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

M. Poupard [(a,b)], M. Ferrari [(a,c)], J. Schlüter [(a)], R. Marxer [(a)], P. Giraudet [(a)],
V. Barchasz [(e)], V. Giès [(e)], G. Pavan [(d)], H. Glotin [(a,e)]∗

[(a)] Univ. Toulon, Aix Marseille Univ., CNRS, LIS, DYNI team, SABIOD, Marseille, France
[(b)] Biosong, France ; [(c)] CNRS, LAMFA, Université de Picardie Jules Verne, France
[(d)] DSTA-CIBRA, University of Pavia, Italy ; [(e)] SMIoT, University Toulon, France

ABSTRACT

Detecting and localizing the echolocation clicks of sperm whales provides insight into 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 the 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 foraging success close to vessel traffic.

Index Terms— 3D passive acoustic tracking, ASV, tortuosity, *Physeter macrocephalus*, sperm whale, clicks, biosonar, embedded real-time system, high velocity sound card, small surface array.

1. INTRODUCTION

Sperm whales (*Physeter macrocephalus, Pm*) are categorized as vulnerable due to human threats including commercial whaling, interaction with fisheries [1], noise [2] and chemical pollution, global warming as well as ship collisions [3]. For assessing population status, modeling behavior, or preventing ship collisions, passive acoustic monitoring provides a viable option: *Pm* spends more than 70% of its life in deep waters, foraging, diving downwards up to -2 km, using echolocation for orientation and prey localization. For echolocation, *Pm* emits sequences of transients ("clicks") that travel far under water [4].

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

In this work, we propose to use a small nonuniform array of 5 hydrophones mounted directly under an Autonomous Surface Vehicle (ASV). This setup is made challenging by the mobility of the vessel, the small aperture and the recording of sound done 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 using a moving, near-surface hydrophone array. This mobile observatory opens up new possibilities for analyzing the movement and behavior of cetaceans, whether it is natural or modified by nearby human activities. Thus, it may assist in the design of mitigation and conservation programs.

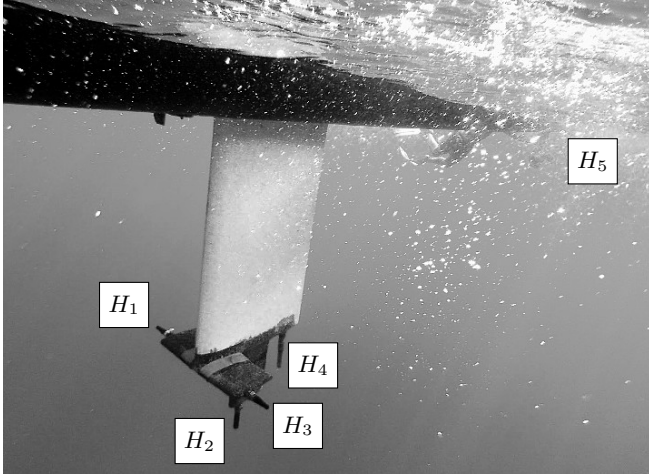**Fig. 1**. The ASV Sphyrna, 17m long, Polynesian design.

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.

Fig. 3. The JASON sound card (SMIoT), 5 x 2 MHz sampling rate at 16 bits resolution, placed into the drone. The luxmeter was not used in this work.

2. DATA ACQUISITION

We mounted 5 hydrophones under the ASV "Sphyrna" of the society Sea Proven (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 and pitch variances were about 1 degree, and yaw 24. We mounted four hydrophones (H_1 to H_4, 2 Cetacean 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 used the high-resolution sound card JASON (Fig. 3), designed by the scientific SMIoT platform [12]. It allows recording 5 uncompressed channels with a sampling rate (SR) of up to 2 MHz and a resolution of 16 bits; here we recorded at 300 kHz. A high SR is crucial for localization with small aperture. Along with the sound, we recorded 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, while the ASV was drifting at an average speed of 0.26 m/s.
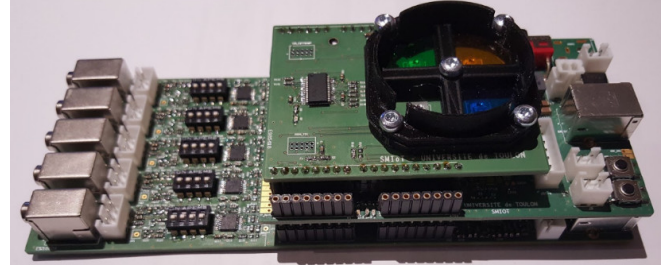
3. DATA ANALYSIS

