Applied Acoustics 71 (2010) 994–999

Contents lists available at ScienceDirect

Applied Acoustics

journal homepage: www.elsevier.com/locate/apacoust

Whale 3D monitoring using astrophysic NEMO ONDE two meters wide platform with state optimal filtering by Rao-Blackwell Monte Carlo data association

Frédéric Bénard-Caudal [a,*], Pascale Giraudet [b], Hervé Glotin [a]

[a] System & Information Sciences Lab. (LSIS) – UMR CNRS 6168 & Université du Sud-Toulon Var, France
[b] Department of Biology, Université du Sud-Toulon Var, BP 20132, 83957 La Garde Cedex, France

ARTICLE INFO

Article history:
Available online 11 August 2010

Keywords:
Angle of Arrival
Rao-Blackwell
Monte Carlo
Marine mammal localization
Short-spaced hydrophone array
Passive acoustic

ABSTRACT

This paper proposes an algorithm for whale monitoring by passive acoustic using four short-spaced hydrophones. It is successfully tested on the NEutrino Mediterranean Observatory – Ocean Noise Detection Experiment (NEMO ONDE) underwater acoustic platform. In past years, interest in marine mammals has increased, leading to the development of passive acoustic methods of localization which permits to study whales' behavior in deep water (several hundreds of meters) without interfering with the animals. In this paper, we propose a robust Angle of Arrival (AoA) tracking algorithm, which provides an estimation of elevation and azimuth angles in spite of the short distance between hydrophones, which does not allow direct 3D localization. Moreover we show that, in some cases, the time-delay between direct signal and its reflection on the sea surface allows the range estimation, and thus the exact localization of the whale. The classic AoA estimation is performed with a non-linear regression, and it is compared with a more sophisticated algorithm, the Rao-Blackwell Monte Carlo Data Association. We demonstrate that the second method is very robust to the presence of clutter. According to the Pavi 2009 workshop challenge, our algorithm is the only one performing relevant tracks of whale.

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

Processing of Marine mammal signals for passive oceanic acoustic localization is a problem that has recently attracted attention in scientific literature. Motivation for processing marine mammal acoustic signals stems from the difficulty to study animals that spend a large amount of time underwater, and increasing interest in the behavior of endangered species. In [6,10,5], the authors recently developed a robust algorithm for tracking one or more whales in space using a widely-spaced hydrophone array. Our experiment is conducted on a real deep ocean recordings of 5 min duration from the NEMO ONDE platform using a tetrahedron of short-spaced hydrophone array [8]. Those recordings contain series of click sequences of sperm whales (Physeter macrocephalus). Intensive work is currently done in the field [14], and would need automatic processing like the one presented in this paper, especially in order to extend the use of astrophysical detectors to marine mammals' studies [3]. The experiments in this paper consist in estimating and tracking the azimuth α and elevation ϕ of the whale and evaluating the performance of the algorithm. In a first time, we detect each click with a simple energy detector. In a second time, identical clicks are associated in the four hydrophone recordings and the Time Delay Of Arrival (TDOA) is computed for each pair of hydrophones. Finally, the azimuth and elevation angles (and range when possible) are estimated with a classic method, and a particle filtering method. This last method presents more robustness to clutter.

2. Material

We use several dataset files in wav standard, each one containing four channels, 16 bits coding, corresponding to the four hydrophones (H1–H4). The sampling frequency is 96 kHz. Fifteen datasets are available and time duration for each file is 5 min. Most datasets contain sperm whale sounds; one contains other clicks, similar to those emitted by Cuvier's beaked whales. These signals were recorded in 2005 by the CIBRA,[1] at 2000 m depth 25 km offshore Catania (Sicily, Italy), from the NEMO ONDE acoustic module, forming a tetrahedral array about 1 m wide. Hydrophones H1, H2, and H4 lie in the same plane at about 2.5 m from the seabed, and H3 is placed on the top vertex at about 3.2 m from the seabed. Fig. 1 shows the tetrahedral station configuration. The azimuth α, x- and y-axes are relative to the array position. The z-axis origin

* Corresponding author. Tel./fax: +33 (0) 4 941428 24/97.
E-mail addresses: frederic.benard83@gmail.com (F. Bénard-Caudal), giraudet@univ-tln.fr (P. Giraudet), glotin@univ-tln.fr (H. Glotin).

[1] Centro Interdisciplinare di Bioacustica e Ricerche Ambientali (CIBRA), Pavia, Italy.

