Sonar Point Cloud Processing to Identify Sea Turtles by Pattern Analysis

Dror Kipnis, Yaniv Levy, and Roee Diamant, *Senior Member, IEEE*

***Abstract*—Abundant in coastal areas, sea turtles are affected by high-intensity acoustic anthropogenic sounds. In this article, we offer a pattern-analysis-based detection approach to serve as a warning system for the existence of nearby sea turtles. We focus on the challenge of overcoming the low signal-to-clutter ratio (SCR) caused by reverberations. Assuming that, owing to low SCR, target reflections within the point cloud are received in groups, our detector searches for patterns through clustering to identify possible "blobs" in the point cloud of reflections, and to classify them as either clutter or a target. Our unsupervised clustering is based on geometrical and spectral constraints over the blob's member relations. In turn, the classification of identified blobs as either a target or clutter is based on features extracted from the reflection pattern. To this end, assuming that reflections from a sea turtle are stable but include spectral diversity due to distortions within the turtle's body, we quantify the stability of the blob's members and the entropy of their reflection spectrum. We test our detector in both the modeled simulations, and at sea, for the detection of sea turtles released after rehabilitation. The results show robustness to highly fluctuating target intensity and ability to detect at low SCR.**

***Index Terms*—Active sonar, clustering, detection of sea turtles, pattern recognition, point cloud processing, spectral diversity.**

I. Introduction

OWING to the increase in marine anthropogenic activities, it was found that megafauna are threatened by high acoustic intensities with both the physical and behavioral impacts [1]. While marine mammals and pelagic fish can quickly escape high-intensity acoustic sources, this is not the case for sea turtles. Hence, there is a need for a detection solution that can serve as a warning system to alert the presence of sea turtles to stop or mitigate disturbance or detriment during marine operations. This type of system can be applied over vessels performing seismic surveys that may lead to the interruption of marine turtles' normal behaviors [1], within areas of seawater intake to desalination factories that can injure turtles [2], within the proximity of areas containing hot water from power plants where turtles aggregate [3], as well as within ports and ship channels and dredging zones where turtles are abundant [4]. The identification of sea turtles' presence in their natural habitat can also be an important tool for population monitoring and conservation. This article takes a first step toward such a goal and offers a pattern recognition approach for detecting sea turtles by active acoustics.

The active acoustic detection of sea turtles was studied for commercial echosounders, where target strength (TS) measurements are conducted at a frequency of 200 kHz in a tank and in the open sea [5], [6]. Other related studies [7] use active acoustics to detect fish, based on the resonance of their swimming bladder. However, when operating at frequencies that differ from the resonance or for species without a swimming bladder, acoustic reflection is expected to be less intense. Detection methods to identify reflected sonar echoes for very low observable (VLO) targets are studied extensively and include methods to reduce the false alarm rate (FAR) by utilizing information on reverberation and noise power [8], classification of sonar contacts to target or clutter [9], [10], tracking based on a motion model [11], or by probabilistic analysis to identify the sequences of reflections based on dynamic programming [12]. However, to the best of our knowledge, no consideration has been given to date to the time-varying characteristics of reflections from targets like sea turtles. In particular, a main challenge of the active sonar detection of sea turtles is to identify reflections originating from a turtle within a point cloud of clutter reflections. The problem becomes further compounded when detection is omnidirectional and the signal-to-clutter ratio (SCR) is low [13]. On top of this, for a swimming sea turtle that changes its orientation and depth with respect to the receiver, the intensity of the acoustic reflections varies over time. Thus, rather than a continuous reflection pattern, we expect the reflection pattern from the turtle to form clusters, or "blobs," within the point cloud of received sonar echoes.

In this article, we introduce a novel pattern-recognition-based solution for sea turtle presence detection. Our solution considers the expected time variation of the reflections from sea turtles to reduce the FAR through point cloud clustering and classification. We start by accumulating a sequence of reflections from the active sonar signals and arrange them such that each reflection is identified by its ping number and estimated range. This results with a sparse representation of the acoustic reflections by their (range, time) coordinates, which we refer to as a point

Manuscript received 30 May 2022; revised 28 September 2022; accepted 6 October 2022. Date of publication 8 December 2022; date of current version 14 April 2023. This work was supported in part by the Israeli Ministry of Energy, Action on Environmental Impact Assessment, under Grant 219-17-013. *(Corresponding author: Roee Diamant.)*

Associate Editor: A. Trucco.

Dror Kipnis and Roee Diamant are with the Hatter Department of Marine Technologies, University of Haifa, Haifa 3498838, Israel (e-mail: dkipnis@campus.haifa.ac.il; roee.d@univ.haifa.ac.il).

Yaniv Levy is with the Morris Kahn Marine Research Station, Department of Marine Biology, Leon H. Charney School of Marine Sciences, University of Haifa, Haifa 3498838, Israel, and also with the Israel Sea Turtle Rescue Center, National Nature and Parks Authority, Beit Yannai National Park, Beit Yanai 4029300, Israel (e-mail: yaniv@npa.org.il).

Digital Object Identifier 10.1109/JOE.2022.3214274

cloud. Then, spectral characteristics and geometrical relations between the reflection's temporal and spatial patterns are tested against limitations on the animal's speed and size. Reflections are represented as the nodes of a weighted graph, and clustering over the resulting graph is performed to identify the blobs of reflections associated with potential targets. The identified blobs are, then, classified as either a target or clutter by examining features derived from the statistics of the cluster's members and from the reflections' spectral diversity. The latter is based on our key assumption that, owing to distortions within the turtle's body, the spectral response of reflections originating from a sea turtle is significantly more diverse than that of the clutter. Based on this assumption, we use the entropy of the reflections' spectral response as a feature to classify between targets and clutter. Furthermore, since we expect clutter reflections to be independent while that of a valid blob to show statistical relations, we validate a blob to be related to a target by the stability of its members.

Our contributions are as follows:

1) the design of a graph-based clustering solution that accounts for time variations within a point cloud of sonar's reflections;
2) the utilization of the statistical dependence between the target's reflections as a measure for pattern recognition;
3) the design of a classification metric that uses as a feature the expected spectral diversity of acoustic reflections originating from a sea turtle.

We explore the performance of our method in both simulations and a designated sea trial. The simulations are based on a physical model for reflections from a sea turtle and involve real clutter collected during multiple sea trials. The sea trial involved the detection of two sea turtles we released after rehabilitation. Results show a favorable tradeoff between the false alarm and detection rates and resilience to low and time-varying SCR.

The rest of this article is organized as follows. Section II surveys the related work in the fields of active sonar detection, classification of sonar reflections, and track-before-detect (TkBD) approaches for VLO targets. Section III describes the system's model and assumptions, as well as a model for acoustic reflections from sea turtles. Section IV details our algorithm, starting from the clustering solution, and followed by the classification of clusters. Section V reports the simulation and sea trial results. Finally, Section VI concludes this article.

Throughout this article, we use the following notations: a small letter indicates a scalar, while vectors and matrices are indicated by bold small and uppercase letters, respectively. The entries of a vector or a matrix are indicated by subscript indices.

II. Related Work

The common approach for detecting the reflection of active sonar in Gaussian noise is by a matched filter (MF) [14], which is used widely for active sonar applications [9], [10], [15]. However, since sound propagation in shallow water is considerably affected by reverberations, the MF's performance deviates from its optimum processing gain [8]. A method to detect sonar echoes in a reverberant environment is proposed in [8], where the Page test is applied on the MF output to identify the start and end of a signal. Alternatively, for the specific case of linear frequency modulation (LFM) signals, emitted for object detection, a fractional Fourier transform (FrFT) can be applied to separate signals overlapping in time [16]. An application of the FrFT for sonar signal processing is given in [17] for the parameter estimation of chirp signals. Detection based on simulations and experimental data is illustrated in [18], and a better accuracy than that of the MF is observed in terms of the delay estimation. While the above detection methods aim to identify echoes in the low SCR and in conditions of high interference, in practice, owing to noise fluctuations, high FARs [10] are reported in real sea environments.

