

REAL-TIME PASSIVE ACOUSTIC 3D TRACKING OF DEEP DIVING CETACEAN BY SMALL NON-UNIFORM MOBILE SURFACE ANTENNA

M. Poupard ^(a), M. Ferrari ^(a), J. Schlüter ^(a), R. Marxer ^(a), P. Giraudet ^(a),
V. Barchasz ^(c), V. Giès ^(c), JM. Prévot ^(c), T. Soriano ^(d), G. Pavan ^(b), H. Glotin ^{(a,c)*}

^(a) Univ. Toulon, Aix Marseille Univ., CNRS, LIS, DYNI team, SABIOD, Marseille, France

^(b) DSTA-CIBRA, University of Pavia, Italy

^(c) SMIoT, University Toulon, France, ^(d) Cosmer, University Toulon, France

ABSTRACT

Detecting and localizing the echolocation clicks of sperm whales allows to research their diving behavior, but existing methods are limited in range, imprecise, or costly. In this work, we demonstrate that we can obtain a high definition 3D track of deep diving cetaceans from a five-channel, small-aperture hydrophone array on a moving autonomous surface vehicle (ASV), enabled by the vessel’s hydrodynamic quality and a high recording sample rate. Real-time processing is achieved by splitting our non-uniform array into two parts for time delay of arrival estimation. Resulting 3D tracks depict the behavior of the cetacean in abyss (-1.2 km), with one position per second. This high resolution allows us to observe a correlation between the repetition rate of the predator’s biosonar and the tortuosity of its track. Our proposed mobile observatory may offer new insights about whale behavior and its success of foraging close to vessel traffic.

Index Terms— 3D passive acoustic tracking, ASV, tortuosity, *Physeter macrocephalus*, sperm whale, clicks, biosonar, embedded real-time system

1. INTRODUCTION

Sperm whales (*Physeter macrocephalus*, *Pm*) are categorized as vulnerable due to human-made threats such as commercial whaling, interactions with fisheries [1], noise [2], chemical pollution, global warming or collisions with ships [3]. In order to assess population status, to model behavior, or to prevent ship collisions, passive acoustic monitoring provides a viable option: *Pm* spends more than 70% of its life in deep

foraging dives down to -2 km, using echolocation (by far-travelling transients [4]) for orientation and prey localization.

Existing work uses different methods to record and localize the echolocation clicks of cetaceans. Static hydrophone arrays using underwater buoys provide stable recordings and observations of foraging [5, 6, 7] but can only monitor a fixed location, making it difficult to track animals over longer distances. Moreover, a permanent quadriphonic sonobuoy array [8] obtains a track of the animal only in 2 dimensions. A bottom-mounted array with small aperture (2 m) and high sampling rate was used for 3D localization of multiple whales [9, 10]. More flexibility can be achieved by mounting hydrophones on a vessel, albeit introducing more challenges. In [11], a wide-aperture towed array of two hydrophones was used for tracking 2D *Pm* dive profiles, taking advantage of surface-reflected paths, but not allowing 3D localization. In [12], a first attempt of a moving deep 6-hydrophone array resulted only into range estimation of the cetacean.

In this work, we use a small nonuniform array of 5 hydrophones mounted directly under an autonomous surface vehicle (ASV). This setup is made challenging by the small aperture and the recording of sound close to the surface above the thermocline. To the best of our knowledge, this is the first demonstration that cetaceans can be tracked precisely in 3D from a moving, near-surface hydrophone array.

This mobile observatory opens up new possibilities to analyze the movements and behaviors of cetaceans, either when expressing their natural behavior or when disturbed by human activities nearby. Thus, it may assist in the design of mitigation and conservation programs.

2. DATA ACQUISITION

We mounted 5 hydrophones under the ASV “Sphyrna” of Seaproven SA (Fig. 1). The autonomous vehicle is 17 m long and very stable. According to the Motion Processing Unit (MPU), during 1 hour of drift at Beaufort sea state 2, roll variance was 1 degree, pitch below 1, yaw 24.