0003-682X/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.apacoust.2010.05.018

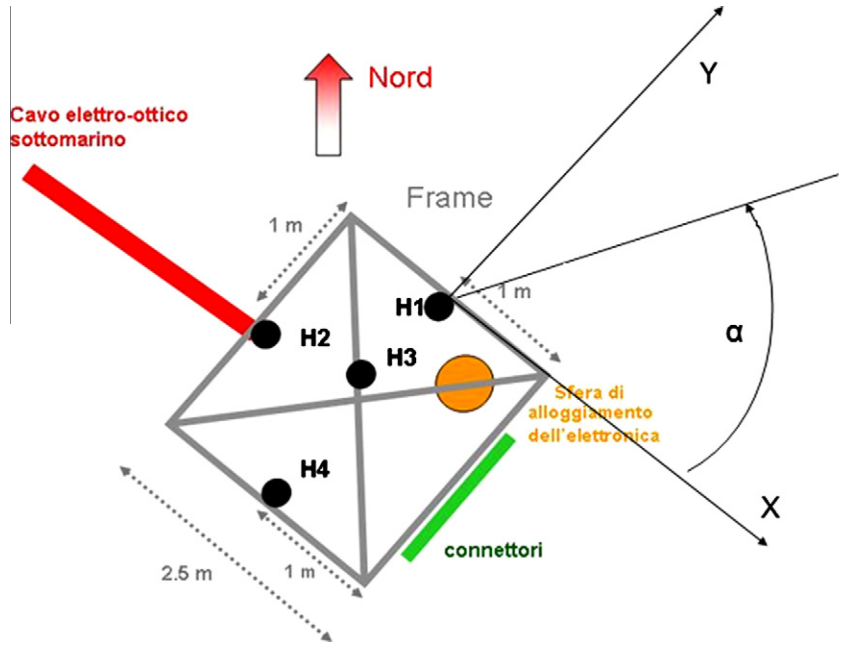

Fig. 1. The NEMO ONDE array projected onto the horizontal plane with the chosen x- and y-axes. The azimuth orientation α is also presented. The z-axis origin is the sea-floor depth, and its positive direction is toward the sea surface. The elevation ϕ is the angle with the horizontal plane.

is the sea-floor depth, and its positive direction is toward the sea surface. The elevation ϕ is 0° in the horizontal (sea-floor) plane and its value is 90° in the vertical plane (positive from the floor toward the surface). Our concern in this paper is the dataset 3, which contains one emitting whale (dataset provided by the 4th DCL workshop, recorded on August 9, 2005 at 9 am).

3. Signal filtering and source coordinates estimation

3.1. Signal filtering

A sperm whale click is a transient increase of signal energy 3–20 ms long (Fig. 2 shows the relative amplitude of the click). The clicks are generally emitted every 1 s forming series called trains. Each click is generally followed by its reflection on the sea surface. New analysis algorithms are being developed to maximize the SNR ratio and to track the movements of impulsive acoustic sources to reveal the movement of sperm whales whilst in the detection range. In this paper, the detection of clicks is performed with an energy detector on octave bands and an adaptative threshold. Sliding 20 ms windows are used for the detection. We can write for the energetic operator:

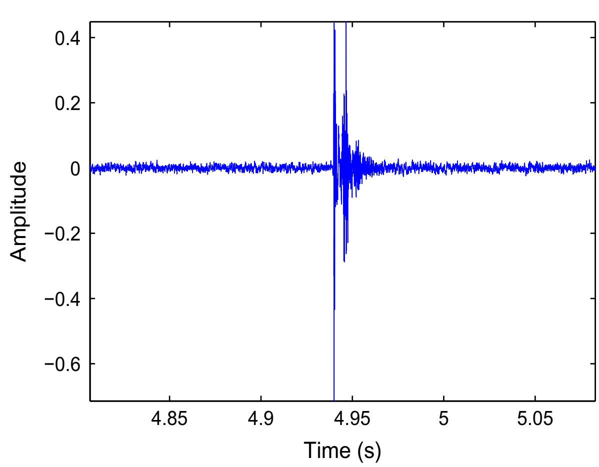

Fig. 2. Example of signal for a single click without reflection.

$$\psi = \sum_{n=1}^{N} s(n)^2, \tag{1}$$

where ψ is the detection value, and s is a 20 ms window (length $N = 1920$). The first octave is centered at 100 Hz. The threshold is constant while detection occurs, but when there is no click detected, it is updated in function of the value of the current window energy (10% over the energy value of the window). This method allows to adapt the threshold to both long-term and short-term variations of ocean noise level (due for instance to meteorological conditions and engine noises, respectively) and to the decrease of click energy when the whale signal grows fainter.