One way to reduce false detections is by classifying the acoustic reflections to separate between a target and clutter. In [10], a machine-learning approach was demonstrated to reduce the FAR of active sonar reflections by using the backscatter intensity as a classification feature. Similarly, the range-bearing neighborhood of MF sample outputs that passed a detection threshold is used in [15] to extract features for classification by a support vector machine. In [9], a similar feature set is offered as the input of a neural network to classify the sonar data after MF operation. Alternatively, De Magistris et al. [13] implemented a convolutional neural network to classify a target or clutter by the sonar's spectrogram images. The drawback is the need for training data, which is often hard to obtain due to the spatial and temporal diversity of the underwater acoustic channel. Unsupervised anomaly detection [19] is used to overcome this limitation, making it more practical for real-world applications. However, the classification is limited to individual detected instances and does not exploit the information of the reflection pattern.

The temporal information of the reflection pattern is obtained when emitting a train of sonar pings whose reflections, after being synchronized, are stacked in a reflection matrix. Analyzing such a matrix can dramatically increase detection performance by using the TkBD framework. Common approaches to TkBD include dynamic programming, particle filter (PF), probabilistic multihypothesis tracking (PMHT), and probabilistic data association (PDA). Dynamic programming performs tracking over a discrete grid, assuming that the target's dynamics can be modeled as a Markov process with some limitations on its velocity [20]. Although straightforward, it has the drawback of having a very large computational cost. Application for weak target detection by active sonar can be found in [21] on simulated data. Viterbi algorithm (VA)-based TkBD for underwater targets is presented in [12], where detection decision is based on the calculated likelihood of optional target paths. For lower complexity, the PF samples the state space to estimate the probability density function (PDF) for target reflections [11].

The tracking of an underwater moving source by the PF is demonstrated in [22], while Northardt and Nardone [23] demonstrated a PF-based TkBD algorithm on a multitarget passive sonar scenario. In turn, the PMHT approach [24] models a combination of target and noise using the expectation–maximization algorithm and estimates the target's path by a Kalman smoother. Extensions include the histogram probabilistic multihypothesis tracking (H-PMHT) [25], [26] and the Poisson-H-PMHT [27]. Finally, the PDA approach considers the overall probability of

the set of measurements to infer the existence of a target [28]. The applications of ML-PDA for real bistatic and multistatic sonar data are demonstrated in [29] and [30], respectively. A comparison of ML-PDA versus PMHT on synthetic multistatic active-sonar scenarios is presented in [31]. An underlying assumption for these tracking-based methods is that a target exists, and that its probability of detection is roughly constant within the observation time (i.e., the series of pings). While such an assumption may be valid for large targets such as submarines, this is not the case for small targets as considered here, i.e., a sea turtle. This is because of the high impact of orientation changes, caused by the animal's motion, on the intensity of the acoustic reflections.

III. System Model

Our setup includes an omnidirectional monostatic sonar, composed of a projector and hydrophone pair. A number, N_{p}, of LFM chirp signals are emitted at a constant ping rate interval (PRI), T_{PRI}, such that reflected echoes allow nearly uniform temporal sampling of the surrounding. In this article, we set $N_{\mathrm{p}} = 20$ for simulations and $N_{\mathrm{p}} = 37$ for the experiment. For each emission, we record the reflected signal, such that the analysis is performed in blocks of time windows equal to $N_{\mathrm{p}} \cdot T_{\mathrm{PRI}}$. To utilize the reflections' expected time variations, we arrange the N_{p} reflections as a set of signals $\{\mathbf{y}^{(p)}\}_{p=1}^{N_{\mathrm{p}}}$, attributed by their pulse index p. Our goal is to find target reflection patterns within these signals.

A. Main Assumptions

Our model assumes that, owing to the dynamics of the sea turtle, the intensity of the acoustic reflections forms "blobs" along the target's path at the MF output. We also assume that, owing to the complex structure of the sea turtle's body, the acoustic intensity of these blobs' members is frequency dependent. Under this model, our detector combines classification with tracking.

We assume that a valid reflection that originates from a sea turtle follows three conditions. First, the detected path should follow boundaries in terms of the turtle's maximum speed. Specifically, the range associated with a valid reflection must be distanced up to $v_{\max} \cdot T_{\mathrm{PRI}}$ [m] from its previous or following valid reflection. Second, a valid reflection must show spectral diversity that differs from that of the clutter. Finally, we expect the samples along a valid target's path at the MF output to be statistically dependent. In terms of the channel model, we focus on the practical case in which 1) the reflections are received at low SCR, such that energy-based detection is not possible, and that 2) there is nonnegligible ambient noise; therefore, statistical analysis over a number of emissions is required.

B. Model for Acoustic Reflections From Sea Turtles

Studies on the distribution of acoustic reflections from sea turtles focus on a commercial ecosounder at a carrier frequency of 200 kHz [5], [6]. Especially for omnidirectional detection, this frequency greatly limits the detection range, as the acoustic attenuation above the 100 kHz becomes a dominant factor in the attenuation loss. Instead, our transmissions are performed at lower frequencies of 7–17 kHz and between 48 and 78 kHz, where the acoustic wavelength is on the same order as the dimensions of the target sea turtle. For this frequency range, the turtle's reflections are characterized as frequency dependent. The frequency dependence of acoustic reflections is mostly studied for fish [32], [33], [34]. Here, the acoustic reflections primarily originate from the fish's swim bladder, which, arguably, resemble the reflections from a sea turtle's lungs. In terms of the reflection pattern, a major difference between turtles and fish is that, while reflections from a fish mainly originate from its swim bladder, reflections from sea turtles originate from both the turtle's lungs and its hard bony shell. We argue that, owing to this difference, the frequency response of the reflection pattern from a sea turtle is more diverse than that of a fish. The general model for the spectral diversity of acoustic reflections we refer to is described in [35]. For completeness, we now outline this model in detail below.

The frequency dependence of reflections from a finite-sized object can be divided into three regions, depending on the size of the reflecting object relative to the wavelength. When the acoustic wavelength λ is shorter than the object's size, l, the reflection is modeled as Rayleigh scattering, in which the TS rapidly increases with the frequency. Geometric reflection, which is not frequency dependent, occurs at the optical limit when the wavelength is much smaller than the object's size. Between these two extremes, there is an interference region in which multiple reflections from the object's boundaries interfere, thus creating a diverse frequency selective pattern. For a well-defined object, there is a clear division between the above regions. For example, for a sphere with a radius of a, the interference region is in the range of $1 \leq 2\pi a/\lambda \leq 10$. However, for complex shapes like a marine animal, this division is not clear, and the animal's internal organs are expected to create a diverse interference pattern. Empirical results [35] identify the interference region for fish as $0.7 \leq l/\lambda \leq 200$, where the lower limit may be even smaller. For our first transmission range, 7–17 kHz, the size of an object that is within the interference region is between

$$l_{\min} f_{\min} \geq 0.7c \rightarrow l_{\min} \geq \frac{0.7c}{f_{\min}} = \frac{0.7 \cdot 1500}{7000} = 0.15\,\mathrm{m} \quad (1)$$

$$l_{\max} f_{\max} \leq 200c \rightarrow l_{\max} \leq \frac{200c}{f_{\max}} = \frac{200 \cdot 1500}{17000} = 17.647\,\mathrm{m}\,. \quad (2)$$

Similarly, for the frequency range of 48–78 kHz, the interference region fits a target size between 2 cm and 3.8 m. We, therefore, conclude that the acoustic reflections expected from the sea turtle are within the interference region, and a high diversity in the frequency domain is expected. In contrast, for clutter composed of smaller elements, the acoustic reflections are within the Rayleigh scattering region, and its frequency response is expected to be different than that of a sea turtle. Our approach utilizes this difference in the frequency diversity characteristics of the two signals to separate the turtle's reflections from those of the clutter.

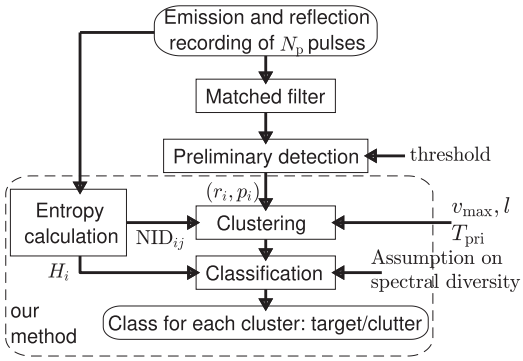

Fig. 1. Block diagram of the processing chain.