We mounted four hydrophones (H_1 to H_4 , 2 Cetacean

(*) Corr. author glotin@univ-tln.fr. This research has been possible due to Seaproven and TechnoLaval. We thank Port-Cros Nat. Park, H. Bergère and coll. for logistic help. We thank Albert II foundation, Pôle ‘Information Numérique Prévention et Santé’ INPS UTLN, Accobams, Pelagos, CNAM LAVAL and TES minister Nicolas Hulot for their label. We thank F. and V. Sarano and Longitude 181 ONG for their help in the mission. We thank P. Cosantino, N. Boizot, J. Patris, M. Grimaldi, J. Olson for their advises. This project is in the continuity of FUI AAP14 Sycié, takes place into SABIOD and EADM MADICS CNRS groups, and is into FUI Abyssound and ANR Smiles.



Fig. 1. The ASV Sphyrna, 17m long, Polynesian design.

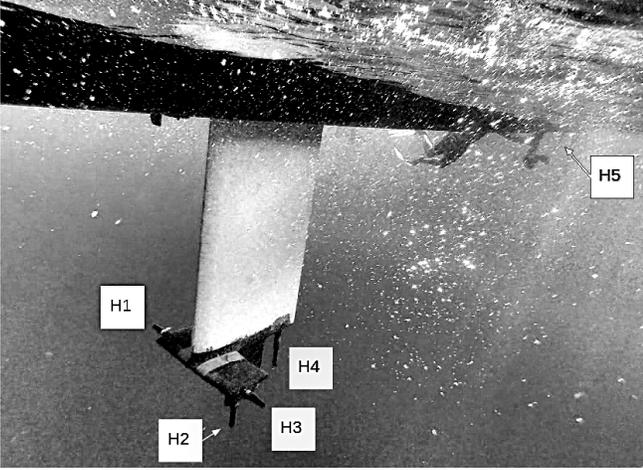


Fig. 2. Layout of the 4+1 hydrophone array. The inter-hydrophone distances under the keel are 35, 59, 59, 63, 63, 70 cm. Hydrophone H_5 is placed at the stern, 7 m away.

Res. CR75 and 2 CR57) under the keel at a depth of -2 m, spaced 35–70 cm apart, and a CR75 (H_5) at the stern of the ASV at a depth of -0.2 m and 7 m far from the keel (Fig. 2).

We use the high-resolution sound card JASON Q. (Fig. 3), designed by the scientific SMIoT platform [13]. It allows recording 5 uncompressed channels with a sampling rate (SR) of up to 2 MHz and a resolution of 16 bits. We record at 300 kHz SR which is a minimal SR for localization with small aperture. Along with the sound, we record the ship’s location and orientation provided by the MPU.

We acquired our data on August 12, 2018 at 11:00–13:00 local time, offshore Cap Sicié, Toulon, France, with the ASV drifting at an average speed of 0.26 m/s.

3. DATA ANALYSIS

Trajectories are derived through the three following steps: (1) detection of the echolocation clicks in each channel, (2) estimation of time delay of arrival (TDoA) of each click between the five channels, and (3) reconstruction of the 3D position



Fig. 3. The JASON sound card (SMIoT), 5 x 1 MHz sampling rate at 16 bits resolution, placed into the drone. The luxmeter was not used in this work.

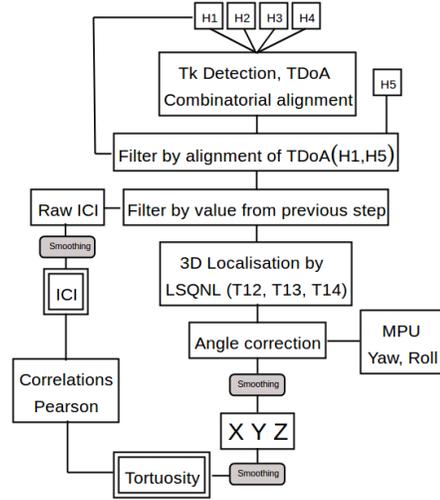


Fig. 4. Summary diagram of the analysis.

(Fig. 4). In addition, we analyze the time complexity of our real-time approach.