Since the hydrophones are very close one to another (a few meters), only detections occurring at the same approximate time on the four hydrophones (±5 ms) are validated and associated. Thus, for each click detected, we obtain four times of arrival (one for each hydrophone).

TDOAs are computed for each pair of hydrophones. Each pair (i,j), $i = 1$, $j = 2{:}4$ is cross-correlated to estimate the TDOA value $\mathrm{TDOA}(i,j) = t_i - t_j$, where t_i and t_j are the times of arrival of the same click on hydrophone i and j. Results obtained for dataset 3 are presented in the Fig. 3.

3.2. Angle of Arrival (AoA) estimation

Three-dimensional tracking of whales was previously based on a geometric method of calculation of (x,y,z) coordinates from three large enough TDOAs (a few seconds) recorded by a widely-spaced hydrophone array [9]. In the present study, the short distances between hydrophones drastically reduce the TDOAs (they cannot exceed 1 ms) and prevent us from applying the same method: (x,y,z) coordinates would be too strongly affected by the inaccuracies in the hydrophone positions (about 5 cm) and the TDOA values (about 0.01 ms). We therefore replace direct calculation of (x,y,z) coordinates with an estimation the direction of the whale.

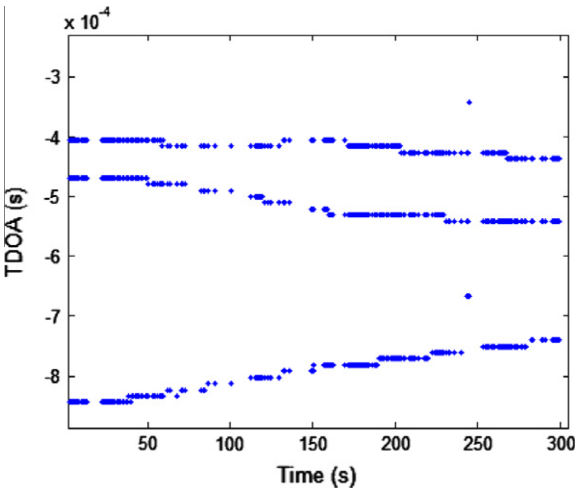

Fig. 3. Results of TDOA computation for dataset 3. As we have four hydrophones, three TDOAs are independent. The low value (less than 1 ms) is due to the short distance between hydrophones (a few meters). The stair shape stems from the sampling frequency (96 kHz) which limits the TDOA precision to 0.01 ms.

Estimation of the TDOA between each hydrophone pair of the dataset allows the calculation of the click Angle of Arrival (AoA), and the tracking of the source movements. Knowing the hydrophone coordinates, the relation between the TDOA and the AoA is:

$$c * \mathrm{TDOA}(i,j) = c * (t_i - t_j) = \vec{\omega} \cdot \overrightarrow{H_iH_j}, \quad (2)$$

where $c = 1500$ m/s is the sound velocity, and $\overrightarrow{H_iH_j}$ is the vector coordinate binding the hydrophone i and j (vector coordinate $\vec{H_j}$ minus vector $\vec{H_i}$), and

$$\vec{\omega} = [cos(\alpha)cos(\phi), sin(\alpha)cos(\phi), sin(\phi)], \quad (3)$$

with $\|\vec{\omega}\| = 1$.

Considering a constant sound speed (1500 m/s), the AoA were finally calculated from the measured TDOA according to Eq. (2), using a multiple non-linear regression with the Gauss-Newton method (Levenberg–Marquardt) [9,13]. The residuals are approximated using a Chi-square distribution with $Nc - d$ degrees of freedom, noted X^2_{Nc-d}, where Nc is the number of hydrophone couples considered and $d = 3$ the number of unknowns (x,y,z) coordinates. The position is accepted if the residual is inferior to a threshold A, that is calculated solving $P = prob(X^2_{Nc-d} > A)$ with $P = 0.01$ (we keep 99% of the estimates).

