained from collecting additional data on an area that had been obscured by ripple shadows during the first survey.

This benefit is calculated using the binary information map, J, constructed during the first survey. Specifically, the benefit at a location (x, y) covered by a track's swath is $B(x, y) = J(x, y) - J'(x, y)$, where $J'(x, y)$ and $J(x, y)$ are the information contents at the given location before and after the new track is run, respectively. For any location (x, y) covered by a track, the information map will be updated to be $J(x, y) = 1$, indicating that the location has now been re-surveyed at the optimal orientation. (It should be noted that ripple shadows may yet persist at the location, but it is not theoretically feasible to continually re-survey until *every* ripple shadow is eliminated, since ripples exist at a continuum of orientations.)

Because a track need not necessarily span the entire length of the mission area, unlike in the first survey pass, the length of each potential track is made to be the minimum length needed to achieve the full benefit of the track. That is, a track whose end portion would cover only areas free of ripples—and hence would not increase the information content—would be terminated prematurely; this reduces energy costs.

The cost of the track is again defined to be the energy costs (e.g., battery)—in units adapted to be consistent with those of the benefit—required to perform the track, which are proportional to the distance of the track (plus transit distance between tracks). This formulation naturally leads to an intuitive termination criterion: perform tracks as long as they would produce a positive utility (i.e., the benefit exceeds the cost).

In practice, the relationship between the benefit and cost of a track depends on the specific mission objectives and must be specified by the controller of the AUV. Defining this relationship will effectively determine the relative importance of re-surveying areas obscured by ripple shadows, and in turn, the number of tracks that can be executed before the algorithm should terminate.

Alternatively, it can be assumed that the AUV has sufficient energy to re-inspect all areas initially deemed to contain sand ripples; in this case, the cost can be ignored so that the utility is equivalent to the benefit.

To maintain the focus on the general nature of the track-selection algorithm, as opposed to effects of application-specific costs, we assume we are operating in this regime hereafter. That is, we assume we can re-inspect all areas initially deemed to contain sand ripples.

---

[2] If multiple sections of the mission area are characterized by substantially different principal ripple orientations, the second survey can be adjusted accordingly to allow tracks of different orientations in different sections of the mission area. In our experience with sonar data, multi-modal ripple orientations have not been observed.

Thus, the challenge of the second survey lies solely on determining the best tracks for the AUV to execute.

5.3.2 Track selection

In the first survey pass, each track was adaptively determined based on the image quality that was obtained. In the second survey pass, we take an alternative approach to more fully exploit the information obtained about the site under investigation during the first survey pass. Specifically, an informed estimate of the maximum range for which image quality was reliably achieved will be possessed after the first survey pass. Therefore, we assume this maximum range in the second survey, which allows the pre-planning of the entire set of second survey pass tracks before any of them are executed.

Thus, the proposed track-selection algorithm consists of a myopic (greedy) search for the best tracks. At each iteration, the expected benefit (and cost) of every possible track is calculated. The track that maximizes the expected utility (or benefit when costs are ignored) at each iteration is chosen and added to the set of tracks to execute. When it is determined that the objective will have been met with the set of selected tracks—i.e., that all of the ripple areas will have been re-surveyed—the algorithm terminates.

Because of the greedy approach employed, the set of selected tracks is necessarily ordered in terms of decreasing utility. Even so, it is not guaranteed that a (local) optimum of the objective function has been reached. To further improve upon the initial set of tracks selected, and to compensate for the adverse effects of the greedy approach, a small geographical displacement of each track is considered. One at a time, we consider the result if a given track is shifted to the port or starboard side one spatial increment, δ. If the objective function increases, the displaced track replaces the previous version of the track. This iterative process continues sequentially for all tracks, one at a time, until no such new shift improves the objective function. This adjustment stage can be thought of as improving closer to the (local) optimum.

A second optional adjustment can also be made. The set of selected tracks are ordered in terms of decreasing utility. However, because we possess this entire set of tracks before any track is executed, the set of tracks can easily be resorted in terms of geography so that the total transit distance (i.e., between the end and beginning of consecutive tracks) is minimized. Although this rearrangement will not impact the *benefit* of the set of the tracks, it *will* minimize energy consumption costs.

It should be noted that the exhaustive, brute-force track search that comprises the main portion of the track selection process for the second survey pass is not as computationally demanding as one may expect. This fact is because the swath coverage of many tracks will not intersect with the swath coverage of most other tracks. As a result, the benefit of certain

tracks can be calculated once and saved. Only the tracks for which the coverage of the previously selected track intersects must be recomputed in a given iteration. The computational savings from exploiting this insight increase proportionally to the size of the mission area to be re-surveyed.

In summary, the proposed track-selection algorithm for the second survey pass consists of three parts. First, an exhaustive greedy search is conducted until the coverage criterion is reached. Second, minor track-adjustments via small centerline displacements are considered. Third, the set of tracks is reordered in terms of geography.

6 Experimental results

In April–May 2008, the NATO Undersea Research Centre (NURC) conducted the Colossus II sea trial in the Baltic Sea off the coast of Latvia. During this trial, high-resolution sonar data was collected by an AUV called MUSCLE, operating at an altitude of 13 m. This AUV is equipped with a 300 kHz sonar with a 60 kHz bandwidth that can achieve image resolution of approximately 3 cm.

In total, 8170 sonar images covering a collective area of 44.93 km^2 were collected during the trial (each sonar image covers approximately 50 m in the along-track direction). Sonar data from this trial will be used to experimentally evaluate the proposed approach.

A subset of this data set from one particular site off the coast of Liepāja, composed of 526 sonar images, will receive additional attention to demonstrate the various components of the proposed data-collection framework as fully as possible. The data collected during the trial at this site, denoted "Area B" here, is particularly appealing because the surveys were performed at five different orientations over the same area of seabed that contained large ripple fields.