3.1. Click detection

To detect clicks, we cross-correlate the signal with one period of a 12.5 kHz sine (central frequency of the click of P_m), followed by a Teager-Kaiser filter [14, 15] and extraction of local maxima in 20 ms windows. To remove maxima caused by background noise, we only keep those above three standard deviations over the mean logarithmic energy of all detections.

3.2. Time Delay of Arrival (TDoA) estimation

Let τ_{ij} denote the TDoA between hydrophones H_i and H_j , and let x_i denote the signal recorded by H_i . We divide the TDoA estimation into two parts: first, we compute the three independent TDoAs $\tau_{21}, \tau_{31}, \tau_{41}$ of the keel antenna (H_1 to H_4), then we compute the remaining TDoA τ_{15} . Compared to estimating all TDoAs jointly, this reduces the time complexity (Sec. 3.4), and accounts for our nonuniform array, with H_5

placed further apart and affected more strongly by noise from the surface (Fig. 2).

For the keel antenna, we first compute cross-correlations between all 6 possible hydrophone pairs, and make them positive by subtracting their minimal values. Then we search for the combination of $\tau_{21}, \tau_{31}, \tau_{41}$ that maximizes their product:

$$\tau_{21}, \tau_{31}, \tau_{41} = \arg \max_{\hat{\tau}_{21}, \hat{\tau}_{31}, \hat{\tau}_{41}} \prod_{i=2}^4 \prod_{j=1}^{i-1} (x_i \otimes x_j - \min(x_i \otimes x_j))_{\hat{\tau}_{ij}}, \quad (1)$$

where \otimes denotes cross-correlation, and the relation $\tau_{ij} = \tau_{i1} - \tau_{j1}$ reduces the search to three dimensions. Ideally, this finds the combination that matches up the echolocation clicks.

For the remaining τ_{51} , we use the same method, but hold $\tau_{21}, \tau_{31}, \tau_{41}$ fixed, only searching over one dimension.

3.3. Localization and ICI

Using a nonlinear solver, we estimate the 3D positions of the whale from the TDoAs [16, 15]. The tracks are stabilized according to the yaw and roll from the ship’s MPU, and smoothed with a running average over 20 s.

We also compute the inter-click interval (ICI, the time between two clicks) at each click. To remove artifacts from the solver or false positives, the ICI sequence is smoothed with a running average of 32 seconds.

3.4. Time complexity

In practice, we need to evaluate Eq. 1 for a limited range of TDoAs only: τ_{ij} cannot be larger than the distance $d_{H_i H_j}$ between H_i and H_j divided by the speed of sound c . When the distance between each pair of hydrophones is of the same order, the time complexity of this evaluation is:

$$O\left(\prod_{i=1}^T (S \cdot M_i)\right) = O(S^T M^T), \quad (2)$$

where S is the sampling rate, T is the number of independent TDoAs, and M_i amounts to twice the maximal TDoA between channels i and 1:

$$M_i = 2 \frac{d_{H_i H_1}}{c}. \quad (3)$$

All that is left is to compute the fourth TDoA τ_{51} holding the others fixed. This requires a search along a single coordinate with a wider range. The final complexity of our TDoA filter is:

$$O(S^3 M^3 + S M_5) = O(S^3 M^3). \quad (4)$$

In summary, the complexity of our (4+1) hydrophones TDoA estimator is a polynomial with a degree lower than if we had computed all 4 independent TDoAs jointly. In our

setup, separating the estimation of τ_{51} provides the best time gain because the distance between the fifth hydrophone and the others is one magnitude higher than distances within the keel antenna. Of course, this also leads to some TDoA combinations not being evaluated at all. Thus, splitting the antenna may only be done if the first sub-antenna provides trustworthy TDoA by itself – the last channel now only serves to validate or reject wrong TDoA estimations.

4. RESULTS

In the following, we present the results of the analysis (Sec. 3) of our recordings (Sec. 2). First, we will demonstrate how we can filter out spurious detections using the (4+1) hydrophone array, then we will show the obtained trajectory of the recorded animal, and finally give an example of what we might deduce about its behavior.