Trajectories are derived using three steps described in the following: (1) detection of the echolocation clicks in each channel, (2) estimation of the time delay of arrival (TDoA) of each click between the five channels, and (3) reconstruction of the 3D position (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 (bandwidth of *Pm* is 10–15 kHz [13]), followed by a Teager-Kaiser filter [14, 15] and the extraction of local maxima in 20 ms windows (twice the inter-pulse interval of 10 ms [16]). To remove maxima caused by background noise, we only keep those above three standard deviations [17] 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 surface noise (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:

$$\tau_{21}, \tau_{31}, \tau_{41} =$$
$$\operatorname*{arg\,max}_{\hat{\tau}_{21}, \hat{\tau}_{31}, \hat{\tau}_{41}} \prod_{i=2}^{4} \prod_{j=1}^{i-1} \left(x_i \circledast x_j - \min\left(x_i \circledast x_j \right) \right)_{\hat{\tau}_{ij}}, \quad (1)$$

where $\circledast$ denotes cross-correlation, and the relation $\tau_{ij} = \tau_{i1} - \tau_{j1}$ reduces the search to three dimensions. Ideally,
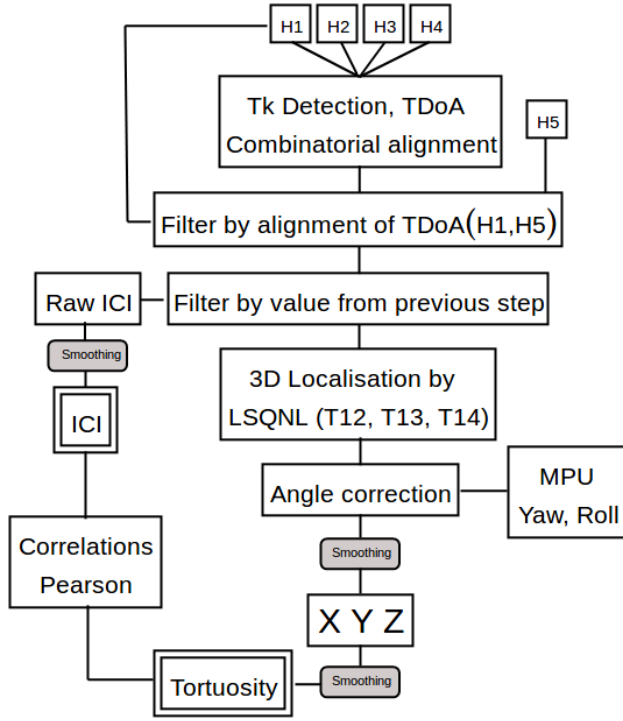
Fig. 4. Summary diagram of the analysis.

this finds the combination that matches up the echolocation clicks: unlike background noise, they are consistent between channels. 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 (X, Y and depth Z) from the TDoAs [18, 15]. The tracks are stabilized according to the yaw and roll from the MPU of the ship, and smoothed with a running average over 20 s. The noise of the raw data is low due to the stability of the Sphyrna and the good weather conditions.

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 the same order, the time complexity of this evaluation is:

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

where S is the sampling rate, K is the number of independent TDoAs, and M_i amounts to twice the maximal TDoA between channels i and 1: $M_i = 2.d_{H_i H_1}/c$.

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

$$O\left(S^3 M^3 + S M_5\right) = O(S^3 M^3). \qquad (3)$$

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} is the most time efficient because the distance between the fifth hydrophone and the others is one magnitude higher than the distances within the keel antenna. However, 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 TDoAs by itself – the last channel now only serves to validate or reject wrong TDoA estimations.

We do not compare results with a joint computation of all 4 independent TDoAs, as this is infeasible: Computation for our (4+1) split took 4 hours, a joint computation would take 2800 times longer.[1]

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 resulting trajectory of the recorded animal,[2] and finally give examples 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 $\tau_{21} \ldots \tau_{41}$, τ_{51}, shows two modes: one from echolocation clicks, the other from noise (blue histograms on the left and bottom). We filter out spurious detections by thresholding low correlations within the sub-antenna, and low correlations with H_5 (with thresholds chosen manually according to the histograms).

We see 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 (yellow horizontal lines in Fig. 5, bottom). We also see parallel click trains caused by reflections at the surface; their computed location is above sea level, so we can easily filter them afterwards.