6.1 Image quality

To demonstrate the fact that sonar data collected at sea is rarely flawless, for all 8170 images from the Colossus II sea trial, we plot the distribution of the maximum range for which image quality (in terms of the the coherence) was deemed sufficient—i.e., where $\rho > 2/3$—in Fig. 6. The average maximum range (mean $\pm$ one standard deviation) among all images was 135.79 ± 18.13 m, meaning that 13% of the data collected during the trial was of insufficient quality, since each image was assumed to be of good quality out to 150 m.

The maximum range for which good image quality was obtained at Area B, as a function of survey orientation, is shown in Table 1. As can be seen from the table, the image quality was consistent for all five survey orientations. This result justifies the plan, in the adaptive data-collection method, to use the maximum sonar range that reliably
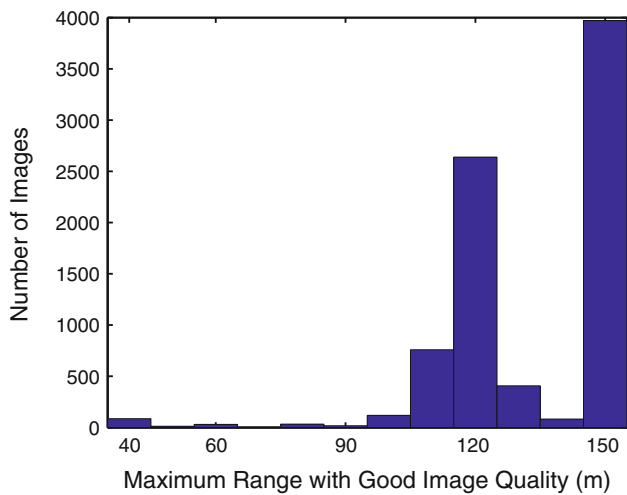
Fig. 6 Of images from the Colossus II sea trial, distribution of the maximum range for which image quality is good

Table 1 Image quality (mean ± one standard deviation) at Area B

Survey orientation	Maximum range (m) with good quality	Fraction of data with good quality
0°	120.63 ± 10.12	0.7330
45°	117.57 ± 18.65	0.7052
70°	127.29 ± 13.84	0.7936
90°	117.18 ± 24.13	0.7017
135°	114.12 ± 20.90	0.6738
All	119.68 ± 18.43	0.7244

portions of the mission area. The existence of such gaps in the data coverage highlights the gains that can be achieved by an adaptive survey strategy. These results also provide support for the manner in which the the first survey pass of the adaptive data-collection plan was constructed, namely based on the image quality.

resulted in good quality data during the first survey pass as the fixed, assumed maximum sonar range for the second survey pass (however, it should be noted that this trend may not hold at sites with steep seabed slopes).

It can also be seen from the table that only 72% of the data collected in this area was of good quality, when it was assumed that quality data would be achieved out to 150 m. As a result, quality data was missing over significant
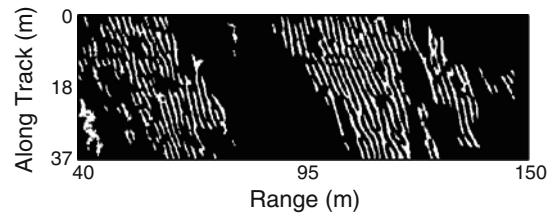
6.2 Ripple detection

To qualitatively demonstrate the ripple detection algorithm, we show the results of the algorithm on two side-scan images in Fig. 7, which contains ripples, and Fig. 8, which does not.
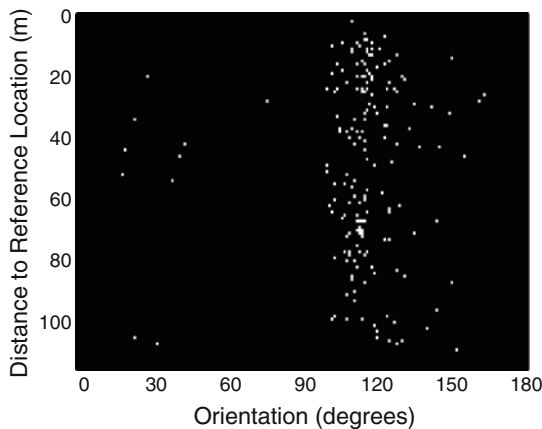
These results support our claim that the lower resolution side-scan imagery is adequate for performing the ripple
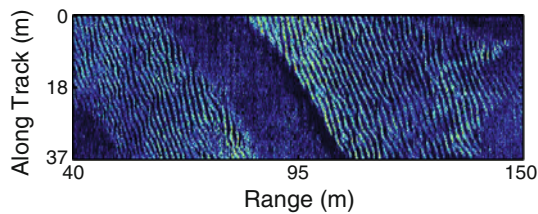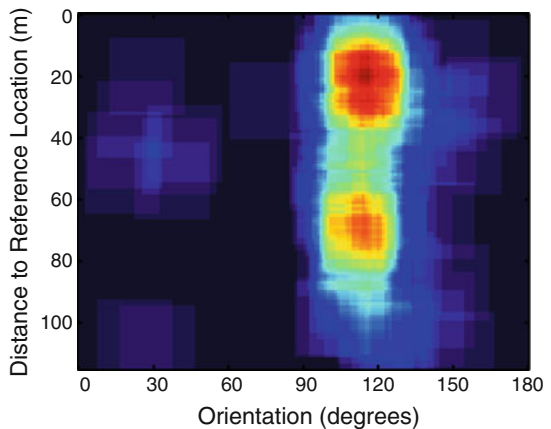