IV. Methodology

A. Key Idea

Our processing chain flow is shown in Fig. 1. Our solution utilizes our assumption that the acoustic reflections from sea turtles tend to appear as occasional blobs along the turtle's path. This is due to expected temporal changes in the turtle's orientation and in the volume of its lungs. The input for our classifier is the set $\{(r_i, p_i)\}_{i=1}^{N}$, which is the point cloud after MF thresholding. Here, r_i represents the range of the point from the sonar, and p_i is the ping number, as illustrated in Fig. 2(a) and (b).

Referring to Fig. 1, ours is a two-stage process. First, we segment target indications into the clusters of potential targets. To this end, we identify reflections that are close in range and time and, thus, are more likely to be from the same object. As part of the clustering solution, we rank the connectivity between data points by their mutual information. Instead of presetting the number of clusters, which in our case is not possible, we determine it by minimizing a utility function for the "cost" of the clustering solution. Once clusters are formed in the shape of "blobs," we move to the second phase where we analyze the blob's members to draw features about their statistics and to determine whether the blob is an indication of a sea-turtle-like target. These features comprise, on the one hand, the spatial, temporal, and statistical connectivity of the acoustic reflections along the blob and, on the other hand, the spectral diversity that is expected for the reflections from a sea turtle. The former is calculated by summing all the connections within a cluster, while the latter by the entropy of the acoustic reflections at different frequencies. In doing so, we achieve a processing gain by analyzing groups of reflections.

We note that our solution can manage two cases: one when an exact spectral diversity model is available, and one when no model is available, but spectral diversity is expected. The first calls for ranking the alignment between the measured diversity and a given template, while the second compares the spectral diversity of the blob to that of the clutter. In turn, clutter is identified by clusters whose members' connectivity is low, or whose members' entropy is high. An important notification is that, while we demonstrate our solution for the detection of sea turtles, our solution may also be applied to a variety of targets—ranging from pelagic fish to top predators. In each case, the user should provide a mobility model (limitation on the speed and movement pattern) and a spectral model, if any.

B. Clustering

The goal of the clustering phase is to segment the thresholded data points in blobs of possible targets. These blobs should contain data points that are close in range and time and have similar reflection spectrum. To measure this similarity, we use the normalized information distance (NID) [36], which is calculated for the frequency spectrum of the reflections from each pair of data points. Denote the time domain signal corresponding to the ith data point by vector $\mathbf{y}^i$ and its normalized power spectrum as $\mathbf{z}^i$ with M frequency bins, z^i, such that $(\mathbf{z}^i)^{\mathrm{T}}\mathbf{1} = \mathbf{1}$. We calculate the PDF of $\mathbf{z}^i$ by dividing it to J histogram bins, ξ_j, and determine its entropy by

$$H_i = -\sum_{j=1}^{J} \mathrm{p}\left(z^i = \xi_j\right) \log_2\left(\mathrm{p}\left(z^i = \xi_j\right)\right) \tag{3}$$

where $\mathrm{p}(z^i = \xi_j)$ is the numerically calculated PDF of $\mathbf{z}^i$. The NID used to quantify spectral similarity between data points i and j is calculated by

$$\mathrm{NID_{ij}}\left(\mathbf{z}^{\mathrm{i}}; \mathbf{z}^{\mathrm{j}}\right) = 1 - \frac{\mathrm{I_{ij}}\left(\mathbf{z}^{\mathrm{i}}; \mathbf{z}^{\mathrm{j}}\right)}{\max\left[\mathrm{H_i}\left(\mathbf{z}^{\mathrm{i}}\right), \mathrm{H_j}\left(\mathbf{z}^{\mathrm{j}}\right)\right]} \tag{4}$$

where I_{ij} is the mutual information between $\hat{\mathbf{z}}^{(i)}$ and $\hat{\mathbf{z}}^{(j)}$ obtained by

$$I_{ij}\left(\hat{\mathbf{z}}^i; \hat{\mathbf{z}}^j\right) = H_i\left(\hat{\mathbf{z}}^i\right) + H_j\left(\hat{\mathbf{z}}^j\right) - H_{ij}\left(\hat{\mathbf{z}}^i, \hat{\mathbf{z}}^j\right) \tag{5}$$

and $H_{ij}(\mathbf{z}^i, \mathbf{z}^j)$ is the joint entropy of $\mathbf{z}^i$ and $\mathbf{z}^j$

$$\begin{aligned} H_{ij}\left(\mathbf{z}^i, \mathbf{z}^j\right) = & -\sum_{l}^{J}\sum_{m}^{J} \mathrm{p}\left(z^i = \xi_l, z^j = \xi_m\right) \\ & \times \log_2\left(\mathrm{p}\left(z^i = \xi_l, z^j = \xi_m\right)\right) \end{aligned} \tag{6}$$

where $\mathrm{p}(z_l^i, z_m^j)$ is the joint PDF of the normalized reflection values from points i and j, calculated numerically by a joint histogram with J bins at each dimension. The reason we use the NID measure is because it relates to the mutual information of the reflection spectrum of each pair of data points [36]. Specifically, a value of NID close to 0 indicates a high level of mutual information, while an NID close to 1 is a case where $\mathbf{z}^i$ and $\mathbf{z}^j$ are independent. The range difference of two data points is denoted by $\Delta r_{ij} = |r_i - r_j|$, and their time difference by $\Delta t_{ij} = |p_i - p_j| T_{\mathrm{PRI}}$. We define the distance of two data points i and j by

$$\Delta_{ij} = \alpha \mathrm{NID}_{ij} + \beta \Delta r_{ij} \mathrm{u}\left(\Delta r_{ij} - l\right) + \tau \Delta t_{ij} \tag{7}$$

where $\mathrm{u}(\cdot)$ is the Heaviside step function and l is the expected size of the target, such that Δ_{ij} is not sensitive to $\Delta r_{ij} < l$. The parameters α, β, and τ tune the relative importance of the frequency diversity, the range difference, and the ping difference, respectively.

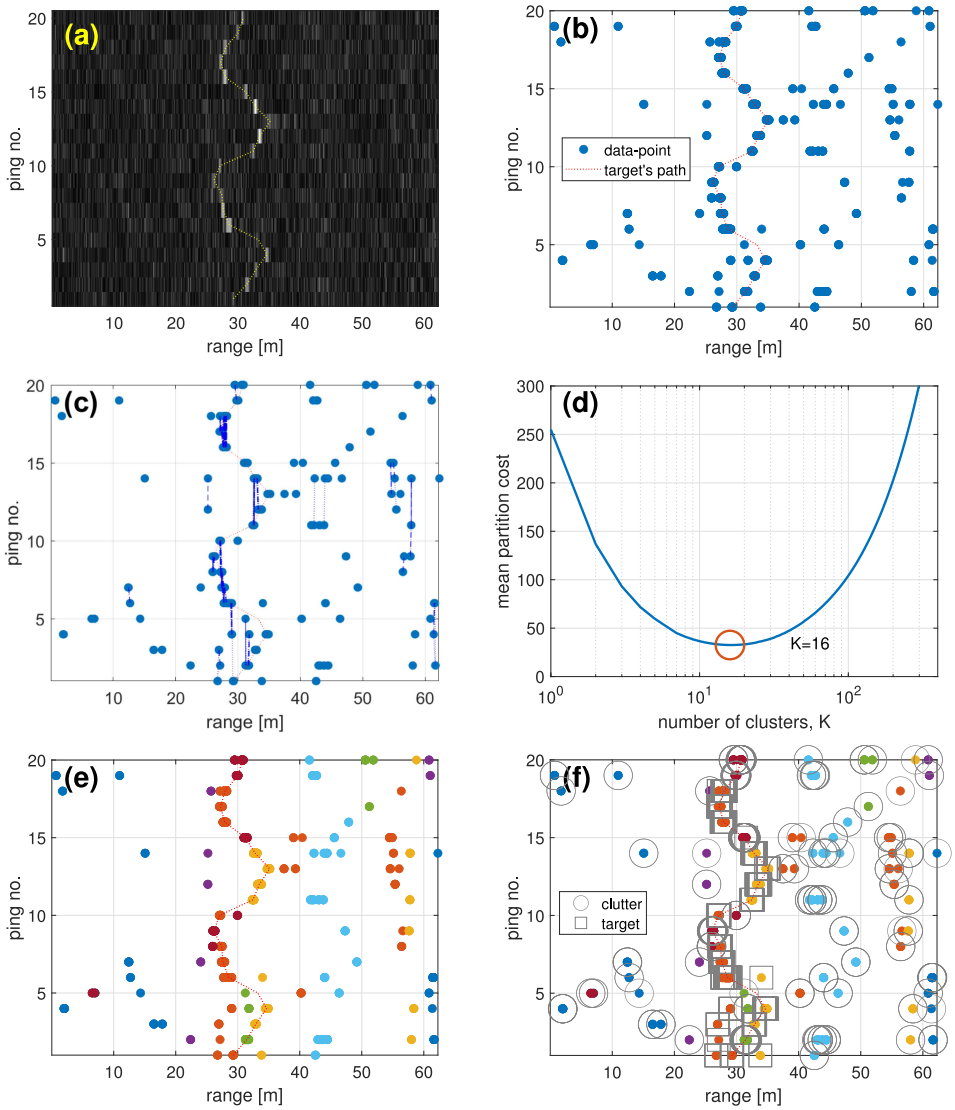

Fig. 2. Processing stages. (a) Reflection matrix after MF. (b) Reflection matrix after preliminary detection. (c) Graph formation. (d) Choosing model order ($\epsilon = 1$). (e) Clusters (color coded). (f) Classification.

