MULTIPLE REAL-TIME 3D TRACKING OF SIMULTANEOUS CLICKING WHALES USING HYDROPHONE ARRAY AND LINEAR SOUND SPEED PROFILE

Frédéric Caudal, Hervé Glotin

USTV - System & Information Sciences Laboratory (LSIS - UMR CNRS 6168)
BP 20132 - 83957 La Garde Cedex - France.
frederic.caudal@gmail.com, glotin@univ-tln.fr

ABSTRACT

This paper provides a real-time passive underwater acoustic method to track multiple emitting whales using four or more omni-directional widely-spaced bottom-mounted hydrophones. After a non-parametric Teager-Kaiser-Mallat signal filtering, rough Time Delays Of Arrival are calculated, selected and filtered, and used to estimate the positions of whales for a constant or linear sound speed profile. The complete algorithm is tested on real data from the NUWC[1] and the AUTEC[2]. Our model is validated by similar results from the US Navy and Hawaii univ labs in the case of one whale, and by similar whales counting from the Columbia univ. ROSA lab in the case of multiple whales. At this time, our tracking method is the only one giving typical speed and depth estimations for multiple (5) emitting whales located at 1 to 5 km from the hydrophones.

Index Terms— Delay estimation, marine animals, acoustic tracking, real-time systems, acoustic propagation

1. INTRODUCTION

Processing of Marine Mammal (MM) signals for passive oceanic acoustic localization is a problem that has recently attracted attention in scientific literature and in some institutes like the AUTEC and the NUWC. Motivation for processing MM signals stems from increasing interest in the behavior of endangered MM. One of the goals of current research in this field is to develop tools to localize the vocalizing and clicking whale for species monitoring. In this paper we propose a low cost time-domain tracking algorithm based on passive acoustics. The experiments of this paper consist in tracking an unknown number of sperm whales (Physeter catodon). Clicks are recorded on two datasets of 20 and 25 minutes on a open-ocean widely-spaced bottom-mounted hydrophone array. The output of the method is the track(s) of the MM(s) in 3D space and time. In Section 2 we briefly review studies of source separation methods and the main characteristics of MM signals. In Section 3 we propose a real-time algorithm for MM transient call localization. In Section 4 we show and compare results of tracks estimates with results from specialized teams.

[1]Naval Undersea Warfare Center of the US Navy

[2]Atlantic Undersea Test & Evaluation Center - Bahamas

2. PROBLEM FORMULATION

This papers deals with the 3D tracking of MM using a widely-spaced bottom-mounted array in deep water. It focuses on sperm whale clicks; detection and classification are not a concern. There were previous algorithms developed in the state of art [1, 2, 3, 4] but none are able to have satisfying results for multiple tracks. Most of them are far from being real-time. The main goal is to build a robust and real-time tracking model, despite ocean noise, multiple echoes, imprecise sound speed profiles, an unknown number of vocalizing MM, and the non-linear time frequency structure of most MM signals. Background ocean noise results from the addition of several noises: sea state, biological noises, ship noise and molecular turbulence. Propagation characteristics from an acoustic source to an array of hydrophones include multipath effects (and reverberations), which create secondary peaks in the Cross-Correlation (CC) function that the generalized CC methods cannot eliminate.

3. MATERIAL

The signals are records from the ocean floor near Andros Island - Bahamas (Tab.1) , provided with celerity profiles and recorded in March 2002. Datasets are sampled at 48 kHz and contain MM clicks and whistles, background noises like distant engine boat noises. Dataset1 (D1) is recorded on hydrophones 1 to 6 with 20 min length while dataset2 (D2) is recorded on hydrophones 7 to 11 with 25 min length. We will use a constant sound speed with $c = 1500ms^{-1}$ or a linear profile with $c(z) = c_0 + gz$ where z is the depth, $c_0 = 1542ms^{-1}$ is the sound speed at the surface and $g = 0.051s^{-1}$ is the gradient. Sound source tracking is performed by continuous localization in 3D using Time Delays Of Arrival (T) estimation from four hydrophones.