3.3. Range estimation

In some cases, when the surface reflections are clearly associated to individual clicks, it is possible to exactly locate the animal, i.e. to know its distance and depth, instead of having only azimuth and elevation information. Such information is important to assess the real detection range of the ONDE station and to improve the tracking of the animals. Knowing the time-delay $\tau = \Delta/c$ (c is the sound velocity) between a click and its surface reflection, we can calculate Δ, the difference between the direct and reflected sound path lengths, and then estimate the horizontal range of the whale. Reflections are hand-labeled on a hydrophone (here we chose H1) and the time-delays between direct clicks and reflections are computed (a total of 20 delays τ are labeled). Horizontal range d is also function of the elevation ϕ:

$$d = \frac{(4h^2 - \Delta^2) * \cos(\phi)}{2\Delta - 4h * \sin(\phi)}, \quad (4)$$

where h is the depth of the hydrophone. Knowing d, ϕ, and α, the computation of the (x,y,z) coordinates of the whale is obvious. The Fig. 4 is the range results from the dataset 3, after the computation of the elevation, azimuth, and time-delays between clicks and their surface reflection.

3.4. Review of the results and performances

In the Figs. 5 and 6, we show, respectively, the AoA and the localization results from the dataset. The stair shape in the AoA and the track plots is the consequence of the sampling frequency f_e which produces a uniform noise in the TDOA estimate with the cross-correlation, with a variance equal to $\left(\frac{1}{f_e\sqrt{12}}\right)^2$ second. This noise can also be observed in the TDOA plot (Fig. 3).

The confidence regions are computed for the dataset with Monte Carlo method. The noise variance for the simulation is $\left(\frac{1}{f_e\sqrt{12}}\right)^2$. The regions are only calculated for the elevation and azimuth. The ellipses in the Cartesian coordinates are not a concern with the performance evaluation because of the dependency with the range estimation, and thus it could give approximate results. The ellipses maxima for the elevation is 1.8° and 1.6° for the azimuth, we can consider that this performance is good. But we should also consider the Cramér-Rao Lower Bound given the array configuration. Indeed, these values of elevation and azimuth performance are reasonable, but this will also depend on the range of the emitting whale. One degree of precision could be revealed to be a huge inaccuracy in the Cartesian coordinates.

4. State optimal filtering with the RBMCDA

The question one could ask is 'What if the records produce noisy TDOA measurement?' To tackle this issue, a method consists in using the Rao-Blackwellized Monte Carlo Data Association (RBMCDA). In a nutshell, we solve the tracking part with an Extended Kalman Filter (EKF) filtering [16,11], and the data association (whale or clutter association) with a particle filter. This can be achieve thanks to the Rao-Blackwell theorem [4,1]. This method has been exhaustively described in articles such [15]. Its principle is briefly recalled below. In multiple target tracking (MTT) we are estimating the states of several targets (whales) through measurements. If we know the targets which produce each measurement the problem reduces to single target tracking and we can use a standard filtering algorithm (e.g. Kalman or Extended Kalman Filter) for estimating the states of the targets independently.

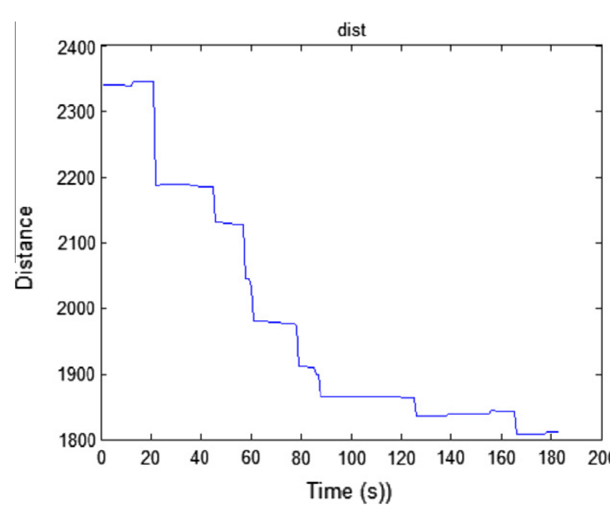

Fig. 4. Range d computed during the whole dataset 3. The stair shape is due to the low amount of τ hand-labeled (a total of 20), and is different from the one engendered by the TDOA variance.