Our clustering solution is based on the method proposed in [37], which performs spectral clustering to a predefined number of clusters. Here, we perform an additional search to find the most appropriate number of clusters. This spectral clustering method was selected because it enables the formation of nonconvex clusters as expected in our application (i.e., the maneuvering sea turtle does not move in straight lines). In this method, the data are encoded as a weighted graph, where data points are the nodes and their edges are weighted according to distances Δ_{ij} defined in (7). We define the weight of the edge between the ith and jth data points by

$$w_{ij} = e^{-\Delta_{ij}} \mathrm{u}\left(v_{\max} - \frac{\Delta r_{ij}}{\Delta t_{ij}}\right) \tag{8}$$

where the step function constrains the maximal velocity allowed by the model $v_{\max}$. This processing stage is illustrated in Fig. 2(c), where the graph edges are marked as dashed lines.

Following [37], we define a matrix $\mathbf{A}$ of entries

$$a_{ij} = \begin{cases} w_{ij}, & i \neq j \\ 0, & i = j \end{cases} \tag{9}$$

and a diagonal matrix $\mathbf{D}$ with entries

$$d_i = \sum_{j=1}^{N} a_{ij}. \tag{10}$$

We, then, calculate the symmetric graph Laplacian by

$$\mathbf{L} = \mathbf{D}^{-1/2}(\mathbf{D} - \mathbf{A})\mathbf{D}^{-1/2}. \tag{11}$$

The spectral clustering algorithm [37] identifies clusters' centers in the space of the graph Laplacian's eigenvectors, as detailed in the Appendix. However, the number of clusters should be provided, as the final stage of [37] relies on k-means clustering.

We estimate the appropriate number of clusters as a tradeoff between the accuracy of the model's prediction and its complexity, which is the underlying principle in model-order selection methods, such as Akaike's information criteria [38, Ch. 7]. Here, for reasons of model interoperability, we consider the total "cost" of the clustering solution as a tradeoff between the intercluster stability and the total number of clusters. We note that the term $\mathbf{D}^{-1/2}\mathbf{A}\mathbf{D}^{-1/2}$ in (11) is the ratio of the inner to outer connectivity of each pair of nodes, $0 \leq a_{ij}/\sqrt{d_i d_j} \leq 1$. Denote the kth cluster by an $N \times 1$ binary vector $\mathbf{b}_k$, whose entries are a binary indicator that is set to 1 if the corresponding data point belongs to the cluster. We relate to $\mathbf{b}_k^{\mathrm{T}}\mathbf{L}\mathbf{b}_k$ as a cost function for forming nodes i and j in the same kth cluster. The appropriate number of clusters, $\hat{K}$, is determined by minimizing the mean partition cost

$$\hat{K} = \arg\min_{K} \frac{1}{K}\sum_{k=1}^{K} \mathbf{b}_k^{\mathrm{T}}\mathbf{L}\mathbf{b}_k + \epsilon K$$
$$\text{s.t. } b_i \in [0,1],\ \mathbf{b}_k^{\mathrm{T}}\mathbf{b}_l = 0 \quad \forall k \neq l \tag{12}$$

where ϵK is a penalty for forming "too many" clusters. This model-order selection process is illustrated in Fig. 2(d), while the formed clusters are shown in Fig. 2(e).

C. Classify Clusters

The classification of the formed clusters to targets of interest, illustrated in Fig. 2(f), is based on the statistics of the data points within the identified cluster. The selection of features for classification is application dependent and should reflect the difference between the desired target and clutter. Recall that for the task of distinguishing between reflections from turtles and clutter, we expect the following.

1) Reflections from a sea turtle will form clusters with relatively high connectivity, while clutter-related data points will form clusters that are less connected.
2) Reflections from a sea turtle will exhibit a diverse frequency spectrum, due to interference from the turtle's internal organs, while clutter-related reflections will have a more uniform frequency spectrum.

To quantify how well a cluster is connected, we define the *connectivity* of a cluster as the sum of all connections (8) within the cluster. For the kth cluster, the connectivity is calculated by

$$c_k = \sum_{i,j \in \text{cluster } k} w_{ij} \tag{13}$$

or, in a matrix form,

$$c_k = \mathbf{b}_k^{\mathrm{T}}\mathbf{W}\mathbf{b}_k \tag{14}$$

where $\mathbf{W}$ holds the weights w_{ij} from (8). To characterize the cluster's frequency diversity, we rank the entropy of the spectrum of the cluster's members. Here, a low entropy reflects high-frequency diversity. This information is obtained by performing the median

$$\bar{H}_k = \text{median}\left[\{H_i\left(\mathbf{z}^i\right)\}_{i \in \text{cluster } k}\right]. \tag{15}$$

For our application, since stability and spectral diversity are not necessarily related, we combine the two above features by

$$\text{class}\left(\mathbf{b}_k\right) = \begin{cases} \text{target}, & (c_k > \eta_c) \wedge \left(\bar{H}_k < \eta_h\right) \\ \text{clutter}, & \text{otherwise} \end{cases} \tag{16}$$

where thresholds η_c and η_h are user defined. A suggested approach to calculate these thresholds is by the distribution of features of the clutter-related data points. When the SCR is low, and most data points are clutter related, we expect the target's features to be anomalously relative to the clutter. Since both connectivity and entropy are precalculated for all the data points, we can determine the thresholds based on, e.g., the empirical cumulative distribution functions of these features. Alternatively, when the SCR is high, a large fraction of the data points is target related, and we cannot rely on anomaly detection. In such cases, the clutter's entropy distribution can be estimated from a sample of the MF outputs that did not pass the threshold. Still, to calculate connectivity, we need to apply the clustering algorithm. Consequently, we recommend learning the connectivity of the clutter offline by using recordings from diverse sea experiments.

D. Discussion

The parameters of our above algorithm are α, β and τ, which control the relative importance of the frequency diversity, range difference, and ping difference, respectively. The role of these parameters is to penalize the weights of the graph's edges when pairs of data points do not fit the model of a maneuvering target. Setting the value too low may give low impact to the model assumptions during the clustering, while setting a value that is too high may decrease the weights, rendering the clustering ineffective. We advise setting the above values such that $\alpha, \beta\|\Delta r_{ij}\|_\infty$ and $\tau\|\Delta t_{ij}\|_\infty$ are close to 1 and then doing some fine-tuning according to the experimental setup, e.g., by clustering together reflections from large objects, such as a ship hull or the seafloor.

Considering features for cluster classification, we note that, while the features we use are suitable for our application, one may consider additional features, such as the cluster's size and spatial extent. Filtering out seemingly nonrelated data points from clusters is also possible. However, a main source of complexity is the computation of the mutual information, which involves calculating the joint entropy of all the pairs of data points. A suggestion to reduce this bottleneck is to split the computation of (7), such that (5) will not be calculated for $\alpha << \beta\Delta r_{ij}\mathrm{u}(\Delta r_{ij} - l) + \tau\Delta t_{ij}$, where it has a negligible contribution. Finally, we note that our solution allows for using an alternative preliminary detection method than the MF method which we presented. Examples are given in [8], [9], [13], [15], [18], and [19].

