Validation of Localization via Wideband Acoustic Arrays for Underwater Fauna Monitoring at Sea

Elizaveta Dubrovinskaya[‡], Amer Shaddad[§], Shlomo Dahan[§], Roee Diamant[§], and Paolo Casari[♯]

[‡]IMDEA Networks Institute and University Carlos III of Madrid, Spain

[§]Department of Marine Technology, University of Haifa, Israel

[♯]Department of Information Engineering and Computer Science, University of Trento, Italy

***Abstract*—We present a technique for simultaneous detection, path tracking and accurate 3D underwater localization using wideband arrays of complex geometry based on acoustic reflections clustering. We have extended our previously proposed algorithm for 3D localization designed for arrays that do not meet typical constraints of one half-wavelength spacing between the closest array elements. Inspired by the scope of SYMBIOSIS, a hybrid opto-acoustic system for pelagic fish species monitoring, we added more functionalities that fit the needs of the project.**

The proposed algorithm can automatically discriminate moving targets from stationary environmental features and track them to estimate their possible time of arrival to the system. We test the algorithm in several autonomous deployments including shallow and deep water. The experimental results for marine fauna monitoring have shown a good performance in various environments.

I. Introduction and related work

The use of acoustic arrays for accurate localization and tracking is typical in underwater scenarios [1]–[3], and is the basis for common sonar detection, positioning, and telemetry methods. Localization in 3D with a single emitting array is algorithmically and computationally challenging [4]: however, this may allow easier low-cost deployments. For these applications, hydrophone arrays are typically engineered specifically to provide the desired spatial scanning capabilities. Yet, once an array structure is available, it is often convenient to exploit the same array for different applications and scenarios, and even for a different frequency band, provided that the hydrophones support it. While signal design is often inspired by the application at hand, changing the operational frequencies of the array implies that the spacing of the elements may exceed the optimal $\lambda/2$ spacing, where λ is the wavelength corresponding to the maximum working frequency of the array. This would lead to significant ambiguity in the output of array processing [5].

Another convenient option when repurposing arrays is to join multiple subarrays into a single, larger array, yielding better noise rejection and estimation accuracy capabilities. Yet, physical elements such as support structures, casings, and connectors may constrain the mounting of subarrays and prevent optimal array topologies.

In our previous work [6], [7], we presented a technique to achieve reliable 3D direction of arrival (DoA) estimates using wideband arrays of arbitrary shape. An opportunity to test the algorithm in the wild came in the context of the experimental campaigns carried of the EU H2020 SYMBIOSIS project [8]. This international effort targets the long-term, single-location monitoring of underwater environments, with the objective of detecting and measuring the presence, density, and variety of marine fauna species.

In this paper, we present the results of several sea experiments that validated our DoA-based localization algorithm. These experiments were organized throughout the duration of the SYMBIOSIS project: an early-stage experiment is described in [7]; subsequent experiments carried out in the context of real longer-term deployments of the whole SYMBIOSIS platform are described in this work. In particular, we present a relevant application of our algorithm to the localization and tracking of pelagic fish. Unlike the preliminary version in [7], having the algorithm run unattended in an offshore platform for several days required us to automate many sub-procedures. This includes the discrimination between stationary and moving targets, as well as novel approaches for target path tracking. Specifically, inspired by the broad application of clustering in various localization applications including radar [9] and sonar [10], [11], in this work we suggest clustering-based approaches both for target detection and path tracking, and validate the effectiveness of these schemes.

The remainder of this paper is organized as follows. We will start with the key extensions of previously proposed localization algorithm in Section II. Then we will continue with a description of SYMBIOSIS project experiments in Section III. Finally, we will discuss the results in Section IV.

II. Extensions to the localization algorithm

A. *Summary of previous contribution*

Wideband signals are commonly used for various underwater localization applications, including bio-inspired sonars [12] or vocalizing marine animals localization [13]. Often, the direct application of efficient DoA algorithms for geometrically complex arrays using a wideband signal can be challenging, because the arrays do not meet the conditions of an algorithm, e.g., the distance between the array elements should be less than half the minimum signal wavelength. In our previous work [7], we proposed an algorithm that uses both phase shift information as in classical DoA algorithms and signal arrival times. Combining the unsupervised clustering algorithm density-based spatial clustering of applications with noise