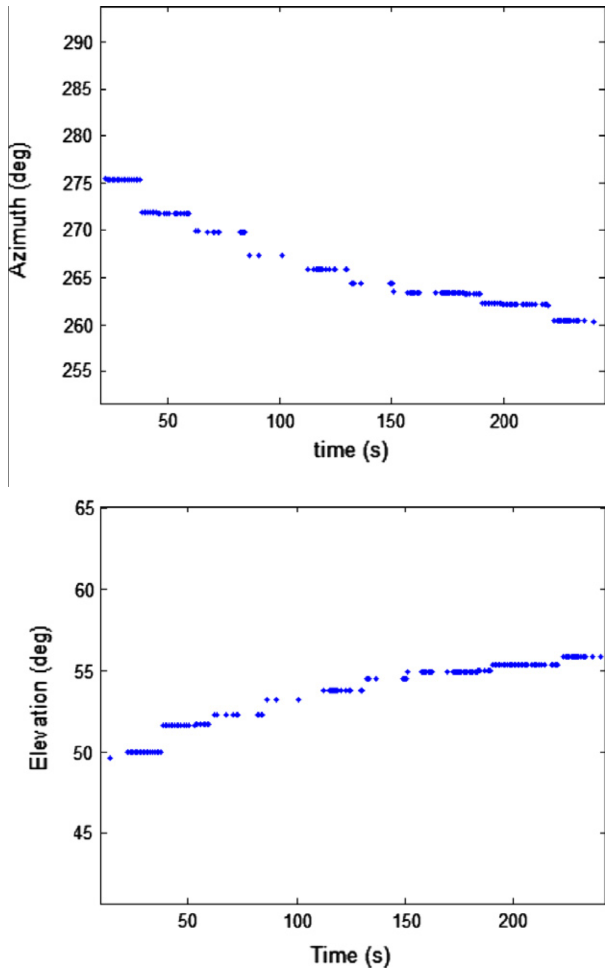

Fig. 5. Azimuth (top) and elevation (bottom) results from dataset 3. The AoA variations are low, and fit with a sperm whale behavior.

Unfortunately, such knowledge is very rarely available in practice, and in many cases some measurements might be also due to clutter, so one is forced to solve the problem of data association. After we have associated measurements to the whale or clutter, we can apply the standard filtering techniques for estimating the target states. In the RBMCDA framework the whale states, data associations and the births and deaths of targets are treated as hidden stochastic processes, which are observed through noisy and indirect measurements. The joint tracking and data association is formulated as a Bayesian estimation problem and the inference is done with Sequential Monte Carlo (SMC) methods (also referred as particle filtering methods) [4,1,7], which give Monte Carlo approximations to the posterior distributions. Furthermore, the accuracy and efficiency of the algorithm are enhanced with the application of Rao-Blackwell algorithm, which allows us to integrate over the target states and use SMC only for estimating the data associations.

4.1. Optimal filtering of the whale states

The system model for one whale is (by discretization) in the space of states:

$$x_k \sim p(x_k|x_{k-1}), \tag{5}$$

$$y_k \sim p(y_k|x_k), \tag{6}$$

where x_k is the whale state at step k; in our case, $x_k = (\alpha_k, \phi_k, \dot{\alpha}_k, \dot{\phi}_k)$. y_k is the measurement at step k, for example the TDOA. $p(x_k|x_{k-1})$ is the dynamic model of the whale. $p(y_k|x_k)$ is the conditional measurement likelihood given the current state. The goal of optimal filtering is to compute the estimate of the current state, using the measurements collected until the current step, i.e., we want to compute recursively the marginal posterior distribution:

$$p(x_k|y_{1:k}). \tag{7}$$

The non-linear system can be express in our case:

$$p(x_k|x_{k-1}) = N(x_k|A_{k-1}x_{k-1}, Q_{k-1}), \tag{8}$$

$$p(y_k|x_k) = N(y_k|h(x_k), R_k). \tag{9}$$

For the linear form of this system, the common algorithm is the Kalman filter [12]. Here we chose to use the EKF thanks to the locally linearization. After the state estimation, it is possible to smooth the states, by calculating the posterior marginal distribution:

$$p(x_k|y_{1:T}), \tag{10}$$

where $T > k$. Those smoothing can be achieved with the Kalman smoother or the Rauch–Tung-Smoother (RTS) [16,2].