V. Performance Analysis

The task we explore can be considered a mixture of detection and classification. *Detection* is considered when we wish to identify the existence/absence of a target. We measure detection per scenario by the tradeoff between the probability of detection,

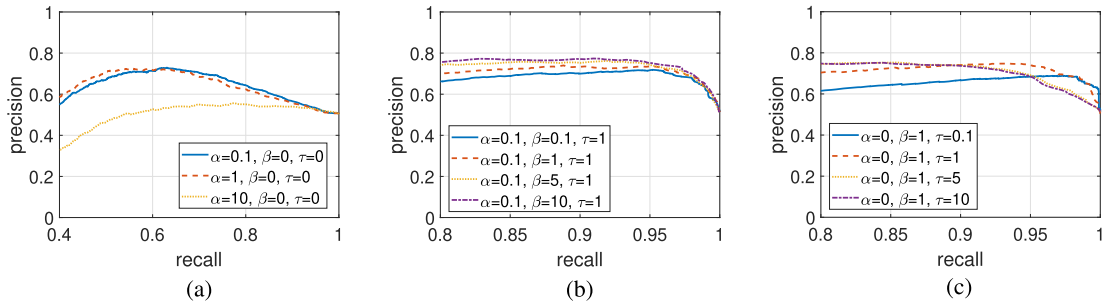

Fig. 3. Precision–recall curves for exploring sensitivity to parameters, while changing the connectivity threshold. SCR $= -9$ dB. (a) Sensitivity to α. (b) Sensitivity to β. (c) Sensitivity to τ.

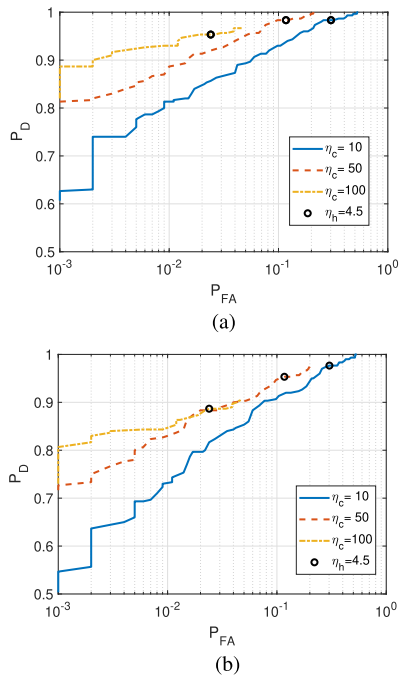

Fig. 4. ROC where $\eta_{\rm h}$ is parameter, and $(\alpha, \beta, \tau) = (0.1, 1, 1)$. Percentage of valid pings is (a) 100% and (b) 70%. SCR $= -9$ dB.

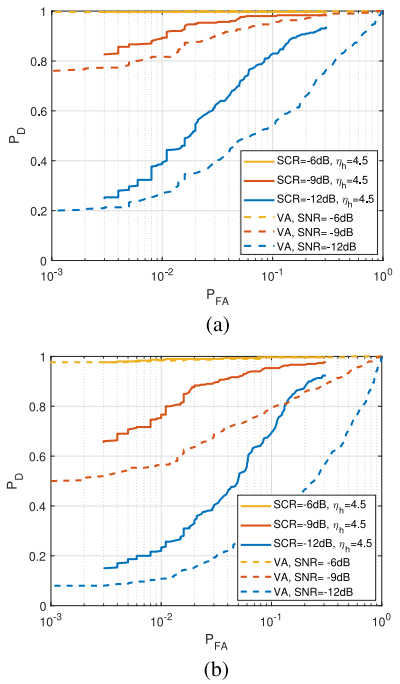

Fig. 5. ROC results when (a) all emissions include a valid reflection and (b) 70% of the emissions are valid. $(\alpha, \beta, \tau) = (0.1, 1, 1)$.

$P_{\rm D}$, and the probability of false alarm, $P_{\rm FA}$. The *classification* of individual data points to either clutter or a target is considered a generalization of detection to account for a soft threshold to the indicated existence/absence of a target. We measure the performance of our classifier per data point and evaluate it by the precision and recall tradeoff. As a benchmark, we use the VA-based detection method [12], which performs TkBD to find a target's path within a clutter by probabilistic analysis. We define a *true detection* as at least one cluster per the emission of N_p pings that is correctly classified as a target. Similarly, we define a *false alarm* as at least one cluster per the emission of N_p pings that is falsely classified as a target. Formally, $P_{\rm D}$ and $P_{\rm FA}$ are calculated by

$$P_{\rm D} = \frac{N_{\rm TP}}{N_{\rm target}} \tag{17a}$$

$$P_{\rm FA} = \frac{N_{\rm FP}}{N_{\text{no-target}}} \tag{17b}$$

where $N_{\rm TP}$ is the number of scenarios with targets that are detected, $N_{\rm target}$ is the total number of simulated scenarios with targets, $N_{\rm FP}$ is the number of scenarios without target but detection is determined, and $n_{\rm no-target}$ is the total number of simulated scenarios without targets.

To evaluate classification, we define *precision* by

$$\text{precision} = \frac{n_{\rm TP}}{n_{\rm TP} + n_{\rm FP}} \tag{18}$$

where $n_{\rm TP}$ is the number of data points correctly classified as a target, and $n_{\rm FP}$ is the number of data points falsely classified as a target. The recall is defined by

$$\text{recall} = \frac{n_{\rm TP}}{n_{\rm TP} + n_{\rm FN}} \tag{19}$$

Fig. 6. Sea turtles that were released, and one of the release activities.

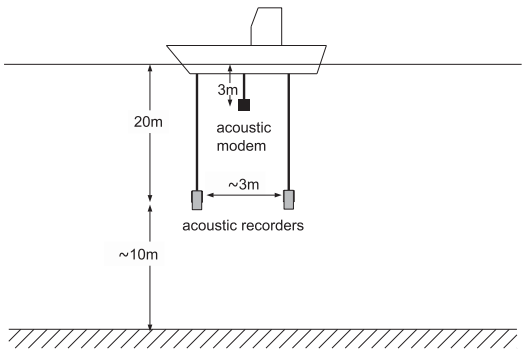

Fig. 7. Sea experiment setup.

where n_{FN} is the number of data points that are indeed a target, but are falsely classified as a nontarget. The ground truth for classification is obtained according to the distance of a determined valid data point to the simulated path of the target, such that

$$\text{ground truth class for } (r_i, p_i) = \begin{cases} \text{target}, & |r_i - \rho(p_i)| < \rho_0 \\ \text{clutter}, & |r_i - \rho(p_i)| \geq \rho_0 \end{cases} \tag{20}$$

where $\rho(p_i)$ is the simulated target's path at ping p_i, and $\rho_0 = 0.5$ m.

A. Numerical Simulations

1) Simulation Setup: The simulation facilitates acoustic reflections from a maneuvering target. We use a chirp signal, $s(t)$, of a $T_{\mathrm{s}} = 10$ ms duration and a frequency band of 7–17 kHz. A single simulation scenario is determined as a set for $N_{\mathrm{p}} = 20$ pings. The time segment duration of data for each ping is 90 ms, corresponding to a two-way path range of ∼70 m. For each scenario, the target's path is simulated by a random process, $\rho_{m+1} = \rho_m + n_m$, where

$$n_m \sim \mathrm{N}(0, \sigma_{\mathrm{n}} = 2\,\mathrm{m}), m = 1, \ldots N_{\mathrm{p}}.$$

We simulate the acoustic signal associated with ping m, $y^m_{\mathrm{sim}}(t)$, by

$$y^m_{\mathrm{sim}}(t) = \begin{cases} s(t - T_m) * h(t - T_m), & T_m \leq t \leq T_m + T_{\mathrm{s}} \\ y_{\mathrm{clutter}}(t), & \text{otherwise} \end{cases} \tag{21}$$

where $T_m = 2\rho_m/c$, $y_{\mathrm{clutter}}(t)$ are clutter samples, and $h(t)$ is the simulated target's impulse response. We consider the case where the turtle's impulse response distorts the transmitted signal $s(t)$ in a manner which we cannot predict. Therefore, we simulate $h(t)$ by a multivariate random variable, $\mathbf{h}_{100\times1}$, that is uniformly distributed in $[-1\ 1]$ and then normalized such that $\mathbf{h}^{\mathrm{T}}\mathbf{h} = 1$. The same impulse response is used for all pings within a single simulation scenario. To explore performance over a realistic scenario, clutter samples $y_{\mathrm{clutter}}(t)$ are taken from sea experiments. In these experiments, we deployed our sonar to a depth of roughly 20 m and transmitted the same chirp signals used in the simulation. The bottom depth was 125 m, and a manual inspection ensured that the reflection pattern contained only clutter. The overall duration of obtained clutter samples is 257 s, from which we randomly uniformly choose in each simulation run. The desired SCR is achieved by controlling the amplitude of the simulated target's echo. No ambient noise is added. We run 300 simulation scenarios per SCR value including target, and 1000 scenarios to explore false alarms.