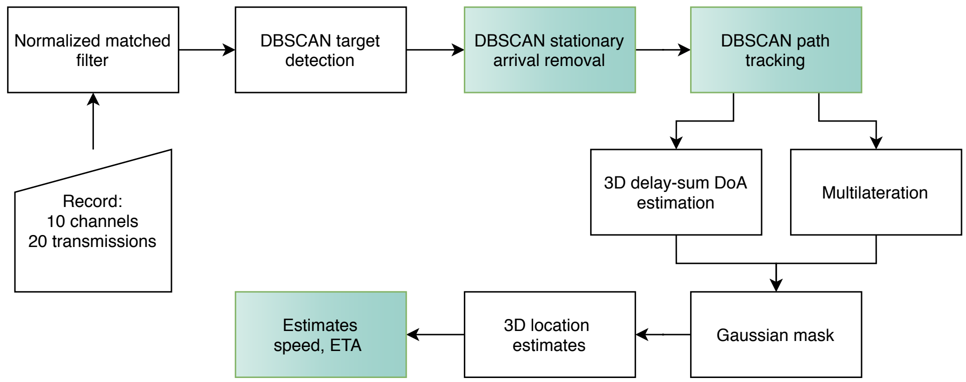

Fig. 1. Extended version of 3D localization algorithm with tracking. The new parts of the algorithm are colored in green.

(DBSCAN) (which yields little if any computational overhead) with classical DoA estimation and multilateration, we were able to achieve simultaneous detection and localization at low signal-to-noise ratio (SNR) levels. We have tested our algorithm on active and passive sonar simulated and emulated data with real clutter, as well as performed an experiment in fresh water Werbellin lake. The results have shown that using the proposed algorithm, an array of opportunity that was assembled without meeting half-wavelength spacing constraints performed as good as a properly spaced array and often outperformed it. However at this stage the algorithm was not able to discriminate targets from stationary environmental features.

B. Main extensions to our algorithm

In the following, we detail how we extended and modified the previously proposed algorithm.

Target detection. With reference to Fig. 1, the basic detection step of the localization algorithm works by clustering peaks from a normalized matched filter (NMF) output aggregated from multiple acoustic channels. Specifically, we collect acoustic samples by transmitting 20 linear frequency-modulated chirps in the 7-17 kHz band, and by listening to the reception of any reflection of such signals off relevant targets or other reflective elements in the surrounding environments. We then collectively analyze a total of 10 channels, where each acoustic channel is the output of a different hydrophone within our array. After processing the NMF outputs through clustering, we record each target detection into a database. Due to the multipath propagation that affects typical underwater acoustic channels, and to the multiple reflections from environmental features, the algorithm may output multiple detections even when there is no fish around the array. These detections correspond to stationary reflections from the water surface, or from parts of the platform itself, and tend to appear stable at the same observation epoch relative to the transmission time of a wideband signal from the array.

Stationary arrival removal. Stationary arrivals are expected to be very stable and consistent over time. In order to remove such detections, we propose a conservative stationary arrival definition: if a detection is present in all of the 20 recordings we perform, and the speed of the moving target is below 0.01 m/s, we consider the object as a stationary target, and remove it from further processing. In order to do so, we represent the target detection information as a three-dimensional matrix whose entries are spanned by the range of detected arrival, the recording time, and the total NMF value of all peaks clustered during the detection step.

Then, we perform another clustering step based on DBSCAN, where we set the algorithm's cluster radius parameter to fit the speed limit. Moreover, we impose that the minimum number of points in each cluster should be equal to the minimum number of records in which the detection is present. Our stationary arrival definition is inspired by several previous experimental results, which yielded an informed estimate of how fast we can expect an actual pelagic fish specimen of interest to move.

Path tracking. After removing stationary arrivals, there still remain multiple detections that do not comply with typical target characteristics. For example, these include spurious reflections from bubbles, or targets moving exceedingly fast, such as motor boats. In order to filter these targets out, and keep only targets whose speed fits a set of desired parameters, we perform an additional clustering step. This is aimed at identifying clusters of points that abide by a chosen maximum speed by being spaced in time no less than a given amount.

Once the target is tracked, we run the localization algorithm. As an output, the algorithm informs whether any target with desired parameters was detected. If this is the case, it calculates the average speed, the 3D coordinates of the target of interest (using the DoA-based approach in [7]) and estimates the time when the target may approach a range of about 10 m from the array, and thus become visible to the underwater cameras mounted on the platform. This estimated time of arrival (ETA) of the target near the platform serves as a trigger to start camera image recording.