4.1. Click detection and verification

In a first experiment, we check whether the fifth channel can be used to reject spurious detections. The scatter plot (Fig. 5, top) of the maximum of the bin-to-bin correlation products with $\tau_{21} \dots \tau_{41}, \tau_{51}$, shows two modes: one from echolocation clicks, the other from noise. We filter out spurious detections by thresholding low correlations within the sub-antenna, and low correlations with H_5 .

We see (Fig. 5, bottom) that this indeed filters out clicks that do not seem to be part of the dive, but caused by the noise of the boat, or other sources. We also see clicks that seemingly originate above the surface, these are reflections we filter out as well.

4.2. Trajectory of the dive

From the remaining localized click detections (3.3), we compute a trajectory over time, shown in Fig. 6.

After its first dive, the animal spends 15 min at the surface (silence). The second dive starts directly with clicks from the surface, and lasts 32 min, maximal depth is -1.2 km.

4.3. Correlation between ICI, Depth and Tortuosity

Finally, to demonstrate which kinds of insights could be won from such a detailed 3D tracking, we compare the Inter-Click Interval $ICI(t)$ with the depth and the tortuosity $\Phi(t)$, an index of movement behavior. A reliable estimate of $\Phi(t)$ against location error is the mean squared displacement [17, 18], i.e. the sum of variances of three dimensions of the displacement in a time window (here 300 s) centered on time t . Our results (Fig. 6) confirm that the inter-click interval ICI decreases with depth Z [19]. The scatter plot of $Z(t)$, $\log(ICI(t))$ and $\log(\Phi(t))$ (Fig. 7) shows that neighbors in time are also nearest on these three variables. Moreover,

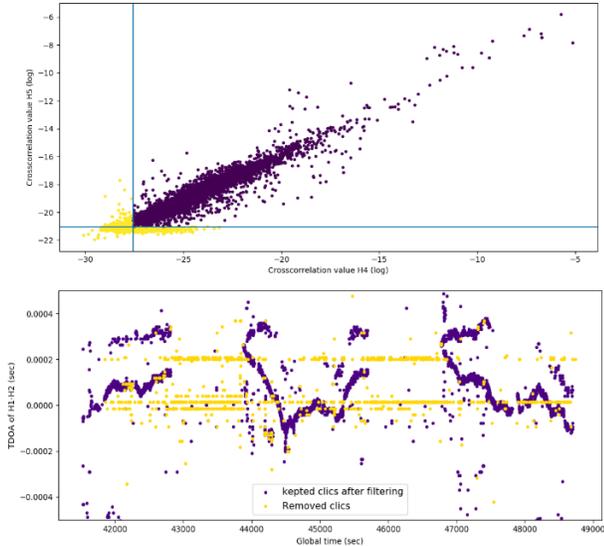


Fig. 5. Top: Filtering of crosscorrelation values (H_4 and H_5). Bottom: TDoA between H_1 and H_2 over time. Yellow points are not used for the analysis.

the Pearson correlation of $\log(ICI(t))$ and $\log(\Phi(t))$ equals -0.46 in the first dive (-0.09 in the next one).

A possible interpretation is that the closer the whale gets to its prey (indicated by a smaller ICI), the more complex its hunt becomes as the prey tries to escape the Pm (Fig. 8).

5. CONCLUSION

This paper demonstrates the first accurate 3D dive track at more than 2 km of distance, by passive acoustic monitoring from an ASV, with a small-aperture array of 5 non-uniformly spaced hydrophones. We show how to take advantage of a fifth hydrophone placed 7 m from the four others to remove false positive detections. Tracking was enabled by the smoothness of the drift of the polynesian-shaped Sphyrna ASV, the high SR of the sound card and adapted algorithms. We thus demonstrate a high definition passive acoustic mobile observatory that may improve our knowledge of cetacean diving behavior in the vicinity of vessel traffic and provide recommendations for ship operators to avoid collisions. This system offers perspectives to monitor the anthropophonic pollution, versus the success of the foraging of the predators. The Sphyrna may be programmed to follow the fauna in silence at a range of up to 6 km during days at a time, and thus assess the quality of their behavior in relation to the environmental context [20]. To the best of our knowledge, this paper is the first to give evidence of a correlation between ICI and Tortuosity of the dive of this mega predator. These clues are related to the prey biomass, and may open new insights of the abyss richness.