(a) Sidescan sonar image


(b) Final ripple-shadow map

(c) Shadows in location-orientation space

(d) Shadow density in location-orientation space

Fig. 7 Example ripple detection result on a sidescan sonar image with sand ripples. The principal orientation of the ripples was estimated to be $\theta_r = 116°$ (with respect to the across-track direction), with the full set of orientations spanning $\Theta_r = [99°, 130°]$

(a) Sidescan sonar image

(b) Final ripple-shadow map

(c) Shadows in location-orientation space

(d) Shadow density in location-orientation space
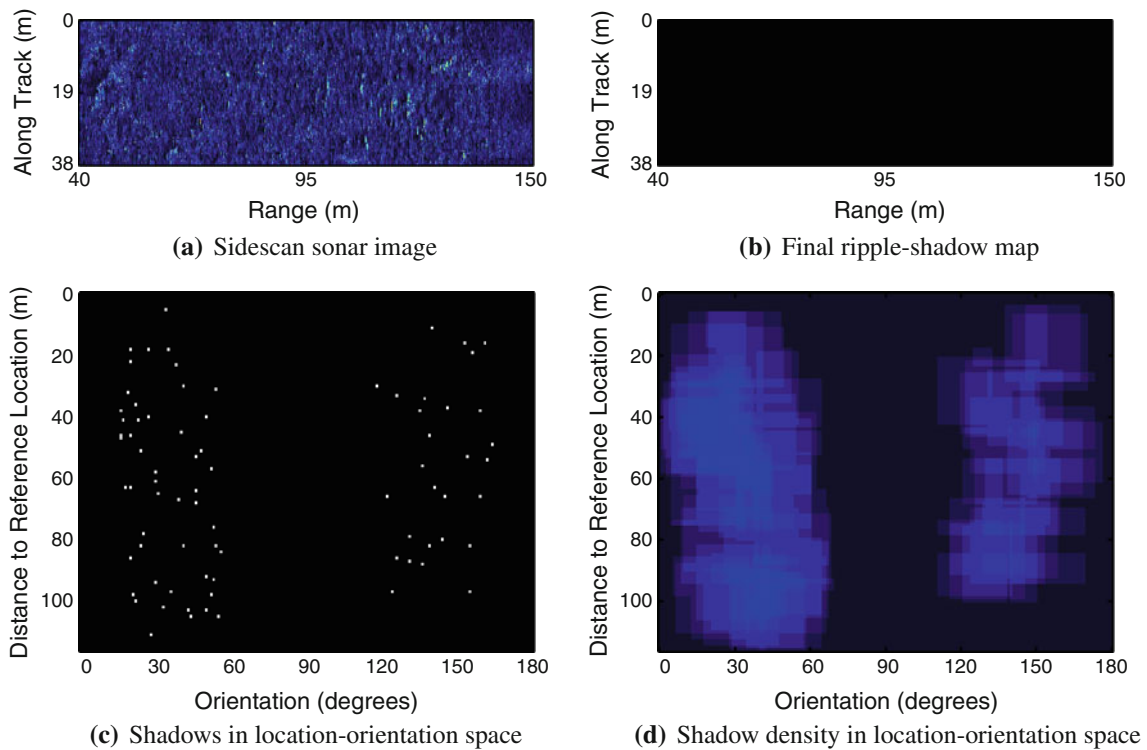
Fig. 8 Example ripple detection result on a sidescan sonar image with no sand ripples

**Table 2** Confusion matrix for ripple detection results at Area B

True status	Predicted status of ripples	
	Present	Absent
Present	0.9425	0.0575
Absent	0	1

detection, thereby allowing us to avoid full SAS image formation onboard the vehicle, if necessary.

We present the results of the ripple detection algorithm, for the 526 images from Area B, quantitatively in the form of a confusion matrix in Table 2. Entries in the table indicate the proportion of images that were classified as containing or not containing sand ripples. Ground-truthing of the images had been performed manually (i.e., by visual inspection).

The results show that the presence of ripples was successfully detected (and the orientation estimated correctly) in 213 images, while 13 images had ripples that were failed to be detected and/or estimated correctly. In the 13 images for which the ripples were not detected, the area in the images covered by ripples was extremely small. All 300 images for which ripples were not present were correctly declared to contain no ripples.

Of the 213 images for which ripples were correctly detected, Fig. 9 shows the proportion of times each orientation—in the reference frame that defines the survey orientations—was included in the images' estimated rip-
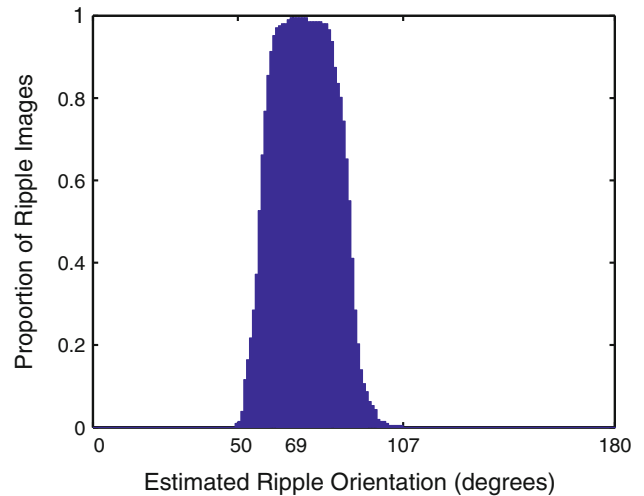
Fig. 9 Proportion of images containing ripples at Area B for which the estimated span of ripple orientations contained each particular orientation

ple spans, Θ_r. The mode of the distribution (and hence the principal ripple orientation at which a "second" survey pass would have been executed) was $\theta_r = 69°$, with the support $[50°, 107°]$. It can be noted that the distribution is unimodal, indicating that the ripples in this data set were indeed directed in a single dominant direction.