C. Preliminary experiment

To validate the clustering-based approach to detection and tracking, we tested the algorithm on an experimental data. In this experiment, we continuously record acoustic data using a single hydrophone while an external transmitter emits chirp

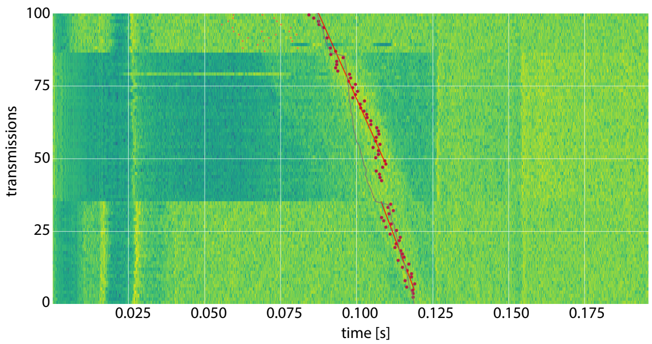

Fig. 2. Diver tracking experiment. Red dots: relevant peaks clustered into a track. Grey line: ground truth from a GPS-equipped surface buoy.

signals in the 7–17 kHz band. This allows us to focus on the stationary target removal, path tracking and target discrimination capabilities of the algorithm, without the additional complexity of fully 3D DoA-based target localization. The experiment took place in shallow north-east Mediterranean waters, with a maximum depth of about 120 m. A diver submerged at a depth of about 10 m depth moved to slowly approach the recorder. The results are shown in Fig. 2. The waterfall matrix in the figure is obtained from the output of the NMF. Here, a higher peak (corresponding to a lighter shade in the figure) represents a higher correlation of the recorded signal with the transmitted signal. Red dots represent relevant clustered arrivals. We observe that these arrivals match a ground-truth track collected from a global positioning system (GPS)-equipped surface-floating buoy the diver was towing while moving towards the platform. Some mismatch such as observed between the 30th and the 70th transmission is expected, as the diver had to periodically stop to pull and relocate the buoy above itself, while swimming at 10 m of depth. After this validation experiment, we concluded that our clustering-based detection algorithm successfully identifies arrivals corresponding to the moving diver, even across multiple recordings in different SNR conditions.

III. Experiments with the SYMBIOSIS platform

A. SYMBIOSIS description

The SYMBIOSIS platform is a hybrid opto-acoustic system: its design revolves around a chain of progressively more accurate detection, localization, tracking and image acquisition steps. The sequence of the main processing steps that the platform periodically carries out is:

1) Coarse detection: the platform emits a 10-ms narrowband signal and processes echoes of this signal from the environment via a neural network, attempting to understand if there exists some target in the surroundings or not.

2) Accurate detection: the platform transmits a sequence of 20 linear chirp signals, in the bandwidth from 7 to 17 kHz, where each chirp signal has a duration of 10 ms, and transmissions are paced every 1.7 s. This phase confirms the presence of a target with greater accuracy than the coarse detection step, using a combination of clustering, neural networks, and signal processing algorithms.

3) Localization and tracking: in parallel to the accurate detection step, we run our localization and tracking algorithm. The overarching objective of this step is to enable the prediction of fish trajectories in the vicinity of the platform. As the algorithm infers that one or more specimens are getting closer to the platform, it signals the platform controller to start image and video acquisitions from the underwater cameras attached to the platform.

4) Optical acquisition and processing: if detections are successful, the system records and processes images and videos: recorded images and frames are input to specifically designed convolutional neural networks, in order to detect the presence of fish specimens, extract bounding boxes that identify the specimens' position within and image, and classify them.

The 3D localization of underwater marine fauna requires to accurately estimate the range and bearing of underwater objects. The SYMBIOSIS platform enables this capability by providing recorded acoustic information from a total of ten hydrophones, grouped into two units of five hydrophones each. Each unit is a commercial off-the-shelf (COTS) USBL pyramidal array having a square base with a side length of 10 cm, and a height of 7.07 cm. The manufacturer of the USBL arrays and SYMBIOSIS project partner EvoLogics GmbH reconfigured each unit to disable pre-programmed USBL functionalities, so that each separate array element can output recorded acoustic samples synchronously.