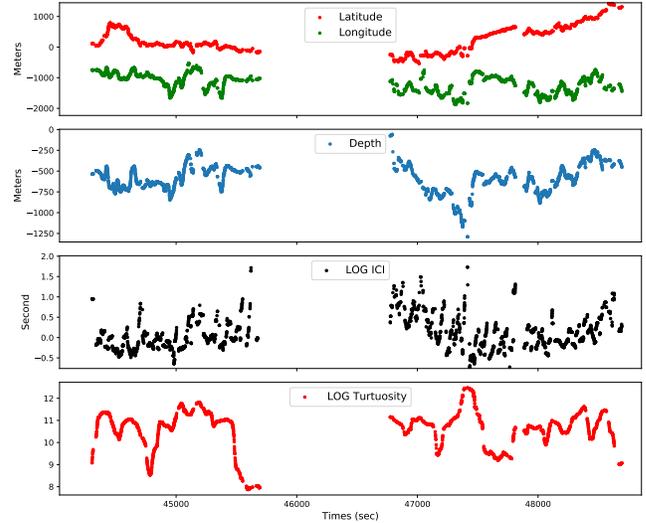


Fig. 6. From Top to Bottom: Position of the Pm , X latitude (Northing), Y longitude (Easting), Depth Z (m), Log ICI (s), Log Tortuosity over time.

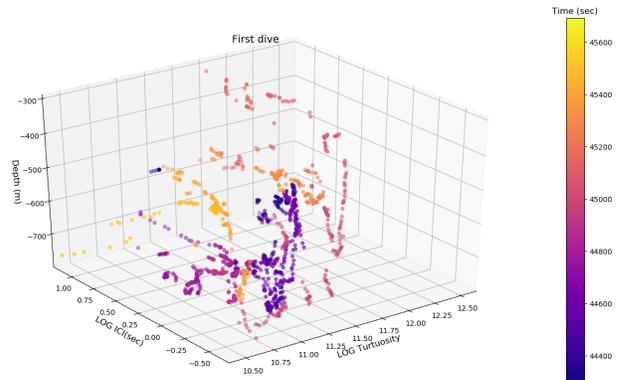


Fig. 7. Depth Z, $\log(ICI)$ and $\log(\text{Tortuosity})$ over time.

6. REFERENCES

- [1] Hucke-Gaete R., Moreno CA., Arata J., and al., “Operational interactions of sperm whales and killer whales with the patagonian toothfish industrial fishery off southern chile,” *Camlr Science*, vol. 11, pp. 127–140, 2004.
- [2] Weilgart LS., “The impacts of anthropogenic ocean noise on cetaceans and implications for management,” *Canadian J. of zoology*, vol. 85, no. 11, pp. 1091–1116, 2007.
- [3] David L. and Di-Méglio N., “Prévention des collisions entre navires et grands cétacés,” 2010.
- [4] Weilgart L. and Whitehead H., “Group-specific dialects and geographical variation in coda repertoire in south

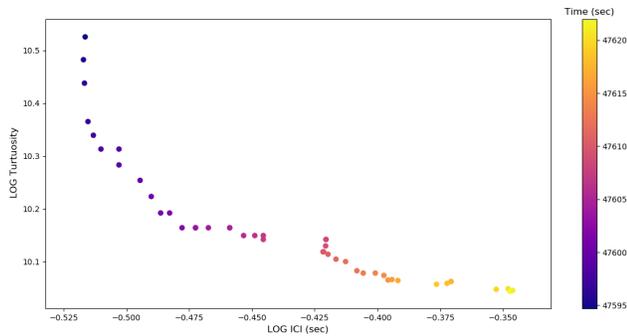


Fig. 8. log-log ICI vs. Tortuosity during 25 s. (from 47595 to 47620 s.) showing high anti-correlation, a progressive decrease of ICI while the tortuosity increases, until the stop of the click train which is supposed to be the capture.

pacific sperm whales,” *Behavioral Ecology and Sociobiology*, vol. 40, no. 5, pp. 277–285, 1997.