4.2. RBMCDA framework with one whale and clutter

We saw that the state at time k is computed with the EKF and then smoothed. The data association is performed with the Monte Carlo data association. We have two associations possible, the whale and the clutter. Birth and death of those associations are modelized with specific distribution. The discretized dynamic model we chose is the rectilinear movement which can be written:

$$x_k = \begin{pmatrix} 1 & 0 & \Delta t & 0 \\ 0 & 1 & 0 & \Delta t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} x_{k-1} + q_{k-1}, \tag{11}$$

where q_{k-1} is a discrete white Gaussian noise. The time step Δt can be chosen at 5 s for example. The equation of measure is:

$$y_k = H(x_k) + r_k, \tag{12}$$

where y_k is the measure vector (the TDOA at step k) and r_k is a Gaussian or uniform noise. H is the measurement transfer function. The system is solved with the RBMCDA, the state vector is computed with EKF and data association is simulated.

4.3. Review of the results and performances

In the Fig. 7, we used the TDOA previously computed (Fig. 3) and we added simulated clutter with a uniform distribution between $-1.5 * 10^{-3}$ to $1.5 * 10^{-3}$s. The number of clutter for each TDOA follows a Poisson distribution with a parameter $\lambda = 1$. We used the TDOA with the synthetic noise, and ran the RBMCDA algorithm on it. The results with the RTS are in the Fig. 8. The clutter is totally eliminated, the algorithm perfectly discriminates the whale's emissions association and false alarm. Moreover, the smoother prevents the stair shape obtained with the classic regression and due to the TDOA variance. We chose not to show the (x,y,z) localization results because the results are nearly the same, except for the TDOA variance sensitivity that is smoothed with the RBMCDA.

The confidence regions are not post-computed, but directly calculated during the elevation and azimuth estimation. Indeed the particle structure allows one to compute the mean and variance of the particle clouds and gives the performance of the current estimation. The value for the azimuth is 1.4° and 1.5° for the elevation. The performance is quite the same than with the non-linear regression method.