To cover an area with about 500 m radius around the platform, in each cycle the system records 0.7 s of acoustic data. Additionally, it takes about 1 s to save it and pass it to the algorithm, thus each record cycle takes about 1.7 s. Given the time constraint, working on a very large acoustic dataset covering a long time span would imply that output location information will be likely outdated. However, the modified algorithm requires a sufficiently large number of acoustic records in order to achieve the desired level of accuracy. In order to find a good tradeoff between the above constraints, it was decided to process data in batches of 20 consecutive records.

B. Experiments at the deep-water THEMO mooring

THEMO is a marine observatory that was designed and installed in the eastern Mediterranean sea, in the context of a collaboration between the Texas A&M University, USA, and the University of Haifa, Israel. The observatory consists of two moorings: the shallow THEMO mooring, located in an area with about 120 m of depth, comparatively close to the Israeli coastal zone, and the deep THEMO mooring, located about 50 km away from the shore, where the sea depth reaches about 1500 m. The full description of the observatory can be found in [14]. During the development stage of the SYMBIOSIS project, several experiments were performed near THEMO deployments.

Our experiment took place near the deep THEMO mooring [15]. Here, SYMBIOSIS personnel released a rehabilitated

(a) (b) (c)

Fig. 3. THEMO SYMBIOSIS deployment: a) THEMO marine observatory b) SYMBIOSIS platform with 2 USBL units during the deployment process c) Scheme of SYMBIOSIS's acoustic array with three USBL (courtesy of EvoLogics GmbH).

turtle from a boat. The acoustic part of the SYMBIOSIS platform consisted of 2 USBL elements with 5 hydrophones each. An underwater transmitter sent a chirp signal every 0.11 s, and each USBL unit record one acoustic file per channel. Each file contains 20 segments of 0.7s each, with a small pause between subsequent audio segments and a slightly longer pause between file records, required to save files on a storage unit. This also explains the uneven time spacing between transmissions and leads to interruptions in detected path segments. We synchronize different recordings using the first NMF arrival as a reference.

Fig. 4 (top panel) shows a comprehensive view of all relevant normalized matched filter peaks from both USBLs. Each peak represents a reflection from the environment or from the target, as recorded from one of the channels of the receiving array. The x-axis represents the peak observation time relative to the start of a given transmission, whereas the y-axis reports the transmission count for each signal sent by the SYMBIOSIS platform (earliest transmissions at the bottom).

The strongest peaks are clearly earliest in the figure, and related to strong reflectors in the environment. Yet, a cluster of peaks starting just before transmission 300 suggests a target getting farther from the ship (as inferred from the increasing observation time of the corresponding peaks). This is compatible with the turtle's release, as the animal swam linearly away from the releasing ship. We note that the trajectory of the animal is composed of two sets of points, one showing a linear movement, and a second set composed of detected target reflections around such a linear trajectory. This is due to some elements of the SYMBIOSIS array being shadowed by other construction components of the array, and showing a slight jitter in the sampling times. Still, the NMF peaks are very well connected and compatible with the trajectory of a moving target.

While our algorithm can operate with multiple sub-arrays joined together, one of the key assumptions for this is that acoustic sampling is synchronous across all elements. In SYMBIOSIS, this was realized through a sync-in signal sent by the central embedded system to all software-defined sub-arrays. Unfortunately, this signal experienced a malfunctioning during the experiment. Therefore, we employed the rest of the data from the deep THEMO site to also refine the capabilities of the localization algorithm. We tuned the parameters of the detection steps (and in particular, of the clustering step) in order to localize targets independently from each software-defined USBL array mounted on the SYMBIOSIS platform. In the worst case, this would allow us to use a single USBL unit to localize targets by choosing the one with the best output in terms of signal-to-noise ratio.