The estimated principal ripple orientation as a function of survey orientation, in the reference frame that defines the

Table 3 Estimated principal orientation (mean ± one standard deviation) of sand ripples at Area B

Survey orientation	Estimated principal ripple orientation
0°	75.20° ± 4.60°
45°	72.25° ± 3.23°
70°	74.31° ± 1.05°
90°	73.56° ± 4.26°
135°	72.48° ± 4.35°

Table 4 Area obscured by sand ripple shadows at Area B, and angle between the principal ripple orientation, θ_r, and the survey's across-track angle, θ_a

Survey orientation	Area ($\times 10^4$ m^2) in ripple shadows	Angle between θ_r and θ_a
0°	1.4240	21°
45°	2.0588	64°
70°	2.2945	89°
90°	1.4820	69°
135°	1.5698	24°

survey orientations, is shown in Table 3 for the data from Area B. As can be seen from the small variations in the table, the ripple detection algorithm reliably produces similar ripple-orientation estimates for all five survey orientations.

The purpose of the second survey pass in the proposed adaptive data-collection method is to reduce the area of seabed obscured by ripple shadows. The theoretical results shown in Fig. 4 demonstrated that the area covered by shadows exhibits a strong orientation dependence, with larger shadows cast when the orientations of the ripples and the survey are similar (i.e., when the across-track direction of the survey is furthest from the ripple orientation).

To substantiate this claim with real measured sonar data, we show the total image-area that was calculated to be obscured by shadows cast by sand ripples at Area B, as a function of survey orientation, in Table 4. As in Fig. 4, these results support our argument that the second survey pass should be performed orthogonal to the orientation of the detected ripples so that the ripple shadows can be minimized and the information content of the data maximized.

For example, by surveying at 0° instead of 70°, the area obscured by ripple shadows would be 8705 m^2 smaller. (However, it must be noted that the relative ripple-shadow area realized in the 90° survey orientation case in Table 4 was less than anticipated by theory.) The ideal orientation to survey at this site would have been 159°, since this survey orientation would have an across-track direction that matched the estimated principal ripple orientation, thereby minimizing the area obscured by ripple shadows.

The preceding results on the data from Area B suggest that the ripple detection algorithm reliably detects sand ripples and accurately estimates their orientation. The results also suggest that the decision to perform a second survey pass based on the ripple orientation observed during the first survey pass, in order to reduce the area obscured by shadows, is also sound. Collectively, then, these results provide support for the manner in which the the second survey pass of the adaptive data-collection plan was constructed.

6.3 Adaptive surveying

The experimental results obtained in the previous sections suggest that the various components of the adaptive surveying algorithm were designed reasonably. Finally, we shall demonstrate the full adaptive surveying algorithm in action, as fully as possible given the limitations of the data.

Ideally, the survey algorithm would be demonstrated at sea (as it will be in the near future). In lieu of that opportunity, experiments were conducted here for a simulated, rectangular mission area that is approximately 3.82 km × 1.71 km. The data that is used in the experiments is real measured sonar data, but it is treated as if it was collected in this artificial mission area.

Given this area of interest to be surveyed, we compare the performance of the adaptive survey algorithm with the standard pre-planned survey approach.

The maximum range at which quality sonar data was assumed to be achieved was determined using real data from Area B. To ensure fair comparisons between the two methods, the same image quality was assumed to have been achieved for each method at a given track number. That is, the area for which quality data was collected during the adaptive method's first track was equal to the area for which quality data was collected during the pre-planned method's first track, and so on.
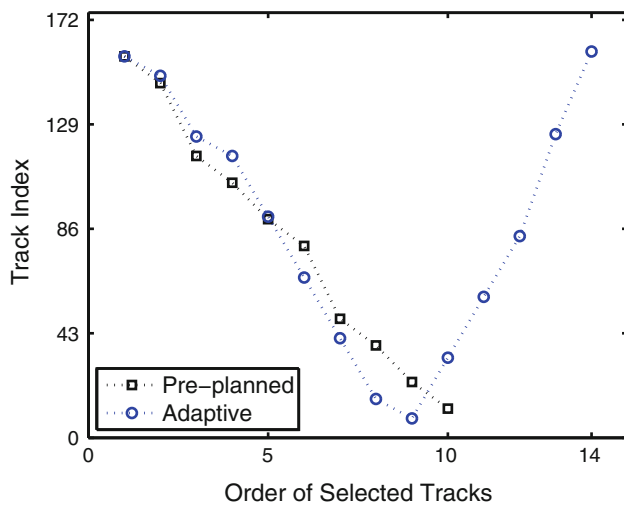
6.3.1 First survey pass

The tracks of the pre-planned survey approach are designed optimally in the sense that the fewest number of tracks must be executed *if* sonar swath coverage is always achieved from 40 to 150 m away from the AUV in both the port and starboard directions. That is, the pre-planned tracks are designed with minimal overlap, fully exploiting the interleaving nature of the swath coverage on consecutive tracks.