Table 1. Hydrophones positions

Datasets	Hydros	Dist to barycenter (m)	X (m)	Y (m)	Z (m)
D1	H 1	5428	18501	9494	-1687
	H 2	4620	10447	4244	-1677
	H 3	2514	14119	3034	-1627
	H 4	1536	16179	6294	-1672
	H 5	3126	12557	7471	-1670
	H 6	4423	17691	1975	-1633
D2	H 7	1518	10658	-14953	-1530
	H 8	4314	12788	-11897	-1556
	H 9	2632	14318	-16189	-1553
	H 10	3619	8672	-18064	-1361
	H 11	3186	12007	-19238	-1522

3.1. Signal filtering

A sperm whale click is a transient increase of signal energy lasting about 20 ms (Fig.1-a). Therefore, we use the Teager-Kaiser (TK) energy operator on the discrete data:

$$\Psi[x(n)] = x^2(n) - x(n+1)x(n-1), \qquad (1)$$

where n denotes the sample number. An important property of TK is that it is nearly instantaneous given that only three samples are required in the energy computation at each time instant. Considering the raw signal $s(n)$ as:

$$s(n) = x(n) + u(n),$$

where $x(n)$ is the signal of interest (clicks), $u(n)$ is an additive noise defined as a process realization considered wide sense stationary (WSS) Gaussian during a short time, by applying TK to $s(n)$, $\Psi[s(n)]$ is:

$$\Psi[s(n)] \approx \Psi[x(n)] + w(n),$$

where $w(n)$ is a random gaussian process [5]. The output is dominated by the clicks energy. Then, we reduce the sampling frequency to 480Hz by the mean of 100 adjacent bins to reduce the variance of the noise and the data size. We apply the Mallat's algorithm with the Daubechies wavelet (order 3). We chose this wavelet for its great similarity to the shape of a decimated click [1, 2]. The signal is denoised with a universal thresholding defined as $D(u_k, \lambda) = sgn(u_k)max(0, |u_k| - \lambda)$, with u_k the wavelet coefficients, $\lambda = \sqrt{(2log_e(Q))}\sigma_N \sigma_{\tilde{N}}$, and Q the length of the signal resolution level to denoise. The noise standard deviation σ_N is calculated on each 10s window on the raw signal with a maximum likelihood criterion. $\sigma_{\tilde{N}}$ is the standard deviation of the wavelet coefficients on a resolution level of a generated, reduced and 0-mean Gaussian noise. This filtering step is very fast without any parameter. Fig.1-c and 1-f are the filtered signals on single (Fig.1-b) and multiple (Fig.1-d) emitting MMs.

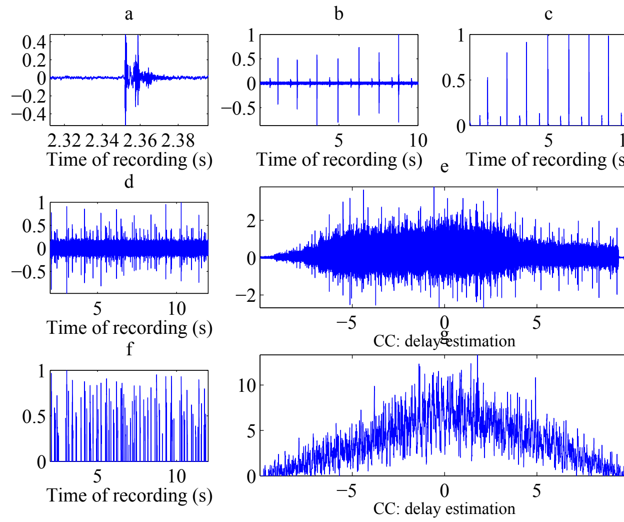

Fig. 1. (a): detail of a click. (b): raw signal (D2) of hydrophone 7 (H7) during the first 10s of recording, containing 7 clicks and their echoes. (c): (b) after filtering. (d): raw signal (D1) of H3 during the first 10s of recording showing multiple emissions. (e): CC between (d) and corresponding raw signal chunk of H1. (f): (d) after filtering. (g): idem than (e) but with (f).