---

[1]We would have to search a range of ± 9.3 ms (7 m/1500 $\frac{m}{s}$), which equals 2800 samples at 300 kHz, for every combination of the other 3 TDoAs.
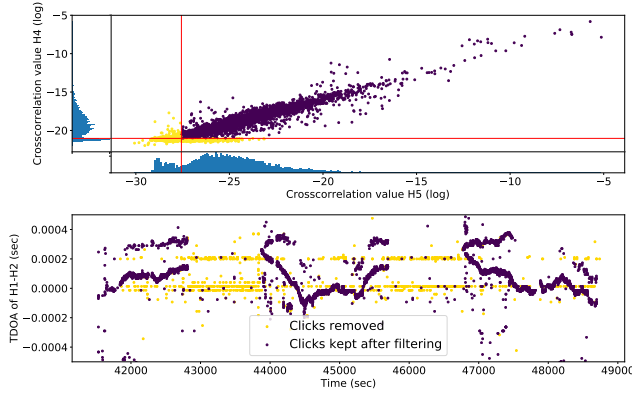[2]Animation video: `https://youtu.be/3ey4LaVhxiA`

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.
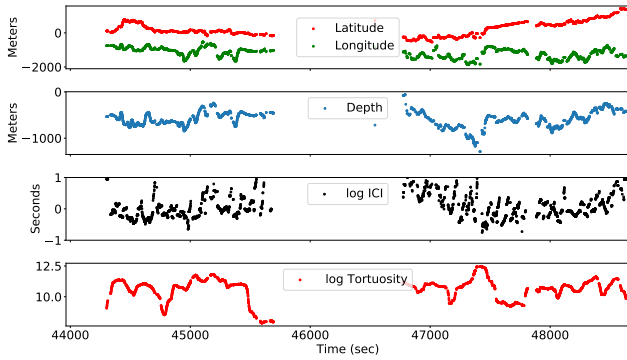

Fig. 6. From Top to Bottom: Position of the *Pm*, X latitude (Northing), Y longitude (Easting), Depth Z (m), Log(ICI) (s), Log(Tortuosity) over time.

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 at the surface, and lasts 32 min, maximum depth is −1.2 km.

4.3. Correlation between ICI, Depth and Tortuosity

Lastly, to demonstrate which kinds of insights can be acquired from detailed 3D tracking, we compare $ICI(t)$ with the depth and the Tortuosity $T(t)$, which is an index of movement behavior. An estimate of $T(t)$ robust against location error is the mean squared displacement [19], i.e. the sum of the variances (V) of the three dimensions (X, Y and Z) of the displacement in a time window (here 300 s):

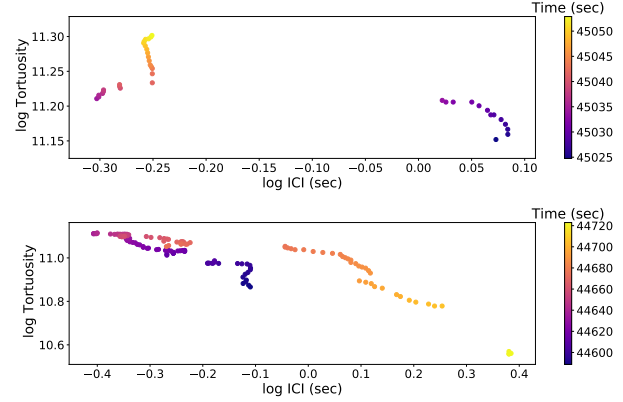$$T = V(X_1, \ldots, X_n) + V(Y_1, \ldots, Y_n) + V(Z_1, \ldots, Z_n) \quad (4)$$

Fig. 7. Scatter plots of log ICI vs. log Tortuosity (T). The sequence of 25 sec (Top) shows that the ICI and T are not dependent by construction. The 2 min sequence (Bottom) shows a strong anticorrelation between ICI and T. It could be interpreted as an approach to a prey, a catch (at the minimum of ICI and maximum of T), and then relaxing ('slack off').

Our results (Fig. 6) confirm that the ICI decreases with depth Z [20]. Moreover, the Pearson correlation of $\log(ICI(t))$ and $\log(T(t))$ varies in ± 0.8 during the dives. A possible interpretation for passages of high anticorrelation is that the closer the whale gets to its prey (smaller ICI, higher T), the more complex its hunt becomes as the prey tries to escape the *Pm* (Fig. 7). After a catch, the animal may slack off (ICI increases, T decreases). Future work shall compare this behavior to other odontoceti [21].