The deep THEMO deployment represented a perfect opportunity to verify and tune the capabilities of the algorithm to discriminate between static and slowly moving targets. Fig. 4 (bottom-right panel) shows the automatic isolation of relevant target reflections and target tracking. We observe a stationary arrival detected (blue dots at about 0.01 s) that is probably a reflection from parts of the deployed platform. It was successfully removed from further processing by our algorithm. For each trajectory segment, our algorithm reports three values corresponding to the movement speed, to the average distance throughout the detection, and to the number of transmissions in which the target was identified. We show such values close to each trajectory segment in the figure for clarity. Fig. 4 (bottom-left panel) shows the result of fine localization algorithm in azimuthal plane. The algorithm

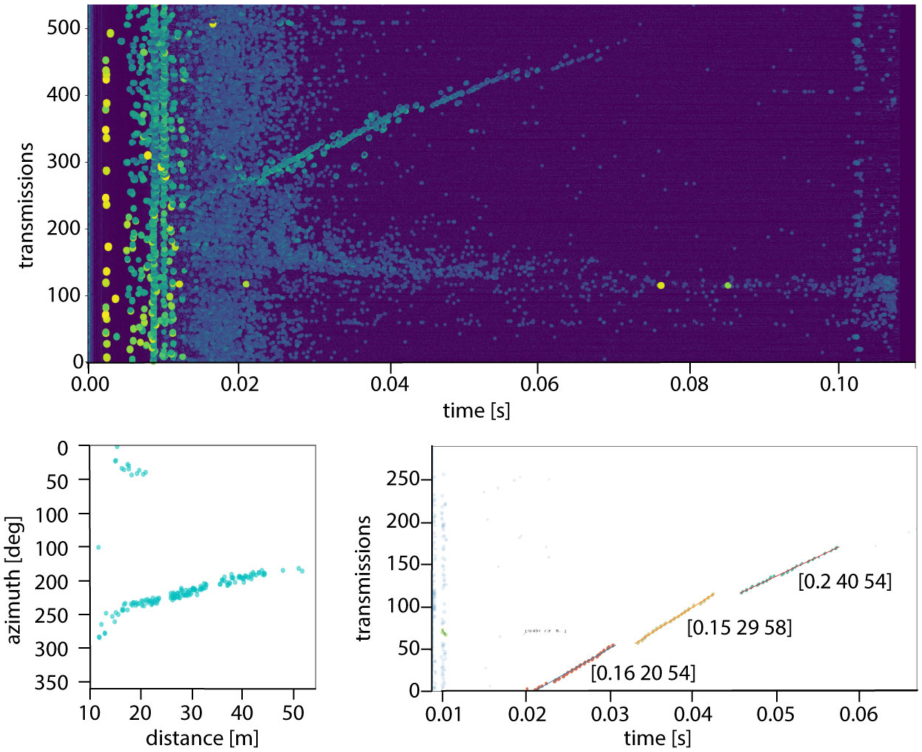

Fig. 4. Deep THEMO dataset. Top panel: NMF output with relevant peaks (yellow and light green hues correspond to stronger arrivals). Bottom-left panel: estimated azimuthal angle of arrival of the signal reflected off the target turtle. Bottom-right panel: turtle path tracking.

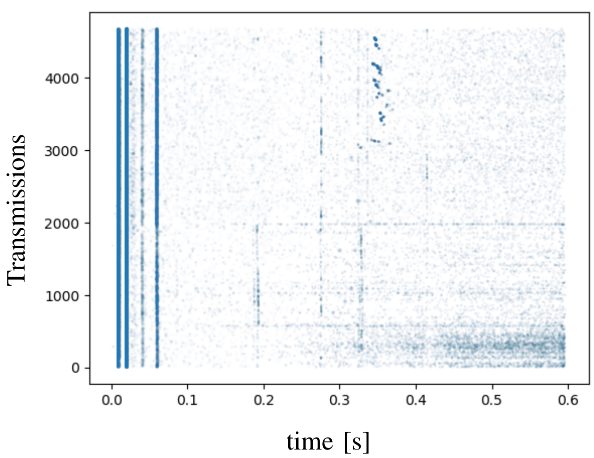

Fig. 5. Eilat dataset. NMF peaks from over 4600 transmissions. Each peak is represented with a light blue transparent dot. The opaqueness of the dots represents the density of NMF peaks.

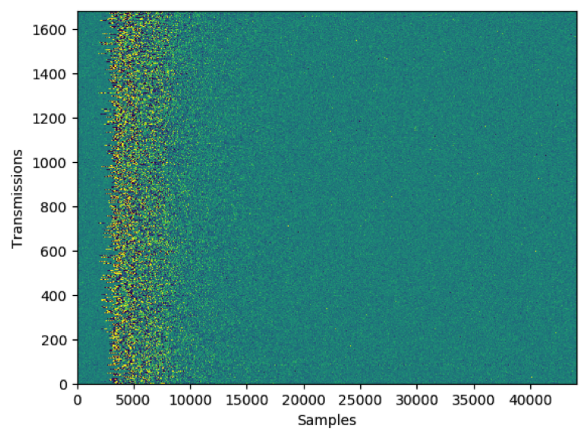

Fig. 6. Eilat dataset. Waterfall matrix that represents raw acoustic data from one of the USBL channels for transmissions 3000-4600.

shows consistent and slowly-changing bearing angle estimates as the turtle swims away. The capability of the algorithm to work in this scenario set the stage for a more complex test of full 3D localization in shallower waters, as detailed in the next subsection.