For this mission area, the pre-planned survey approach would require 10 tracks (in the lengthwise direction) to collect sonar data everywhere in the mission area. The pre-planned approach cannot adapt its tracks if the image quality collected is not perfect. In contrast, the adaptive survey algorithm executes tracks until sonar data of sufficient quality is collected everywhere in the mission area.
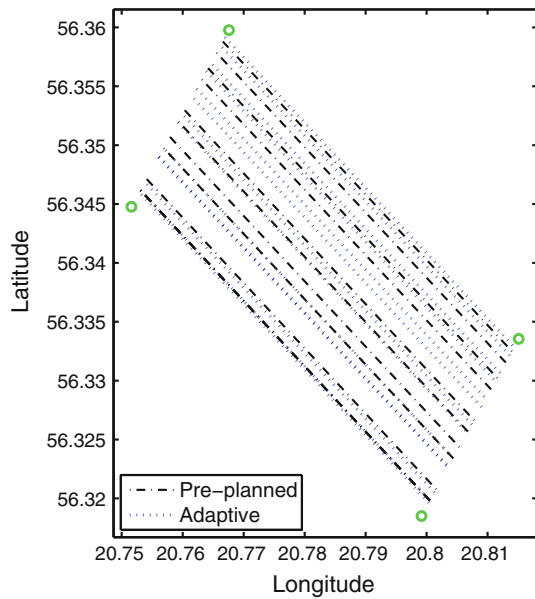
Fig. 10 Map of tracks executed in the first survey pass by the pre-planned and adaptive algorithms


Fig. 11 Order of the tracks that were selected during the first survey pass by each algorithm

The tracks selected by the two competing methods are shown in Fig. 10, where green circles indicate the corners of the mission area of interest.

The universe of possible tracks is discretized at 10 m increments, meaning there exist 172 potential tracks that can be selected by the adaptive algorithm (and the pre-planned algorithm, in fact). These potential tracks are numbered in increasing order vertically. The order of the tracks executed by each method is shown in Fig. 11.

In Fig. 11, it can be observed that the two methods execute different tracks. It can also be observed that the adaptive algorithm executed 14 tracks in order to collect quality data over the entire mission area.
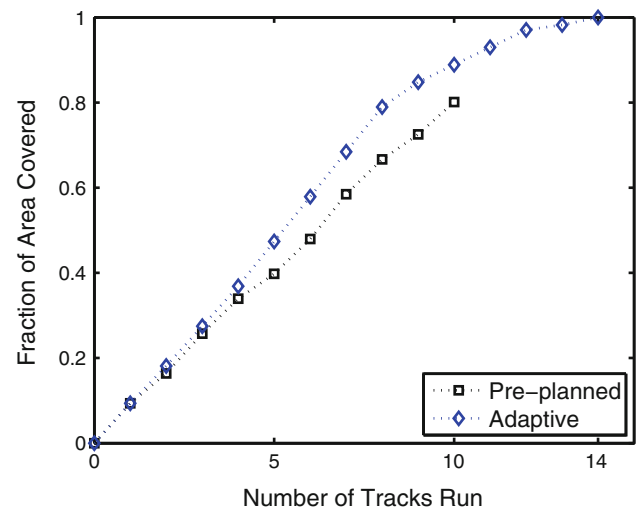
Fig. 12 Progression of the fraction of the mission area for which quality data was collected by each algorithm

It is interesting to note, in Fig. 11, the pattern of the tracks selected by the adaptive method. This method followed a natural progression of tracks down the mission area from the top side of the area to the bottom in its first 9 tracks. Then the method proceeded to execute tracks back toward the top of the mission area. The reason for this is that the algorithm was re-collecting data in the areas for which the image quality had been poor on the first pass through the mission area, effectively filling in the gaps in coverage.

The pre-planned method executed 10 tracks, since that was the number of tracks, in theory, needed to cover the entire area (i.e., collect quality data over 100% of the mission area). But because of poor image quality (cf. Table 1), the pre-planned method did not collect informative data everywhere in the mission area. In fact, after executing its 10 pre-planned tracks, quality data was collected over only 80.12% of the mission area. In contrast, the adaptive survey collected quality data over 100% of the mission area, as it was designed to do.

The progression of the fraction of the mission area for which quality data was collected, as a function of the number of tracks executed by each method, is shown in Fig. 12.

Next, results for the second survey pass will be examined when different environmental scenarios are assumed.

6.3.2 Second survey pass: ripples absent

First assume that the mission area does not contain any ripples.

If sand ripples were not detected during the first survey pass, then the adaptive survey algorithm would not conduct a second survey pass. However, the pre-planned survey method will still cross-hatch the entire mission area. The tracks selected by the pre-planned method for this cross-hatching
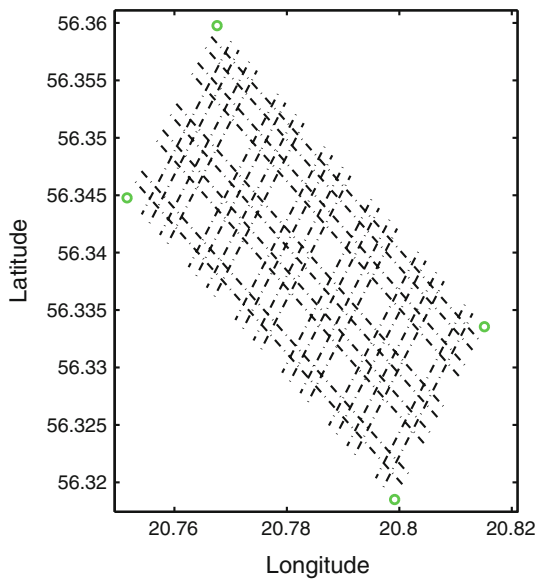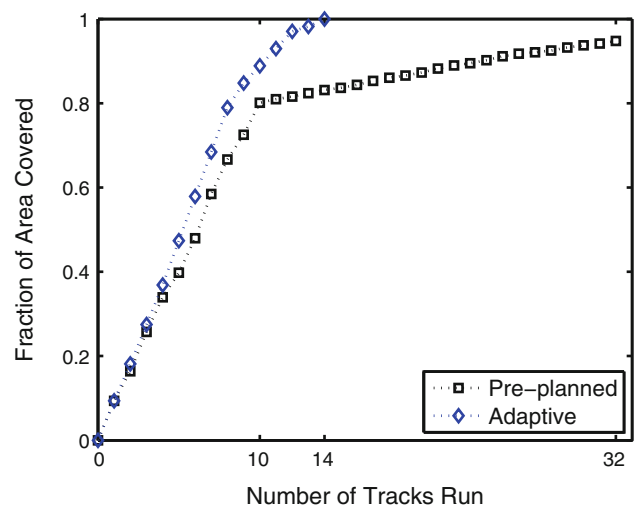
Fig. 13 Map of tracks executed in the second survey pass (and first survey pass) by the pre-planned algorithm
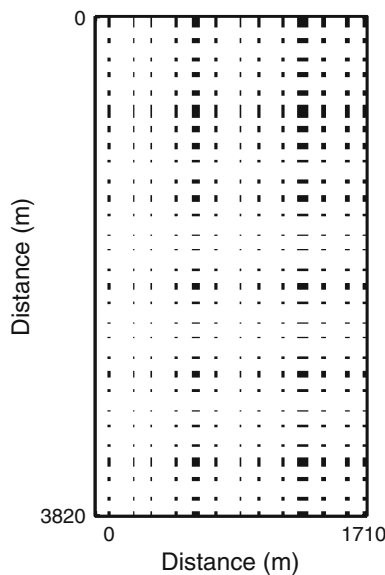
Fig. 14 Coverage map achieved by the pre-planned algorithm after executing two survey passes (*white areas* indicate locations where quality data was collected, *black areas* indicate locations where quality data is lacking)