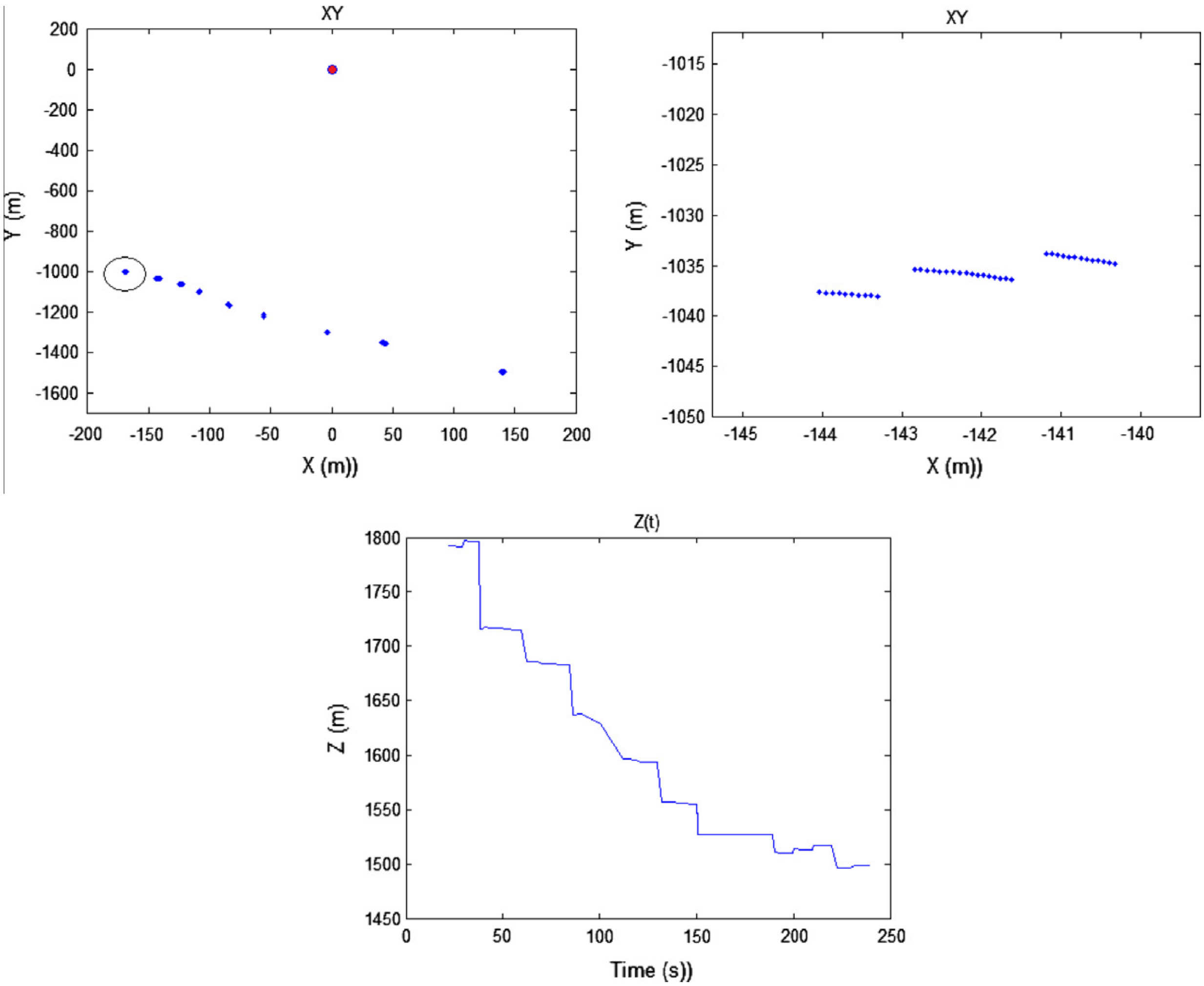

Fig. 6. Plan view of the track (top) and diving profile (bottom). We can see two different stair shapes in the track. The first one corresponds to the largest gaps in the track (about 20 m) and is a consequence of the range estimation about every 30 s. The constant range we consider during this time interval causes these artifacts in the track. The second stair shape is observed in the zoomed area (underlines with a circle in the top left figure) in the figure on the top right. This is the variance in the TDOA estimates that does not allow a better precision (which is quite good although, about 2 m). Whatsoever, the two shapes are from different causes. The mean speed is 6 km/h, and fits well the behavior of a sperm whale.

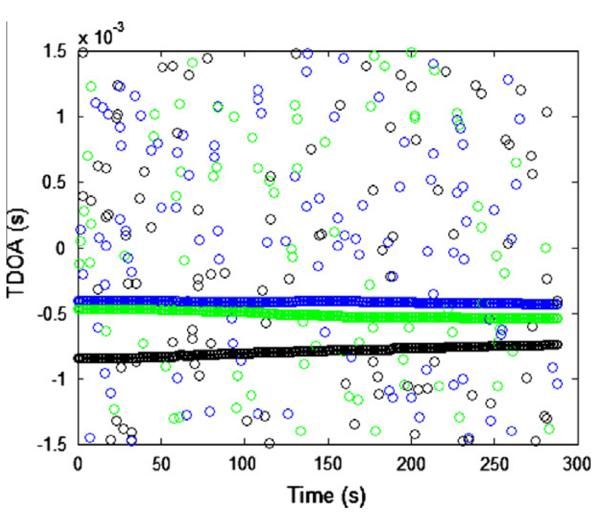

Fig. 7. Real TDOAs with additional simulated clutter. See the Fig. 3 for the real TDOA used to generate this figure.

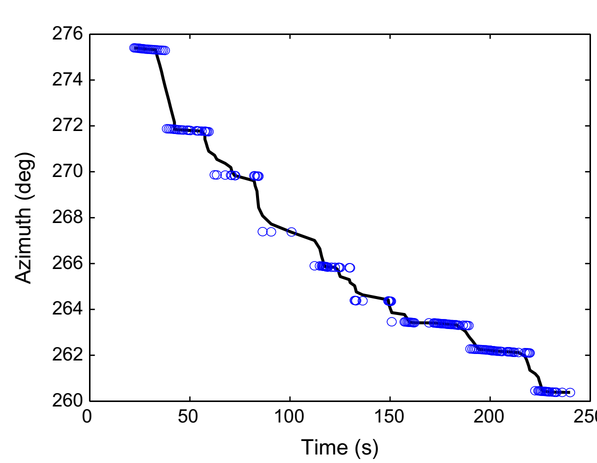

Fig. 8. Azimuth results with the RBMCDA method (bold line) compared to the classic method results (Fig. 5) with the 'o' symbols.