C. Eilat deployment

The Eilat deployment was performed during November 2020 in the Israeli Red sea, and lasted about 2 weeks. The deployment involved the whole SYMBIOSIS platform, including surface Radio Frequency (RF) communication units to enable a radio link to the platform from a shore location. The platform was deployed at about 5 m of depth, and 500 m away from the shore.

The area where the platform was deployed constitutes a very challenging acoustic environment. In contrast with the deep THEMO deployment, where reflections from bottom are barely noticeable, the relatively low depth of the Eilat deployment (below 30 m), and the many coral reefs on the bottom of the Red sea create a very rich multipath acoustic channel. Reflections from environmental features are usually stronger than those from small fish, which makes the detection task particularly complex. Moreover, fish detections are rare in this area at this time of the year, as confirmed by biologists collaborating with the SYMBIOSIS project.

For this deployment, we analyse a part of the dataset that consists of about one working day of recordings. We show the collection of relevant NMF peaks in Fig. 5. The opaqueness of the dots represents the density of peaks in the area of interest. This picture confirms our hypothesis that the peak density is a good indicator of the presence of a reflecting object. We remark that we made several initial attempts at tuning the sensitivity of the platform, as shown by the greater level of noise at the beginning of the data set, as well as by the stationary arrivals (that appear as vertical lines in the figure), whose intensity oscillates between transmissions 800 and 4500. We will now focus on the last 1650 transmissions.

Fig. 6 shows a closer look at the waterfall of acoustic sample values for one acoustic channel of one software-defined USBL unit. In the figure, green hues represent the noise level, whereas yellow hues convey a stronger signal. The figure shows once more that array processing is paramount for underwater target detection even before accurate localization: the signal corresponding to the target is buried in noise, it remains invisible to manual inspection, and needs to be matched with additional acoustic channel outputs in order to improve the signal-to-noise ratio. In fact, simple signal thresholding would not enable reliable target detections.

Fig. 7 shows localization results. The top panel collects relevant NMF peaks, where each dot represents an identified cluster, and thus a possible target detection. In blue, we highlight stationary targets that were correctly identified and removed. We note that some arrivals did not fit the conservative conditions of stationary arrival definition: however, most of them are removed from further processing, and those that were not removed as stationary are not identified as a relevant moving target. Clusters of points crossed by lines represent detections that passed our filtering steps and were tracked to identify movement. For each of these detections, we report the estimated speed and average distance from the platform.

In this set of transmissions, most of the detected targets were located at a range of 260 to 270 meters from the platform, and were slowly drifting away from the platform. The correlation of the distance and azimuthal angle of arrival of the detected targets is shown in the bottom-left panel of Fig. 7. Here, we see that spurious detections surround the platform, and correspond to clutter from the environment or to surface signal reflections that are then correctly detected as out of boundaries. The figure also suggests that the SYMBIOSIS platform experienced some limited rotation along its vertical axis. Finally, in the bottom-right panel of Fig. 7, we report the depth of each localized target. We recall that our algorithm excludes from further processing any targets whose estimated depth falls out of boundaries. Here, we observe that detections distribute throughout the watercolumn, and tend to happen towards the end of the experiment. While ground truth observations are not available for this experiment, we generally observe that a distribution of fish targets at different depths and a general sparsity of targets of sufficient size to be visible to the acoustic sensors is in line with expectations for the Israeli Red sea in the period of the year when the experiment was carried out.

IV. Conclusions

In this paper, we summarized our work on the validation of a 3D localization and tracking algorithm based on DoA estimation in real sea environments. In particular, we focused on the detection, localization and tracking of marine fauna specimens, in order to establish when they would swim sufficiently close to an underwater platform. The algorithm deployed in the wild is a significantly extended version with respect to the one presented in [7].