is shown in Fig. 13, where it was deemed that (because of the wider mission area) 22 tracks were needed to fully cover the mission area in this direction orthogonal to the original survey direction.

If the image quality of the data was sufficient out to a maximum range of 150 m, the pre-planned method would have collected quality data twice for every location in the mission area with the cross-hatching. Because the image quality was flawed, the coverage map that resulted for the pre-planned method is shown in Fig. 14.

Fig. 15 Progression of the fraction of the mission area for which quality data was collected by each algorithm, when the pre-planned method's second survey pass is included

Table 5 Number of tracks executed and total distance traveled for each method

Survey type	Number of tracks		Total distance (km)
	First pass	Second pass	
Pre-planned	10	22	81.35
Adaptive	14	0	61.12

As can be seen in Fig. 14, there exist numerous gaps in coverage because of the image-quality issue and the pre-planned method's inability to react to it. In contrast, the corresponding map of the adaptive algorithm would show coverage everywhere in the mission area after the first survey pass (i.e., the figure would be completely white).

The progression of the fraction of the mission area for which quality data was collected, as a function of the number of tracks executed by each method, when the pre-planned method's second survey pass is included, is shown in Fig. 15. Since no ripples were present in the mission area, the adaptive algorithm did not execute a second pass.

To summarize, the number of tracks executed by each method, for each survey pass, is shown in Table 5. The total distance traveled in executing these tracks (including transit distance between tracks) is also shown. The fraction of the mission area for which data of sufficient quality ($\rho > 2/3$) was obtained is shown in Table 6.

As can be seen from the two tables, the 22 additional tracks the pre-planned method executed during the second survey pass raised the proportion of mission area with quality coverage from 0.8012 to 0.9480, though still short of complete coverage. In contrast, with 18 fewer tracks, the proposed adaptive method was able to achieve complete coverage.

Table 6 Proportion of mission area covered to sufficient image quality by each method

Survey type	Proportion of area covered	
	After first pass	After second pass
Pre-planned	0.8012	0.9480
Adaptive	1.0	1.0
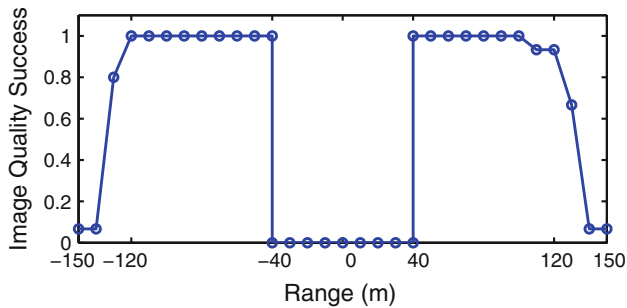
Fig. 16 The final success rate estimating the maximum range at which image quality was achieved, to port and starboard, at the end of the adaptive algorithm's first survey pass

Fig. 17 Map of tracks executed by the adaptive algorithm in the first survey pass (*diagonal dashed lines*) and second survey pass (*horizontal dashed lines*), when sand ripples exist in the area delineated by the *solid box*

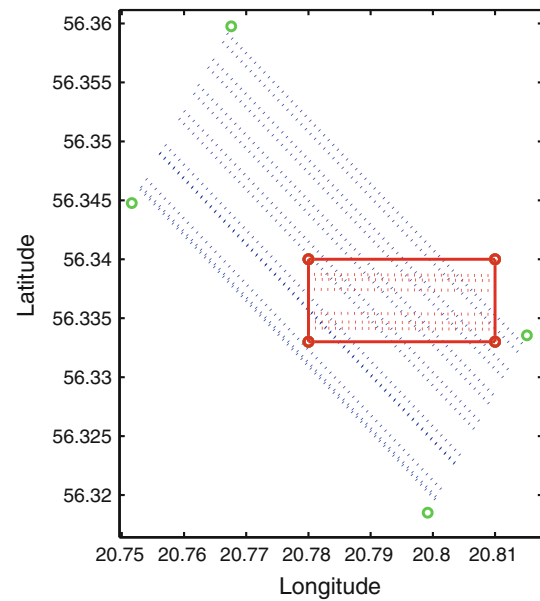### 6.3.3 Second survey pass: ripples present

The results of the second survey pass in the previous section assumed that the mission area did not contain any ripples. Instead, let us suppose that ripples were indeed present at the site. The tracks of the pre-planned method would not change because the method does not react to the data or the environment. But in the case of the adaptive algorithm, the presence of ripples would prompt a second survey pass to be executed.