5. Discussion and conclusion

We defined in this article a real-time tracking algorithm to locate clicking whale with the NEMO ONDE short-baseline platform. First, two different methods are proposed to estimate the AoA. A classic method with a non-linear regression and the RBMCDA algorithm are compared. Moreover, thanks to the presence of reflected signal in the recordings, ranging is estimated and a complete localization of a Physeter macrocephalus whale in a Cartesian space is presented. The normal speed evaluated from the trajectory is a partial validation of this experiment. To evaluate the RBMCDA robustness to noise, we added synthetic noise in the TDOA. The results show that the clutter and the real emissions are well discriminated, which is not the case with the classic method that presents some spurious azimuth and elevation results. The confidence regions for the azimuth and elevation are about 2° for the two methodologies with an advantage for the RBMCDA which reduces the variance of the estimates thanks to the smoother. Our algorithm is working for all clicking marine mammal. Future works will consist in expanding our algorithm to several simultaneous emitting whales in the RBMCA framework. Additional work on the performance will have to be done, particularly with the Cramér-Rao bound considering the array configuration and thus the state observability (azimuth and elevation in our case). Also, the hydrophone position errors evaluation on the performance, could lead to some inaccuracy. Finally, a validation with sightings would be the last step to confirm the performance of the algorithm. Then it shall be usefull for continuous online whale monitoring.

Acknowledgments

We thank Pr. Pavan, president of the Centro Interdisciplinare di Bioacustica e Ricerche Ambientali (Pavi University, Italy) for recording and distributing the dataset. We thank Sébastien Paris from LSIS DYNI team for his discussion on RBMCDA. This research was conducted within the international sea *pôle de compétitivité* at Toulon-France, and is a part of the project ''Platform for Integration of Multimodal information for Cetology (PIMC, http://glotin.univ-tln.fr/PIMC)''. This work is founded by the ''Conseil régional Provence-Alpes-Côte d'Azur'' France, and partly by Chrisar Software Inc.

References

[1] Doucet A, de Freitas N, Gordon N. Sequential Monte Carlo methods in practice. Springer; 2001.
[2] Gelb A. Applied optimal estimation. The MIT Press; 1974.
[3] Adam O, Glotin H. Passive acoustic storey of the ANTARES neutrino detector for real-time cetacean detection, localization and behavior studies. In: PASSIVE'08, IEEE explorer; 2008. 6p.
[4] Ristic B, Arulampalam S, Gordon N. Beyond the Kalman filter. Artech; 2004.
[5] Bénard-Caudal F, Glotin H. Automatic inter-click-interval (ICI) and behavior estimation for one emitting sperm whale. In: Internet wkp PASSIVE 08 IEEE, in IEEE explorer; 2008.
[6] Bénard-Caudal F, Glotin H. Stochastic matched versus Teager–Kaiser–Mallat filters for tracking simultaneous clicking whales. In: Internat wkp PASSIVE 08 IEEE, in IEEE explorer; 2008.
[7] Kitagawa G. Monte Carlo filter and smoother for non-Gaussian nonlinear state space models. J Comput Graph Stat 1996;5:1–25.
[8] Pavan G, La Manna G, Zardin F, Internullo E, Kloeti S, Cosentino G, et al. Short term and long term bioacoustic monitoring of the marine environment. Results from NEMO ONDE experiment and way ahead. Computational bioacoustics for assessing biodiversity. In: Proceedings of the International Expert meeting on IT-based detection of bioacoustical patterns; 2008. p. 7–14.
[9] Giraudet P, Glotin H. Real-time 3d tracking of whales by echo-robust precise TDOA estimates with a widely-spaced hydrophone array. Appl Acoust 2006;67:1106–17.
[10] Glotin H, Bénard-Caudal F, Giraudet P. Whales cocktail party: a real-time tracking of multiple whales. Int J Can Acoust 2008;36(1):141–7.
[11] Jazwinski AH. Stochastic processes and filtering theory. Academic Press; 1970.
[12] Kalman RE. A new approach to linear filtering and prediction problems. Trans ASME J Basic Eng 1960;82:34–45.
[13] Marquardt Donald W. An algorithm for least-squares estimation of nonlinear parameters. SIAM J Appl Math 1963;11(2):431–41.
[14] Nicola Nosengo. The neutrino and the whale. Nature 2009;462:2.
[15] Sarkka S, Vehtari A, Lampinen J. Rao-blackwellized Monte Carlo data association for multiple target tracking. In: Proceedings of the seventh international conference on information fusion, vol. 1; 2004. p. 583–90
[16] Bar-Shalom Y, Li XR, Kirubarajan T. Estimation with applications to tracking and navigation. Wiley Interscience; 2001.