We have tested the proposed algorithm in various natural water basins. Besides sweet water lake experiments in [7], we were able to test it in various salty water deployments. In this work, we consider two different tests environments: one in relatively deep waters in the Mediterranean sea, and another in very shallow waters in the Red sea. The results of both tests proved that the algorithm discriminates the environmental features of acoustic reflections from other target reflections of interest. In both cases the estimated information of target movements corresponds to the expected ones. Notably, these do not include just highly reflective or active targets, but also smaller and weaker targets such as the fish specimens of interest for the SYMBIOSIS project.

Acknowledgment

This work has received support from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement no. 773753 (SYMBIOSIS), and from the Italian Ministry for University and Research under the initiative "Departments of Excellence" (Law 232/2016).

The authors would like to thank Eyal Bigal for his help in conducting the scuba diving experiments.

References

[1] A. Tesei, S. Fioravanti, V. Grandi, P. Guerrini, and A. Maguer, "Localization of small surface vessels through acoustic data fusion of two tetrahedralarrays of hydrophones," in *Proc. ECUA*, Jul. 2012.

[2] X. Zhong, A. B. Premkumar, and W. Wang, "Direction of arrival tracking of an underwater acoustic source using particle filtering: Real data experiments," in *Proc. IEEE Tencon–Spring*, Apr. 2013, pp. 420–424.

[3] W. Chen, T. Luo, F. Chen, F. Ji, and H. Yu, "Directions of arrival and channel parameters estimation in multipath underwater environment," in *Proc. MTS/IEEE OCEANS*, Apr. 2016, pp. 1–4.

[4] H.-P. Tan, R. Diamant, W. K. Seah, and M. Waldmeyer, "A survey of techniques and challenges in underwater localization," *Ocean Engineering*, vol. 38, no. 14-15, pp. 1663–1676, 2011.

[5] S. Stergiopoulos, *Advanced signal processing handbook: theory and implementation for radar, sonar, and medical imaging real time systems*. CRC press, 2018.

[6] E. Dubrovinskaya and P. Casari, "Underwater direction of arrival estimation using wideband arrays of opportunity," in *Proc. MTS/IEEE OCEANS*, Jun. 2019, pp. 1–7.

[7] E. Dubrovinskaya, V. Kebkal, O. Kebkal, K. Kebkal, and P. Casari, "Underwater localization via wideband direction-of-arrival estimation using acoustic arrays of arbitrary shape," *MDPI Sensors*, vol. 20, no. 14, pp. 1–20, Jul. 2020.