From the results in Sect. 6.2, it can be seen that the ripple detection algorithm reliably detects ripples and also estimates their orientation accurately. So let us suppose that the ripple detection algorithm resulted in the detection of ripples at a principal orientation of $\theta_c = 90°$ (vertical on the page), and that the ripple field is delineated by the solid box marked in Fig. 17.

According to the adaptive algorithm, the AUV would then execute a second survey pass in the horizontal direction, 0°, so that the orientation of the ripples matches the across-track direction of the survey.

This second survey pass is designed by exploiting the data collected in the first survey pass. Specifically, the maximum range for which quality data was collected during the first pass is assumed to be the true maximum sonar range for planning the second survey pass. This quantity is determined from the final success rate, ν in (11), for which quality data was obtained in the first pass; for this experiment, ν is shown in Fig. 16.

This figure shows the updated prediction of where quality data would be obtained, based on the experience of the

14 tracks run by the adaptive method during the first pass. For the second pass, the maximum range for image quality on each sonar side is set to the range for which quality data was always achieved (i.e., $\nu = 1$). In these experiments, this corresponds to 120 m on the port side and 100 m on the starboard side.

Using these values for the assumed maximum sonar ranges, in conjunction with the estimated principal ripple orientation and the area for which ripples were detected, the adaptive algorithm then plans the tracks for the second survey pass. The resulting six tracks that were planned and executed by the algorithm are shown as horizontal dashed lines in Fig. 17.

As can be seen from Fig. 17, the second survey pass is performed only in the area for which re-surveying is necessary, namely over the area of ripples. Additionally, the tracks were executed in the orientation dictated by the estimated ripple orientation, in order to minimize the area obscured by the ripple shadows.

Because the adaptive algorithm plans the entire set of tracks in the second survey pass before executing them, the algorithm also ordered the tracks to minimize the transit distance. In this experiment, the second survey pass tracks were run from top to bottom.

In conclusion, we have demonstrated the various components of the proposed data-collection strategy as fully as possible given the limitations of using pre-collected, albeit real, data. A future experimental sea trial will demonstrate the full algorithm on the MUSCLE AUV at sea to further prove the method's value.

The adaptive strategy does have its limitations, however. Currently, ripple fields that are detected are enclosed by a bounding box (parallelogram) to facilitate the second survey pass. A more refined approach is needed that would minimize the amount of unnecessary surveying, as would arise when presented with the case of disjoint ripple fields. Another special case that must be addressed is the possibility of steep seabed slopes. This scenario would cause the sonar coverage success to be dependent more on survey direction than on sonar side. Lastly, to be a reliable algorithm employed at sea requires the inclusion of various safety precautions, including obstacle avoidance and vehicle monitoring checks.

7 Conclusion

A new adaptive strategy for performing data collection with a sonar-equipped AUV was proposed. The approach is general in the sense that it is applicable to a wide range of underwater tasks that rely on subsequent processing of side-looking sonar imagery. With the intelligent allocation of resources and the in situ reaction to the data collected, the proposed approach simultaneously maximizes the information content in the data and decreases overall survey time. These improvements are achieved by adapting the AUV route to prevent portions of the mission area from being either characterized by poor image quality or obscured by shadows caused by sand ripples.

To detect the presence of and estimate the orientation of sand ripples, a new innovative algorithm was also developed. The components of the overall adaptive path-planning algorithm were purposely constructed to permit fast real-time execution assuming only minimal AUV onboard processing capabilities.

Experimental results testing the various components of the algorithm, based on real sonar data collected at sea, highlighted the promise of the proposed approach. An upcoming experimental sea trial will further demonstrate the capability of the proposed adaptive surveying algorithm using the MUSCLE AUV.

Future work will refine the second pass of the adaptive survey algorithm to better address the case of disjoint ripple fields, by approximating the ripple fields to be re-surveyed as a stack of single-track blocks (of varying lengths) rather than a single parallelogram (bounding box). Additional effort will be devoted to addressing the special case of steep seabed slopes and other bathymetric anomalies.

Other future work will focus on incorporating additional application-specific objectives into the information-content quantity. For example, a component of the information content for a mine detection task could be defined to be a function of the expected gain in classification confidence provided by interrogating an object at certain aspects [41].