2) Simulation Results: We start by exploring the sensitivity of our algorithm for parameters α, β, and τ. The classification sensitivity for α is presented in Fig. 3(a), while setting $(\beta, \tau) = 0$ and changing the threshold η_{c}. The observable maxima in the

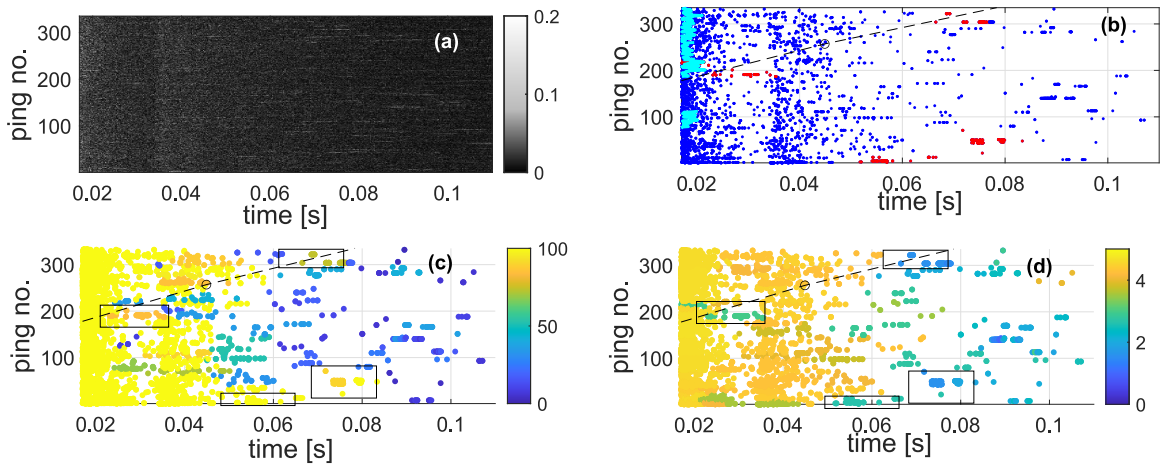

Fig. 8. Results for transmission signals in the range 7–17 kHz. (a) Intensity image of the MF's output. (b) Detections after threshold (our target indications are in red, VA detected path in cyan). (c) Connectivity of clusters. (d) Median entropy of clusters. Target indications are marked by rectangles in (c) and (d).

precision–recall curve indicates that there is a connectivity threshold $\eta_{\rm c}$ under which detection degrades drastically. From this figure, it is evident that the frequency diversity alone, without using the spatial and temporal information, contains sufficient information. We also observe that classification performance is sensitive to α. This is because the exponent in (8) is multiplied by α, and thus, any change in α has an impact on weights w_{ij}. A similar trend is observed in Fig. 3(b), which shows the classification results when only β varies, i.e., exploring the impact of the spatial feature in the reflections on classification. Here, we set $\alpha = 0$, but $\tau = 1$, since we need the temporal connection to form clusters that are geometrically valid. We observe that both precision and recall improve with β, but converge for $\beta > 5$. Next, the sensitivity for the temporal feature in the reflections is explored by changing parameter τ, keeping $\beta = 1$ and $\alpha = 0$. The results in Fig. 3(c) show a tradeoff: a lower τ increases the recall at the cost of precision, i.e., the algorithm tends to classify too many instances as a "target." On the other hand, a higher τ reduces both false positives and true positives. It seems that a favorable tradeoff is $\tau = 1$. In the following, we pick the values $\alpha = 0.1$, $\beta = 1$, and $\tau = 1$.

The effect of the entropy threshold is explored in Fig. 4, where $\eta_{\rm h}$ is used as a parameter, with $2 \leq \eta_{\rm h} \leq 5$, while three values for $\eta_{\rm c}$ are tested. Here, the SCR is -9 dB. Fig. 4(a) shows results for the case where the reflections of each of the $N_{\rm p} = 20$ emissions include a valid target in addition to clutter. In contrast, Fig. 4(b) shows the case where only 70% of the 20 emissions include a valid reflection, such that the target appears as "blobs" of sporadic detections. We observe that the performance of our algorithm degrades in this latter case, since fewer emissions are available for detection. In both cases, it is evident that higher $\eta_{\rm c}$ produces better ROCs, but the maximal $P_{\rm D}$ that is reached is limited. This is because target-related clusters with lower connectivity are misclassified as clutter. The circles markers on the plots are for $\eta_{\rm h} = 4.5$, which we identify as a reasonable tradeoff that keeps $P_{\rm D}$ relatively constant for a large range of $\eta_{\rm c}$ values.

Fig. 5 shows the impact of the SCR level on the detection performance, and the two cases of 100% and 70% valid targets in the 20 emissions are presented in Fig. 5(a) and (b), respectively. Here, we set the entropy threshold to be $\eta_{\rm h} = 4.5$. Additional fixed parameters are $\eta_{\rm MF} = 5 \times 10^{-6}$, $(\alpha, \beta, \tau) = (0.1, 1, 1)$. The TkBD VA benchmark is marked by a dashed line. As expected, both the algorithms perform well for high SCR $= -6$ dB, while performance degrades rapidly as the SCR decreases. It is evident that, for low SCR, our algorithm outperforms the benchmark. In addition, in the case of 70% valid pings, the performance of the VA-based approach highly degrades, while our method, although affected, is more robust. This is because our method is designed to look for 'blobs' of detection, while the VA-based method assumes that the reflections form a path.

B. Field Experiments

The above simulation results explored the performance statistically, but lack a reliable model for the acoustic reflections from the sea turtle. To this end, we conducted a designated sea experiment that involved the release of two rehabilitated sea turtles and measured their acoustic reflections. The experiment was carried out roughly 3 km off the coast of Haifa, Israel, in 30 m of water. The sea turtles' released weight was roughly 40 kg, and their size was roughly 60 cm in diameter. These were mature turtles that had suffered a shock wave and were rehabilitated in the Israeli Rescue Center for Sea Turtles. The release was performed by the center's personnel, under the approval of the Israeli Nature and Park Authority. An image of the release of one of the turtles is shown in Fig. 6.

C. Experimental Setup

The experimental setup is illustrated in Fig. 7. To transmit the chirp signals, we used the Evologics S2CR 7/17 and S2C M 48/78 software-defined acoustic modems to emit trains of chirp signals of 10-ms duration at a PRI of 0.7 s and at frequency ranges of 7–17 kHz and 48–78 kHz, respectively. The

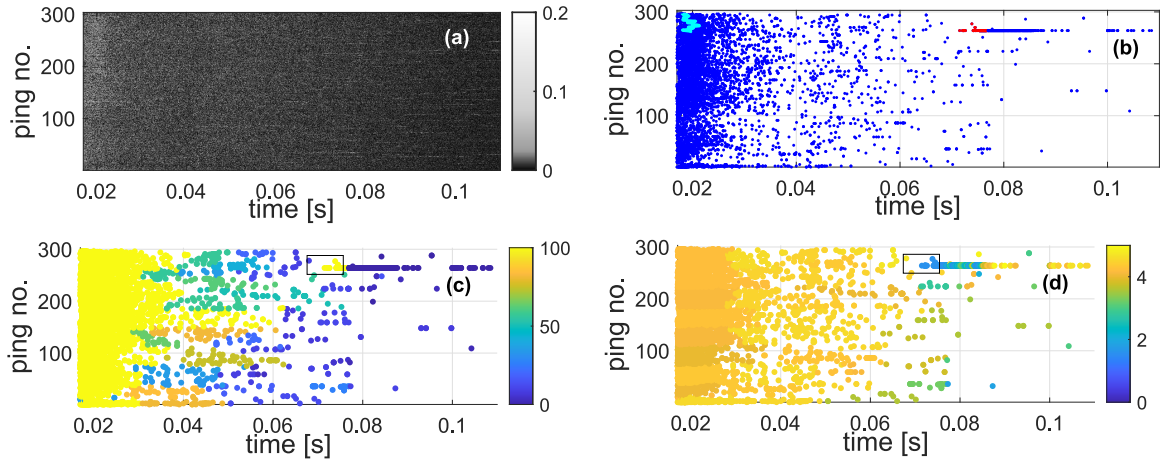

Fig. 9. Results for transmission signals in the range 48–78 kHz. (a) Intensity image of the MF's output. (b) Detections after threshold (our target indications are in red, VA detected path in cyan). (c) Connectivity of clusters. (d) Median entropy of clusters. Target indications are marked by rectangles in (c) and (d).

emission power was 170 dB re. 1μP@1 m. The modems were deployed from the boat to a depth of 3 m. Two omnidirectional self-made receivers were deployed from the boat to a depth of 20 m. Continuous recording took place at 192k samples per second with 3-Byte resolution per sample. The modems started emitting the signals before the release of the sea turtles. During these events, the turtles mostly swam on the surface, but also submerged from time to time. We visually observed their behavior from the vessel and documented the rise and submerge times.