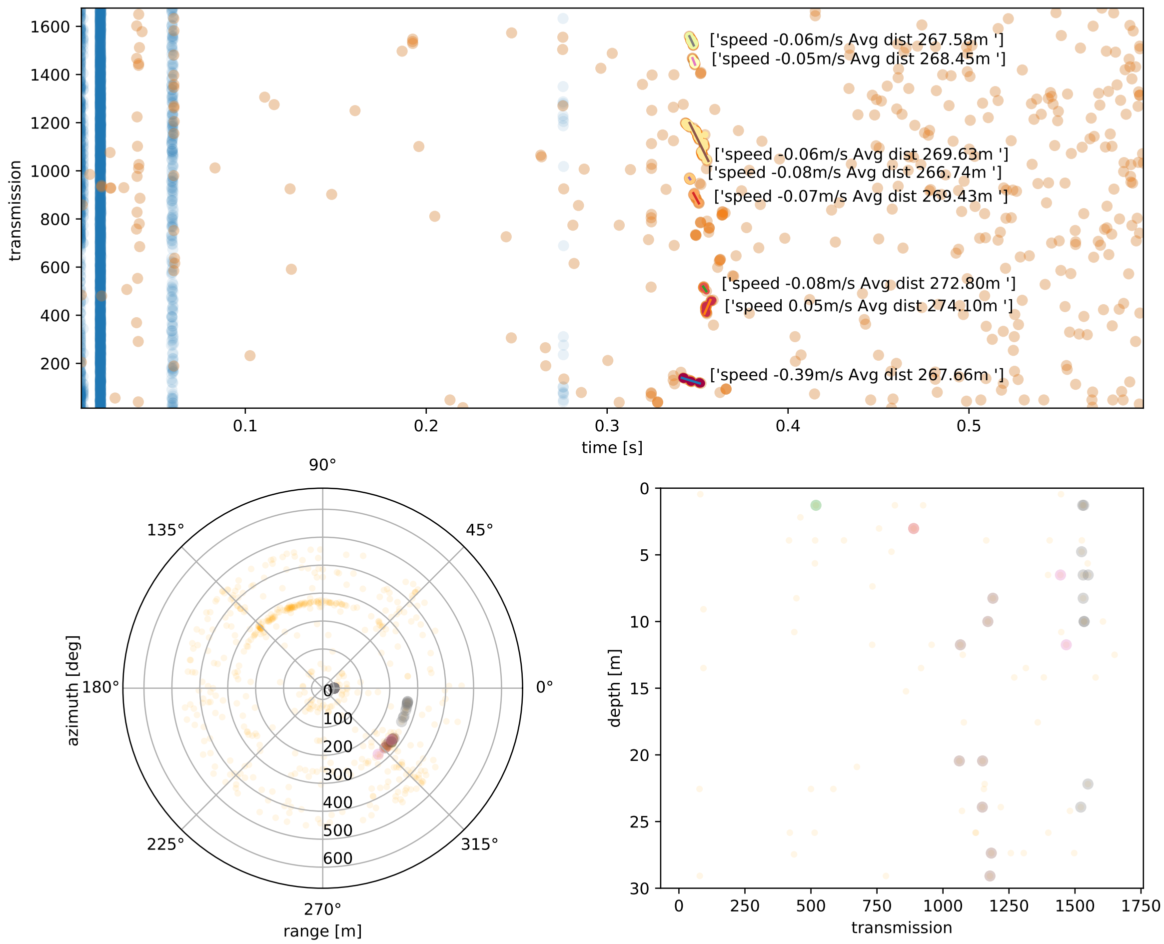

Fig. 7. Eilat dataset. (Top panel) NMF relevant peaks. Each dot represents a cluster of peaks with a possible target detection. Blue dots represent clusters that are defined as stationary. Lines represent clusters that are detected as possible target tracks with relevant speed. (Bottom-left panel) Azimuthal plane localization. Orange points represent all relevant target tracks. Other colors represent corresponding parts of relevant target tracks after filtering out detections whose depth estimates fell out of the sea's boundaries. (Bottom-right panel) Depth plane localization of relevant target tracks limited by environment boundaries.

[8] R. Diamant, V. Voronin, and K. G. Kebkal, "Design structure of SYMBIOSIS: An opto-acoustic system for monitoring pelagic fish," in *Proc. MTS/IEEE OCEANS*, 2019, pp. 1–6.

[9] D. Kellner, J. Klappstein, and K. Dietmayer, "Grid-based DBSCAN for clustering extended objects in radar data," in *2012 IEEE Intelligent Vehicles Symposium*. IEEE, 2012, pp. 365–370.

[10] J. Wang, S. Bai, and B. Englot, "Underwater localization and 3d mapping of submerged structures with a single-beam scanning sonar," in *Proc. IEEE ICRA*. IEEE, 2017, pp. 4898–4905.

[11] Y. Geng, Z. Wang, C. Shi, R. Nian, C. Zhang, B. He, Y. Shen, and A. Lendasse, "Seafloor visual saliency evaluation for navigation with bow and DBSCAN," in *Proc. MTS/IEEE OCEANS*. IEEE, 2016, pp. 1–5.

[12] C. Capus, Y. Pailhas, K. Brown, D. M. Lane, P. W. Moore, and D. Houser, "Bio-inspired wideband sonar signals based on observations of the bottlenose dolphin (tursiops truncatus)," *J. Acoust. Soc. Am.*, vol. 121, no. 1, pp. 594–604, 2007.

[13] R. Hedley, Y. Huang, and K. Yao, "Direction-of-arrival estimation of animal vocalizations for monitoring animal behavior and improving estimates of abundance," *Avian Conservation and Ecology*, vol. 12, no. 1, 2017.

[14] R. Diamant, A. Knapr, S. Dahan, I. Mardix, J. Walpert, and S. DiMarco, "THEMO: The Texas A&M - University of Haifa - eastern mediterranean observatory," in *Proc. MTS/IEEE OCEANS*, Kobe, Japan, May 2018, pp. 1–5.

[15] "Themo - deep sea observatory," http://themo.haifa.ac.il/buoys/mmp, 2020, [Online; accessed 30-July-2020].