References

1. Williams S, Pizarro O, How M, Mercer D, Powell G, Marshall J, Hanlon R (2009) Surveying nocturnal cuttlefish camouflage behaviour using an AUV. In: Proceedings of the IEEE international conference on robotics and automation, pp 214–219
2. Williams S, Pizarro O, Jakuba M, Barrett N (2010) AUV benthic habitat mapping in south eastern Tasmania. Field Serv Robot 62:275–284
3. Reut Z, Pace N, Heaton M (1985) Computer classification of seabeds by sonar. Nature 314:426–428
4. Carmichael D, Linnett L, Clarke S, Calder B (1996) Seabed classification through multifractal analysis of sidescan sonar imagery. IEE Radar Sonar Navig 143(3):140–148
5. Williams D (2009) Bayesian data fusion of multiview synthetic aperture sonar imagery for seabed classification. IEEE Trans Image Process 18(6):1239–1254
6. Dobeck G, Hyland J, Smedley L (1997) Automated detection/classification of seamines in sonar imagery in Proc. SPIE Int Soc Opt 3079:90–110
7. Williams D (2011) Label-alteration to improve underwater mine classification. IEEE Geosci Remote Sens Lett 8(3):487–491
8. Pastore T, Djapic V (2010) Improving autonomy and control of autonomous surface vehicles in port protection and mine countermeasure scenarios. J Field Robot 27(6):903–914
9. Singh H, Adams J, Mindell D, Foley B (2000) Imaging underwater for archaeology. J Field Archaeol 27(3):319–328
10. Petillot Y, Reed S, Bell J (2002) Real time AUV pipeline detection and tracking using side scan sonar and multi-beam echo-sounder. In: Proc IEEE OCEANS, pp 217–222
11. Hansen R, Callow H, Sæbø T, Synnes S (2010) Challenges in seafloor imaging and mapping with synthetic aperture sonar. In: Proceedings of European conference on synthetic aperture radar
12. Camilli R, Reddy C, Yoerger D, Mooy BV, Jakuba M, Kinsey J, McIntyre C, Sylva S, Maloney J (2010) Tracking hydrocarbon plume transport and biodegradation at deepwater horizon. Science 330(6001):201–204
13. Bender A, Williams S, Pizarro O, Jakuba M (2010) Adaptive exploration of benthic habitat using Gaussian processes. In: Proc IEEE OCEANS
14. Bellingham J, Willcox J (1996) Optimizing AUV oceanographic surveys. In: Proc IEEE symposium autonomous underwater vehicles technology, pp 391–398
15. Alvarez A, Caiti A, Onken R (2004) Evolutionary path planning for autonomous underwater vehicles in a variable ocean. IEEE J Ocean Eng 29(2):418–429
16. Ruiz I, Raucourt S, Petillot Y, Lane D (2004) Concurrent mapping and localization using sidescan sonar. IEEE J Ocean Eng 29(2):442–456
17. Reed S, Ruiz I, Capus C, Petillot Y (2006) The fusion of large scale classified side-scan sonar image mosaics. IEEE Trans Image Process 15(7):2049–2060
18. Hayes M, Gough P (1992) Broad-band synthetic aperture sonar. IEEE J Ocean Eng 17(1):80–94
19. Warren C (1990) A technique for autonomous underwater vehicle route planning. IEEE J Ocean Eng 15:199–204
20. Choset H, Pignon P (1997) Coverage path planning: the boustrophedon cellular decomposition. In: International conference on field and service robotics

21. Pêtrès C, Pailhas Y, Patrón P, Petillot Y, Evans J, Lane D (2007) Path planning for autonomous underwater vehicles. IEEE Trans Robot 23(2):331–341

22. Stack J, Smith C(2003) Combining random and data-driven coverage planning for underwater mine detection. In: Proc IEEE OCEANS, pp 2463–2468

23. Ji S, Liao X, Carin L (2005) Adaptive multiaspect target classification and detection with hidden Markov models. IEEE Sens J 5(5):1035–1042

24. Wang L, Davies G, Bellettini A, Pinto M (2002) Multipath effect on DPCA micronavigation of a synthetic aperture sonar. In: Pace N, Jensen F (eds) Impact of littoral environmental variability on acoustic predictions and sonar performance. Kluwer, Dordrecht pp 465–472

25. Bellettini A, Pinto M (2002) Theoretical accuracy of synthetic aperture sonar micronavigation using a displaced phase-center antenna. IEEE J Ocean Eng 27(4):780–789

26. Williams D (2010) Image-quality prediction of synthetic aperture sonar imagery. In: Proceedings of the IEEE international conference on acoustics, speech, and signal processing, pp 2114–2117

27. Synnes S, Callow H, Hansen R, Sæbø T (2010) Area coverage rate of synthetic aperture sonars. In: Proceedings of European conference on underwater acoustics

28. Synnes S, Hansen R, Sæbø T (2009) Assessment of shallow water performance using interferometric sonar coherence. In: Proceedings of international conference on underwater acoustic measurements

29. Groen J, Coiras E, Williams D (2009) Detection rate statistics in synthetic aperture sonar images. In: Proc Intl Conf & Exh underwater acoustic measurements, pp 367–374

30. Coiras E, Groen J (2009) Simulation and 3d reconstruction of side-looking sonar images. In: Kordic V (ed) Advances in sonar technology. IN-TECH 2009

31. Williams D, Coiras E (2010) On sand ripple detection in synthetic aperture sonar imagery. In: Proceedings of the IEEE international conference on acoustics, speech, and signal processing, pp 1074–1077

32. Nelson J, Kingsbury N (2010) Fractal dimension based sand ripple suppression for mine hunting with sidescan sonar. In: International conference on synthetic aperture sonar and synthetic aperture radar

33. Viola P, Jones M (2004) Robust real-time object detection. Int J Comput Vis 57(2):137–154

34. Widmer G, Kubat M (1996) Learning in the presence of concept drift and hidden contexts . Mach Learn 23:69–101

35. Zadrozny B (2004) Learning and evaluating classifiers under sample selection bias. In: Proceedings of the 21st international conference on machine learning

36. Wiberg P, Harris C (1994) Ripple geometry in wave-dominated environments. J Geophys Res 99:775–789

37. Hanes D, Alymov V, Chang Y (2001) Wave-formed sand ripples at Duck, North Carolina. J Geophys Res 106:22-575–22-592

38. Reffet E, du Pont SC, Hersen P, Douady S (2010) Formation and stability of transverse and longitudinal sand dunes. Geology 38(6):491–494

39. Lyons A, Abraham D, Johnson S (2010) Modeling the effect of seafloor ripples on synthetic aperture sonar speckle statistics. IEEE J Ocean Eng 35(2):242–249

40. Williams D (2010) On optimal AUV track-spacing for underwater mine detection. In: Proceedings of the IEEE international conference on robotics and automation, pp 4755–4762

41. Runkle P, Nguyen L, McClellan J, Carin L (2001) Multi-aspect target detection for SAR imagery using hidden Markov models. IEEE Trans Geosci Remote Sens 39(1):46–55