D. Experimental Results

For analysis, we considered $v_{\max} = 2$ m/s, $l = 0.3$ m, and the parameter set $(\alpha, \beta, \tau) = (0.1, 1, 1)$. We applied our algorithm on patches of 37 pings with a duration of 60 ms each. Fig. 8 shows results for the frequency range of 7–17 kHz. In this experiment, there was a plastic buoy in the water with a GPS receiver at a range of no more than 20 m from the sea turtle. The path of this GPS is marked by a dashed line. Panel (a) shows the data-sample matrix after the MF processing, but before the thresholding operation. Here, we merge the reflections from all the emissions. We observe that the main reflection comes from the bottom at roughly 35 ms. The detection indications after thresholding are shown in Panel (b), and detections that are classified as targets are marked in red. The path detected by the VA benchmark is plotted in cyan, using a threshold value of 94. Panels (c) and (d) show the connectivity and the median entropy of the clusters, respectively, and clusters classified as targets are marked by rectangles. We observe two target indications along the path of the GPS: one at emission 191 that is $\sim$7 m from the path and the other at emission 304 that is right on the path. The range to these targets, measured on the surface, is 28.7 ± 2 and 60.2 ± 4.8 m, respectively. In addition, there are two clusters that were falsely classified as targets on emissions 5 and 44 at ranges of 47.1 ± 4.2 and 61.5 ± 1.5 m, respectively. We see that the benchmark VA detects intense reflections at the leftmost part of the data-sample matrix and follows them instead of the assumed turtle. This is because the VA uses only intensities as an input and is not able to detect more than a single target per patch. Instead, our method has the following advantages: it uses spectral information, and multiple targets can be detected.

Results obtained for the 48–78 kHz frequency range are shown in Fig. 9. In this experiment, there was no GPS buoy, and we rely on the experiment's log to identify the events. The sea turtle was released around ping 108 and was swimming on the surface, thus making its detection very challenging. Approximately 90 s after its release, at emission ping 236, the turtle was observed taking air, indicating that it was about to dive. This arguable dive enabled its detection, and indeed, we observed a detection that matches a low-entropy cluster around emission 264 with a time delay of $\sim$ 74 ms, corresponding to a range of 61.9 ± 2.1 m. In this case, no false alarms were observed. Moreover, in this case, the VA benchmark, applied with a threshold of 147, is blinded by high-intensity reflections and cannot detect the turtle.

VI. Conclusion

In this article, we proposed an unsupervised remote sensing algorithm to identify a sea turtle within a point cloud of sonar's reflections under low SCR conditions. We designed a clustering algorithm, which is based on the geometrical relations between a sample within the point cloud that relates to the target's reflections, as well as on the expected spectral diversity of the turtle's reflections. We also utilized the expected stability of target-related reflections as another means of detection verification. We explored the performance of our algorithm in both simulations and in a designated sea experiment, where we released two rehabilitated sea turtles and tracked their trajectory acoustically. Compared to a TkBD benchmark, our results show an improved robustness to fluctuating intensity and a better tradeoff between the false alarm and detection rates. Future research may improve classification by adding more features and exploring a multivariate feature combination.

APPENDIX

DETAILS OF THE SPECTRAL CLUSTERING ALGORITHM

For the sake of completeness, we now detail our implementation to the algorithm in [37].

1) Define the matrix $\mathbf{A}$ according to (9).
2) Define the diagonal matrix $\mathbf{D}$ according to (10).
3) Calculate the symmetric graph Laplacian by (11).
4) Assuming that the number of clusters is K, calculate the K eigenvectors of $\mathbf{L}$ corresponding to the K smallest eigenvalues (this is because we use (11) instead of $\mathbf{D}^{-1/2}\mathbf{A}\mathbf{D}^{-1/2}$ used in [37]), stack them in the columns of a matrix $\mathbf{X}_{N\times K}$, and normalize its rows to get matrix $\mathbf{Y}$ with entries

$$Y_{ij} = \frac{X_{ij}}{(\sum_{j=1}^{N} X_{ij}^2)^{1/2}}. \tag{22}$$

5) Refer to each row in $\mathbf{Y}$ as a point in $\mathbb{R}^K$, classify by K-means to one of K clusters, and assign the cluster index associated with this row to the corresponding data point.

REFERENCES

[1] S. E. Nelms, W. E. Piniak, C. R. Weir, and B. J. Godley, "Seismic surveys and marine turtles: An underestimated global threat," *Biol. Conservation*, vol. 193, pp. 49–65, 2016.

[2] Y. Levy et al., "A small fishery with a high impact on sea turtle populations in the eastern Mediterranean," *Zool. Middle East*, vol. 61, no. 4, pp. 300–317, 2015.

[3] B. D. MacDonald, R. L. Lewison, S. V. Madrak, J. A. Seminoff, and T. Eguchi, "Home ranges of East Pacific Green turtles Chelonia mydas in a highly urbanized temperate foraging ground," *Mar. Ecol. Prog. Ser.*, vol. 461, pp. 211–221, 2012.

[4] A. B. Bolten, *Seasonal Abundance, Size Distribution, and Blood Biochemical Values of Loggerheads (Caretta Caretta) in Port Canaveral Ship Channel, Florida*, vol. 353. Miami, FL, USA: US Department of Commerce, National Oceanographic and Atmospheric Administration, National Marine Fisheries Service, Southeast Fisheries Science Center, 1994.

[5] A. Mahfurdz, A. Mahfurdz, and S. Saifullah, "Green turtle and fish identification based on acoustic target strength," *Int. J. Adv. Intell. Informat.*, vol. 4, no. 1, pp. 53–62, 2018.

[6] I. Pérez-Arjona et al., "TS measurements and simulations of Mediterranean sea turtles," in *Proc. 38th Scand. Symp. Phys. Acoust.*, 2013, pp. 1–6.

[7] B. A. Jones, T. K. Stanton, J. A. Colosi, R. C. Gauss, J. M. Fialkowski, and J. M. Jech, "Broadband classification and statistics of echoes from aggregations of fish measured by long-range, mid-frequency sonar," *J. Acoust. Soc. Amer.*, vol. 141, no. 6, pp. 4354–4371, 2017.

[8] D. A. Abraham and P. K. Willett, "Active sonar detection in shallow water using the page test," *IEEE J. Ocean. Eng.*, vol. 27, no. 1, pp. 35–46, Jan. 2002.

[9] H. Berg and K. T. Hjelmervik, "Classification of anti-submarine warfare sonar targets using a deep neural network," in *Proc. IEEE/MTS OCEANS Conf.*, Charleston, SC, USA, 2018, pp. 1–5.

[10] K. T. Hjelmervik, K. T. Hjelmervik, D. H. S. Stender, and H. Berg, "Sonar scattering from the sea bottom near the Norwegian coast," in *Proc. IEEE OCEANS Conf.*, Anchorage, AK, USA, 2017, pp. 1–5.

[11] S. J. Davey, M. G. Rutten, and B. Cheung, "A comparison of detection performance for several track-before-detect algorithms," *EURASIP J. Adv. Signal Process.*, vol. 2008, pp. 1–10, 2007.

[12] R. Diamant, D. Kipnis, E. Bigal, A. Scheinin, D. Tchernov, and A. Pinchasi, "An active acoustic track-before-detect approach for finding underwater mobile targets," *IEEE J. Sel. Topics Signal Process.*, vol. 13, no. 1, pp. 104–119, Jan. 2019.