3.2. Rough $\widetilde{T}$ estimation

First, T estimates are based on MM click realignment only. Every 10s, and for each pair of hydrophones (i, j), the difference between times t_i and t_j of the arrival of a click train on hydrophones i and j is referred as $T(i, j) = t_j - t_i$. Its estimate $\widetilde{T}(i, j)$ is calculated by CC of 10s chunks (2s shifting) of the filtered signal for hydrophones i and j [1, 2]. We keep the 35 highest peaks on each CC to determine the corresponding $\widetilde{T}(i, j)$. The filtered signals give a very fast rough estimate of T (precision $\pm$ 2ms). Fig.(1.e) shows the CC with the raw signal and (1.g) with the filtered signal. Without filtering, CC generates spurious delays estimates and the tracks are not correct. The maximum $\widetilde{T}$ rank (Fig.2) in D1, pitching the source localization, are high among the 35 $\widetilde{T}$ kept in the CC which justifies this number.

3.3. $\widetilde{T}$ selection and localization with a constant profile

Each signal shows echoes for each click (Fig.1 b), maybe due to the reflection of the click train off the ocean surface or bottom or different water layers. We use a method based on autocorrelation [1, 2] to eliminate it. Then, thanks to the $\widetilde{T}$ transitivity system described in [1] we keep $\widetilde{T}$ triplets coming from the same source. Finally, thanks to the measured delays and an acoustic model based on a constant sound speed profile, the least squares cost function determines the MM positions using a multiple non linear regression with Gauss-

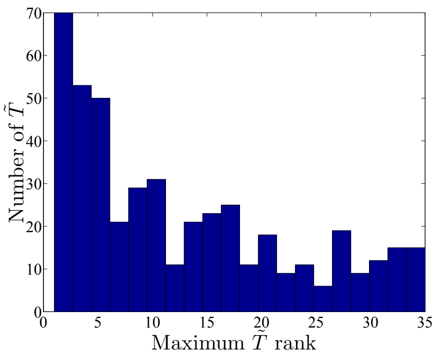

Fig. 2. Maximum $\widetilde{T}$ rank histogram in the CC for each triplet

Newton method (Levenberg-Marquardt) [1]. The residuals are approximatively following a Chi-square distribution with $Nc - d$ degrees of freedom, noted X^2_{Nc-d}, Nc is the number of hydrophones couples considered and d the number of unknowns, here 3 (x, y, z). The position is accepted if the residual is inferior to a threshold x, That is calculated solving $P = prob(X^2_{Nc-d} > x)$ with $P = 0.01$ (we keep 99% of the estimates).

3.4. Source localization with a linear sound speed profile

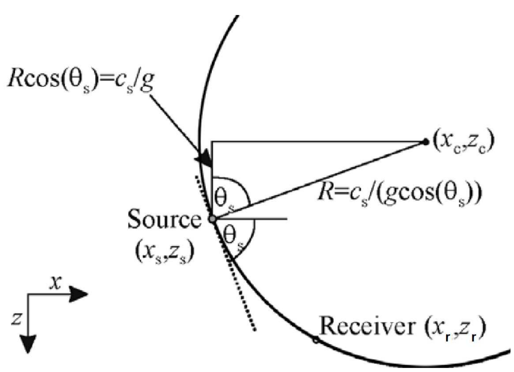

Fig. 3. Geometry for a source and receiver in a linear profile

It is well known that the ray paths in a medium with linear sound speed profile are arcs of circles and further the radius of the circle can be computed (Fig.3). c_s is the sound speed at the source, θ_s is the launch angle of the ray at the source, measured relatively to the horizontal. From the geometry shown in Fig.3, the center of the circle, (xc, zc), along which the ray path is an arc, is:

$$
\begin{aligned}
x_c &= \frac{x_s + x_r}{2} + \frac{(z_s - z_r)}{2(x_s - x_r)}(z_r - z_s + \frac{2c_s}{g}), \\
z_c &= z_s - \frac{c_s}{g}.
\end{aligned} \tag{2}
$$