- [5] Simard Y. and Roy N., “Detection and localization of blue and fin whales from large-aperture autonomous hydrophone arrays: A case study from the St. Lawrence estuary,” *Canadian Acoustics*, vol. 36, no. 1, pp. 104–110, 2008.
- [6] Roy N., Simard Y., and Gervaise C., “3D tracking of foraging belugas from their clicks: Experiment from a coastal hydrophone array,” *Applied Acoustics*, vol. 71, no. 11, pp. 1050–1056, 2010.
- [7] Glotin H. et al, “Bombyx stereophonic sonobuoy: a permanent survey of cetacean in Pelagos sanctuary,” *Pelagos Research Report*, 2017.
- [8] Brunoldi M., Bozzini G., Casale A., Corvisiero P., Grosso D., and al, “A permanent automated real-time passive acoustic monitoring system for bottlenose dolphin conservation in the mediterranean sea,” *PloS one*, vol. 11, no. 1, pp. e0145362, 2016.
- [9] Bénard-Caudal F., Giraudet P., and Glotin H., “Whale 3D monitoring using astrophysic NEMO ONDE two meters wide platform with state optimal filtering by rao-blackwell monte carlo data association,” *App. Acoustics*, vol. 71, no. 11, pp. 994–999, 2010.
- [10] Favali P., Chierici F., Marinaro G., Giovanetti G., and al., “NEMO-SN1 abyssal cabled observatory in the western ionian sea,” *J. Oceanic Eng.*, vol. 38, no. 2, pp. 358–374, 2013.
- [11] Thode A., “Tracking sperm whale (*p. macrocephalus*) dive profiles using a towed passive acoustic array,” *J. of the Acoustical Society of America*, vol. 116, no. 1, pp. 245–253, 2004.
- [12] Zimmer W., “Range estimation of cetaceans with compact volumetric arrays,” *J. of Acoustical Society of America*, vol. 134, no. 3, pp. 2610–2618, 2013.
- [13] Fourniol M., Gies V., Barchasz V., Kussener E., Barthelemy H., Vauché R., and Glotin H., “Low-power wake-up system based on frequency analysis for environmental internet of things,” in *Int. Conf. on Mechatronic, Embedded Systems, App. IEEE*, 2018, pp. 1–6.
- [14] Kandia V. and Stylianou Y., “Detection of sperm whale clicks based on the Teager–Kaiser energy operator,” *Applied Acoustics*, vol. 67, pp. 1144–1163, 2006.
- [15] Glotin H., Caudal F., and Giraudet P., “Whale cocktail party: real-time multiple tracking and signal analyses,” *Canadian acoustics*, vol. 36, no. 1, pp. 139–145, 2008.
- [16] Giraudet P. and Glotin H., “Real-time 3D tracking of whales by echo-robust precise TDOA estimates with a widely-spaced hydrophone array,” *Applied Acoustics*, vol. 67, no. 11-12, pp. 1106–1117, 2006.
- [17] Almeida P., Vieira MV., Kajin, Forero-Medina G., and Cerqueira R., “Indices of movement behaviour: conceptual background, effects of scale and location errors,” *Zoologia*, vol. 27, no. 5, 2010.
- [18] Benhamou S., “How to reliably estimate the tortuosity of an animal’s path: straightness, sinuosity, or fractal dimension?,” *J. of theoretical biology*, vol. 229, no. 2, pp. 209–220, 2004.
- [19] Jaquet N., Dawson S., and Douglas L., “Vocal behavior of male sperm whales: Why do they click?,” *J. of Acoustical Society of America*, vol. 109, no. 5, pp. 2254–2259, 2001.
- [20] Pavan. G, *Ecoacoustics: The ecological role of sounds*, John Wiley & Sons, 2017.