[13] G. De Magistris et al., "Automatic object classification for low-frequency active sonar using convolutional neural networks," in *Proc. IEEE/MTS OCEANS Conf.*, Seattle, WA, USA, 2019, pp. 1–6.

[14] R. Diamant, "Closed form analysis of the normalized matched filter with a test case for detection of underwater acoustic signals," *IEEE Access*, vol. 4, pp. 8225–8235, 2016.

[15] I. Seo, S. Kim, Y. Ryu, J. Park, and D. S. Han, "Underwater moving target classification using multilayer processing of active sonar system," *Appl. Sci.*, vol. 9, no. 21, 2019, Art. no. 4617.

[16] D. M. Cowell and S. Freear, "Separation of overlapping linear frequency modulated (LFM) signals using the fractional fourier transform," *IEEE Trans. Ultrasonics Ferroelectr. Freq. Control*, vol. 57, no. 10, pp. 2324–2333, Oct. 2010.

[17] R. Jacob, T. Thomas, and A. Unnikrishnan, "Applications of fractional fourier transform in sonar signal processing," *IETE J. Res.*, vol. 55, no. 1, pp. 16–27, Jan. 2009.

[18] G. Yu, S.-C. Piao, and X. Han, "Fractional Fourier transform-based detection and delay time estimation of moving target in strong reverberation environment," *IET Radar, Sonar Navigat.*, vol. 11, no. 9, pp. 1367–1372, 2017.

[19] P. Stinco, G. De Magistris, A. Tesei, and K. D. LePage, "Automatic object classification with active sonar using unsupervised anomaly detection," in *Proc. IEEE 28th Eur. Signal Process. Conf.*, 2021, pp. 46–50.

[20] Y. Barniv, "Dynamic programming solution for detecting dim moving targets," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES- 21, no. 1, pp. 144–156, Jan. 1985.

[21] M. Wei, B. Shi, C. Hao, and S. Yan, "A novel weak target detection strategy for moving active sonar," in *Proc. IEEE/MTS OCEANS Kobe Techno-Oceans Conf.*, 2018, pp. 1–6.

[22] C. Jing, Z. Lin, and J. Li, "Detection and tracking of an underwater target using the combination of a particle filter and track-before-detect," in *Proc. IEEE OCEANS Conf.*, Shanghari, China, 2016, pp. 1–5.

[23] T. Northardt and S. C. Nardone, "Track-before-detect bearings-only localization performance in complex passive sonar scenarios: A case study," *IEEE J. Ocean. Eng.*, vol. 44, no. 2, pp. 482–491, Feb. 2018.

[24] R. L. Streit and T. E. Luginbuhl, "Maximum likelihood method for probabilistic multihypothesis tracking," in *Proc. Signal Data Process. Small Targets Int. Soc. Opt. Photon.*, 1994, pp. 394–405.

[25] R. L. Streit, M. L. Graham, and M. J. Walsh, "Multitarget tracking of distributed targets using histogram-PMHT," *Digit. Signal Process.*, vol. 12, no. 2/3, pp. 394–404, 2002.

[26] H. Vu, S. Davey, F. Fetcher, S. Arulampalam, R. Ellem, and C. Lim, "Track-before-detect for an active towed array sonar," in *Proc. Acoust. Conf.*, Victor Harbor, SA, Australia, 2013, pp. 1–7.

[27] H. X. Gaetjens, S. J. Davey, S. Arulampalam, F. K. Fletcher, and C.-C. Lim, "Histogram-PMHT for fluctuating target models," *IET Radar Sonar Navigat.*, vol. 11, no. 8, pp. 1292–1301, 2017.

[28] C. Jauffret and Y. Bar-Shalom, "Track formation with bearing and frequency measurements in clutter," in *Proc. IEEE 29th Conf. Decis. Control*, 1990, pp. 3335–3336.

[29] P. Willett and S. Coraluppi, "Application of the MLPDA to bistatic sonar," in *Proc. IEEE Aerosp. Conf.*, 2005, pp. 2063–2073.

[30] W. Blanding, P. Willett, and S. Coraluppi, "Sequential ML for multistatic sonar tracking," in *Proc. IEEE OCEANS Conf.- Europe*, 2007, pp. 1–6.

[31] S. Schoenecker, P. Willett, and Y. Bar-Shalom, "ML–PDA and ML–PMHT: Comparing multistatic sonar trackers for VLO targets using a new multitarget implementation," *IEEE J. Ocean. Eng.*, vol. 39, no. 2, pp. 303–317, Feb. 2013.

[32] R. J. Korneliussen and E. Ona, "Synthetic echograms generated from the relative frequency response," *ICES J. Mar. Sci.*, vol. 60, no. 3, pp. 636–640, 2003.

[33] T. K. Stanton, D. Chu, J. M. Jech, and J. D. Irish, "New broadband methods for resonance classification and high-resolution imagery of fish with swimbladders using a modified commercial broadband echosounder," *ICES J. Mar. Sci.*, vol. 67, no. 2, pp. 365–378, 2010.

[34] Y. Pailhas, C. Capus, K. Brown, and P. Moore, "Analysis and classification of broadband echoes using bio-inspired dolphin pulses," *J. Acoust. Soc. Amer.*, vol. 127, no. 6, pp. 3809–3820, 2010.

[35] R. H. Love, "Measurements of fish target strength: A review," *Fish. Bull*, vol. 69, no. 4, pp. 703–715, 1971.

[36] N. X. Vinh, J. Epps, and J. Bailey, "Information theoretic measures for clusterings comparison: Variants, properties, normalization and correction for chance," *J. Mach. Learn. Res.*, vol. 11, pp. 2837–2854, 2010.

[37] A. Y. Ng, M. I. Jordan, and Y. Weiss, "On spectral clustering: Analysis and an algorithm," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2002, pp. 849–856.

[38] T. Hastie, R. Tibshirani, J. H. Friedman, and J. H. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, vol. 2. Berlin, Germany: Springer, 2009.

Dror Kipnis received the B.Sc. degree in electrical engineering and the B.A. degree in physics from the Technion—Israel Institute of Technology, Haifa, Israel, in 2003, the M.Sc. degree in electrical engineering from the Department of Electrical Engineering—Physical Electronics, Tel Aviv University, Tel Aviv, Israel, in 2007, and the Ph.D. degree in marine technologies from the Department of Marine Technologies, University of Haifa, Haifa, in 2022.

He has more than 18 years of experience developing algorithms, in the domains of signal processing and machine learning. His current research interests include interference tolerant approaches for acoustic detection of marine fauna.

Yaniv Levy received the B.Sc. degree in environmental studies from Hebrew University, Jerusalem, Israel, in 2000, the M.Sc. degree in zoology from Tel Aviv University, Tel Aviv, Israel, in 2005, and the Ph.D. degree in marine biology from the University of Haifa, Haifa, Israel, in 2018.

He is currently a Research Fellow with the Department of Marine Biology, Leon H. Charney School of Marine Sciences, University of Haifa. He is the Founder and Manager of Israel's Sea Turtle Rescue Center, Nature and Parks Authority, Beit Yannai, Israel. He established the place in 1999, and since then, his work has consisted of conservation, biology, movement ecology, animal welfare, veterinary, and genetics.

Roee Diamant (Senior Member, IEEE) received the B.Sc. and M.Sc. degrees from the Technion—Israel Institute of Technology, Haifa, Israel, in 2002 and 2007, respectively, and the Ph.D. degree from the Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, BC, Canada, in 2013, all in electrical engineering.

From 2001 to 2009, he was a Project Manager and System Engineer with Rafael Advanced Defense Systems, Haifa, where he developed a commercial underwater modem with network capabilities. In 2015 and 2016, he was a Visiting Professor with the University of Padua, Padua, Italy. He is the coordinator of the EU H2020 project SYMBIOSIS (BG-14 track) and leads the Underwater Acoustic and Navigation Laboratory as an Associate Professor with the Department of Marine Technologies, University of Haifa, Haifa. His research interests include underwater acoustic communication, underwater localization and navigation, object detection and classification, and sonar signal processing.

Dr. Diamant received the Israel Excellent Worker First Place Award from the Israeli Presidential Institute in 2009 and the NSERC Vanier Canada Graduate Scholarship in 2010. He has received three Best Paper awards. He is an Associate Editor for IEEE JOURNAL OF OCEAN ENGINEERING.