For linear sound speed profile the course time τ of the ray is:

$$
\tau = \frac{1}{g}\left\{\log\left(\frac{z_c - z_s}{z_c - z_r}\right) - \log\left(\frac{R + x_c - x_s}{R + x_c - x_r}\right)\right\}. \tag{3}
$$

Using Eqs.(2)-(3) allows one to compute the propagation time from the source to any receiver and then the whale position.

4. RESULTS

For D2 (Fig.4), a constant and a linear sound speed profile were used and the results are similar with the Morrissey's [4] and Nosal's [3] methods. The diving profile underlines a bias of about 50 to 100m between the linear and the constant profiles results, emphasizing the importance of the chosen profiles. Moreover with the linear sound speed, the results are about the same as Morrissey's and Nosal's, who used profiles corresponding to the period and place of the recordings.

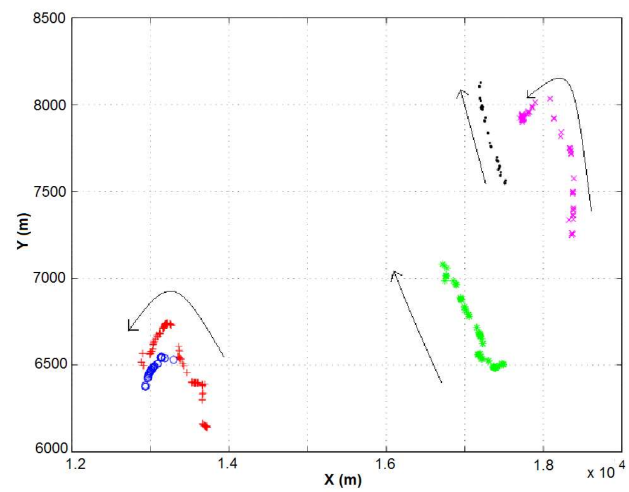

Fig. 6. Plan view in D1. Each symbol corresponds to one of the 5 whales. The arrows stress the directions.

Results on D1 are shown in Fig.6-5 for a linear sound speed profile. We thus localize 5 MM. The confidence regions are computed for the two datasets with a Monte Carlo method. The ellipses maxima (30m) fit with MM length (20m).

5. DISCUSSION AND CONCLUSION

The tracking algorithm presented in this paper is real-time on a standard laptop and works for one or multiple emitting sperm whales. Depth results with constant speed contains a bias errors due to the refraction of the sound paths from the MM to the receivers what a linear speed corrects. An other way to tackle the speed profile issue [6] is to estimate it as a fourth unknown in the regression. Our algorithm has no species dependency as long as it processes all transients. At

