ECHO-ROBUST AND REAL-TIME 3D TRACKING OF MARINE MAMMALS USING THEIR TRANSIENT CALLS RECORDED BY HYDROPHONES ARRAY

Pascale Giraudet [(a, b)] and Hervé Glotin [(a)]

System and Information Sciences Lab - LSIS - UMR CNRS 6168 [(a)],
Department of Biology [(b)],
Univ. Sud Toulon Var, BP 20132 – 83957 La Garde Cedex, France - {giraudet, glotin}@univ-tln.fr

ABSTRACT

Processing marine mammals (MM) signal for 3D tracking has increasing interest for the study of the behavior of endangered species. Practically, real time systems are required, and it has been observed that echoes are very common in underwater signal, and dramatically affect the localization results. Therefore, we develop a robust real-time tracking algorithm, with echo cancellation, processing wide band MM transient signal in time domain, from records of 5 omni-directional widely spaced hydrophones. The complete algorithm is tested on 25 min. of real data from NUWC & AUTEC. The attractive outlines of our method are: it runs 2 times faster than real time, requires small frequency sample data, still estimates similar trajectories than state of the art methods.

1. INTRODUCTION

Processing of marine mammals (MM) signals for passive oceanic acoustic localization is a problem that has recently attracted attention in scientific literature and institutes like AUTEC or NUWC[1]. Motivation for processing MM signals stems from increasing interest in the behavior[2] of endangered MM [1]. The ultimate goal of the current research in this field is to develop tools to analyze the emitted signal for species monitoring. We propose in this paper a time-domain low cost tracking algorithm using passive acoustics. Our method is general for any transient signal, without species dependency. A real application illustrates the efficiency of our method. The organization of this paper is as follows. In section 2 we briefly present previous models of sources separation and the main characteristics of MM signal. In section 3 we propose a new echo robust time-domain algorithm for MM transient analysis. The section 4 gives tracking estimates and running times.

Hydrophone number	X (m)	Y (m)	Z (m)
1	10 658	-14 953	-1 530
2	12 788	-11 897	-1 556
3	14 318	-16 189	-1 553
4	8 672	-18 064	-1 361
5	12 007	-19 238	-1 522

Table 1: 3D position of the 5 hydrophones.

[1] We thank Naval Undersea Warfare Center (NUWC) of the US Navy and Atlantic Undersea Test and Evaluation Center (AUTEC) for providing acoustic data. We thank also O. Adam (Univ. Paris XII) for having organized tracking evaluation workshop session-Monaco 2005 [15].
[2] This study takes place within the international competitivity center on marine researches at Toulon-France, through the project *Platefrom of Integraion of Multimedia data for Cetology* (PIMC).

2. PROBLEM FORMULATION

Propagation characteristics from any acoustic source to an array of hydrophones include multipath effects and reverberations. Thus it seems reasonable to model the propagation characteristics as multivariate FIR filter that has delayed taps. For a real time application it is heavy to run the signal separation filter (like a dynamic recurrent network [8], for which another problem is also the lack of training data on transient MM signals).

A preprocess could consist in recovering sources from their convolutive mixtures, in either time-domain [9], or in frequency-domain [10]. Unfortunately in FFT different permutations occur in each frequency bin [8], which might result in severe performance degradation. In numerous studies [2,5,6] the spectrogram has been used as the interface between MM signal and features extraction, but as all the Cohen's class members, the spectrogram analysis is limited in the case of non-linear time-frequency analysis [3]. It has been observed that the time frequency structures of the MM signals are generally non-linear [4]. Some MM tracking models use a mixture of time and frequency domain clicks and whistle vocalization analysis [6], but finally require multiple computers simultaneously for real time application [5]. Moreover most MM signals are transient clicks at high SNR against ocean noise (Fig.1). Therefore we develop in this paper a time-domain low cost tracking algorithm.

3. MATERIAL AND METHODS

Signals from NUWC and AUTEC are recorded by 5 widely spaced omni-directional hydrophones (Tab. 1), and sampled at 48 KHz. Data set is an extract of 25 min., related to 1 or 2 sperm whale(s) sighted in March 2002 near Andros island-Bahamas. It contains the whale's clicks and whistles, with reverberations and background noises (distant boat noises, tonal at 4,9,13,17,21 KHz from equipments, boat engine at 140Hz).