5. CONCLUSION

This paper demonstrates the first accurate 3D dive track at more than a 2 km 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 ASV, the high sampling rate of the JASON sound card and the 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 we might also use our findings to provide recommendations for ship operators to avoid collisions. The Sphyrna may be programmed to follow the fauna in silence at a range of up to 6 km for days at a time, and thus assess the characteristics of their behavior in relation to the environmental context [22]. To the best of our knowledge, this paper is the first to provide such tracks from an ASV, and to show correlation between ICI and Tortuosity of the dive of the predator. These clues are related to the prey biomass, and may provide new insights on the richness of the abyss.

6. REFERENCES

[1] R. Hucke-Gaete, CA. Moreno, J. Arata, et 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] L.-S. Weilgart, "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] L. David, "Disturbance to mediterranean cetaceans caused by vessel traffic," *Cetaceans of the Med. and Black Seas: State of Knowledge and Conservation Strategies*, vol. 11, 2002.

[4] L. Weilgart and H. Whitehead, "Group-specific dialects and geographical variation in coda repertoire in south pacific sperm whales," *Behavioral Ecology and Sociobiology*, vol. 40, no. 5, pp. 277–285, 1997.

[5] Y. Simard and N. Roy, "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. 36.1, no. 1, pp. 104–110, 2008.

[6] N. Roy, Y. Simard, and C. Gervaise, "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] H. Glotin et al., "Bombyx stereophonic sonobuoy: a permanent survey of cetacean in Pelagos sanctuary," *Pelagos Research Report*, 2017.

[8] M. Brunoldi, G. Bozzini, A. Casale, P. Corvisiero, D. Grosso, et 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] F. Bénard-Caudal, P. Giraudet, and H. Glotin, "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] A. Thode, "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.

[11] W. Zimmer, "Range estimation of cetaceans with compact volumetric arrays," *J. of Acoustical Society of America*, vol. 134, no. 3, pp. 2610–2618, 2013.

[12] M. Fourniol, V. Gies, V. Barchasz, E. Kussener, H. Barthelemy, R. Vauché, and H. Glotin, "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.

[13] P.-T. Madsen, R. Payne, N. Kristiansen, M. Wahlberg, I. Kerr, and B Møhl, "Sperm whale sound production studied with ultrasound time/depth-recording tags," *J. of Exp. Biology*, vol. 205, no. 13, pp. 1899–1906, 2002.

[14] V. Kandia and Y. Stylianou, "Detection of sperm whale clicks based on the Teager–Kaiser energy operator," *Applied Acoustics*, vol. 67, pp. 1144–1163, 2006.

[15] H. Glotin, F. Caudal, and P. Giraudet, "Whale cocktail party: real-time multiple tracking and signal analyses," *Canadian acoustics*, vol. 36, no. 1, pp. 139–145, 2008.

[16] R. Abeille, Y. Doh, P. Giraudet, H. Glotin, J.-M. Prevot, and C. Rabouy, "Estimation robuste par acoustique passive de l'intervalle-inter-pulse des clics de physeter macrocephalus: méthode et application sur le parc national de Port-Cros," *Journal of the Scientific Reports of Port-Cros National Park*, vol. 28, 2014.

[17] F. Pukelsheim, "The three sigma rule," *The American Statistician*, vol. 48, no. 2, pp. 88–91, 1994.

[18] P. Giraudet and H. Glotin, "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.

[19] P. Almeida, MV. Vieira, Kajin, G. Forero-Medina, and R. Cerqueira, "Indices of movement behaviour: conceptual background, effects of scale and location errors," *Zoologia*, vol. 27, no. 5, 2010.

[20] N. Jaquet, S. Dawson, and L. Douglas, "Vocal behavior of male sperm whales: Why do they click?," *J. of Acous. Soc. of America*, vol. 109, no. 5, pp. 2254–2259, 2001.

[21] Y. Doh, F. Delfour, E. Augier, H. Glotin, C. Graff, and Adam O., "Bottlenose dolphin (T. truncatus) sonar slacks off before touching a non-alimentary target," *Behavioural processes*, 2018.

[22] G. Pavan, "Fundamentals of soundscape conservation," in *Ecoacoustics: The ecological role of sounds*, A. Farina and SH. Gage, Eds. John Wiley & Sons, 2017.