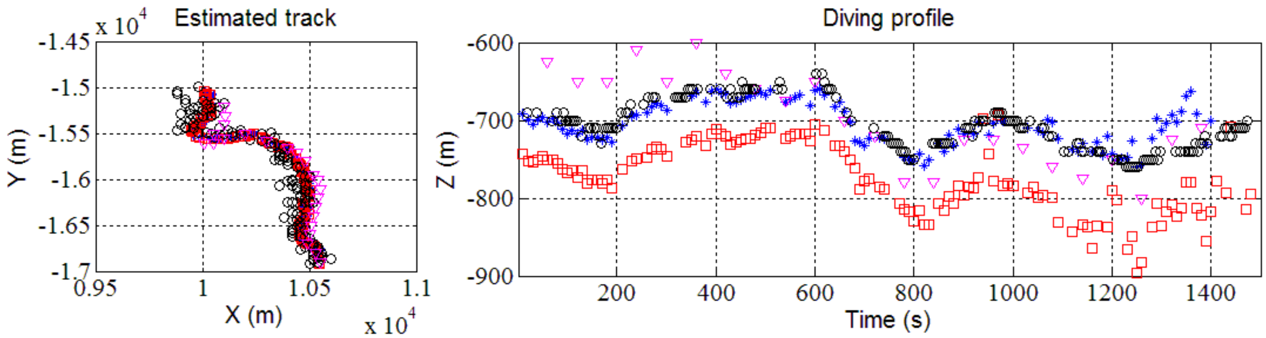

Fig. 4. Plan view and diving profile of the MM in D2, our estimates with a linear ($\ast$) or constant profile ($\square$); and estimates from Morrissey's [4] ($\triangledown$) and from Nosal's [3] methods (o).

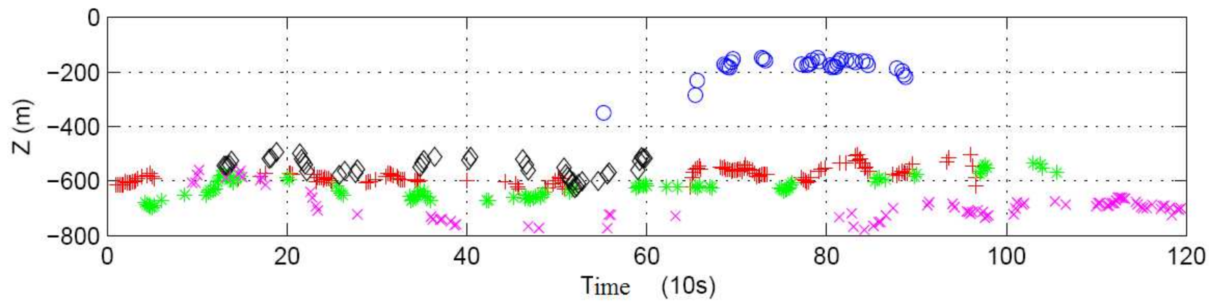

Fig. 5. Averaged diving profile in D1. Each symbol corresponds to one whale in Fig.6 (Bottom to top (+), (o), ($\ast$), (x), ($\diamond$)).

this time, only our algorithm gives results with typical speed and depth estimations for multiple emitting whales. In D2, results indicate that only one sperm whale was present in the area, unless other whales in the area were quiet during the selected 25-min period. Moreover, according to ROSA Lab estimation based on click clustering, averaged number of MM for each 5min chunks on D1 is $[4.3; 5.3; 4; 3.6]$ [7] similar to ours ($[4; 4; 4; 3]$)[8, 6]. Our method provides thus robust online detecting/counting system of clicking MM in open ocean.

6. REFERENCES

[1] 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, pp. 1106–1117, 2006.

[2] P. Giraudet and H. Glotin, "Echo-robust and real-time 3d tracking of marine mammals using their transient calls recorded by hydrophones array," ICASSP, 2006.

[3] E.M. Nosal and L.N. Frazer, "Delays between direct and reflected arrivals used to track a single sperm whale," *Applied Acoustics*, vol. 62, pp. 1187–1201, 2006.

[4] R.P. Morrissey, N. DiMarzio J. Ward, S. Jarvisa, , and D.J. Moretti, "Passive acoustics detection and localization of sperm whales in the tongue of the ocean," *Applied Acoustics*, vol. 62, pp. 1091–1105, 2006.

[5] 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.

[6] H. Glotin, F. Caudal, and P. Giraudet, "Whale cocktail party: Real-time multiple tracking and signal analyses," *Canadian Acoustics*, 2008, In Press.

[7] X.C Halkias and D Ellis, "Estimating the number of marine mammals using recordings form one microphone," ICASSP, 2006.

[8] H. Glotin, F. Caudal, and P. Giraudet, "Localisation, counting and tracking of simultaneous emitting marine mammals using ocean bottom mounted hydrophones array," *3thd int. Wksp on Detection and Localisation of MM using passive Acoustics, Boston*, 2007.