Our signal processing and sound source(s) localization/tracking follow the simple 5-steps linear algorithm described in the next 5 subsections.

3.1. Signal filtering

The signal is firstly filtered in order to extract clicks from whistles and background noises. Since a click has a mean duration of 100ms, the sampling frequency of the signal is reduced to 480Hz to speed up further processing by summation of 100 adjacent bins after a full wave rectification. Then peaks are extracted by thresholding at twice the mean of this new signal (Fig.1).

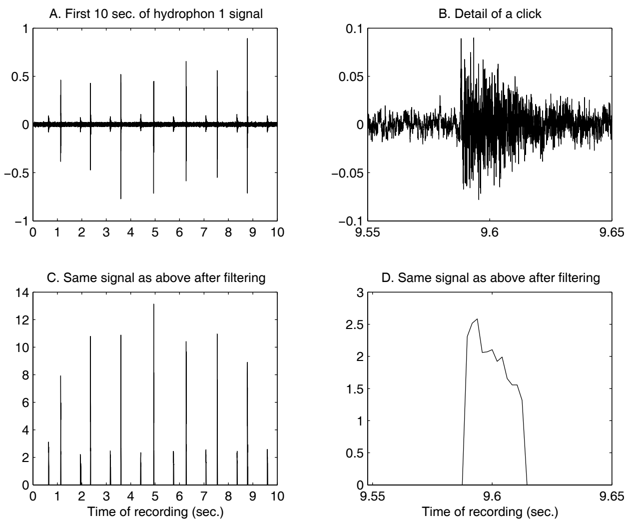

Figure 1: Upper graphs show the raw signal of hydrophone 1 during the first 10 s. of recording, containing 7 clicks and pertaining echoes (A), and the detail of a click (B). Lower graphs show the corresponding signals after filtering (see 3.1).

3.2. TDOA estimations

For each pair of hydrophones signal (i,j), the Time Delay Of Arrival of the clicks TDOA(i,j), are estimated every ten seconds by means of their cross-correlation. Since several MM may have to be localized simultaneously, several TDOA are estimated. We keep only the positions of the 5 highest peaks of the cross-correlation to estimate 5 TDOA with a precision of 2ms: $\{TDOA_n(i,j), n=\{1,..,5\}\}$.

3.3. Echo characterization

Each signal shows a various number of echoes of each click, responsible for the detection of TDOA between a direct signal and an echo. Since TDOA are calculated from clicks position only in this study, echo characterization can be performed by means of a simple auto-correlation (AC) of the filtered signal applied every minute on each hydrophone (Fig. 2). We then use the ratio R_1/R_0 (R_1: 1st maximum of the AC, R_0: signal energy) to detect the first echo. R_1/R_0 has been used for dominant speaker detection in cocktail party [14] or robust fundamental estimation, and robust SNR estimation in ASR systems [13]. We show in Fig. 2 that the 1st echo E(i) of each hydrophone i is visible 0.2 sec. to 1 sec. after the click, and is often followed by others found at multiple delays of the delay of the first echo (Fig 3).

3.4. TDOA selection

Therefore, the 5 TDOA estimated for each pair of hydrophones every 10sec. are selected in order to eliminate these corresponding

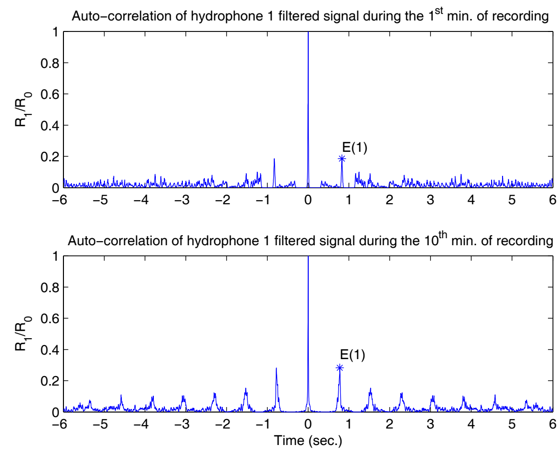

Figure 2: Echo characterization by AC of hydrophone 1, during the 1st (top) or the 10th min (bott.). In both cases, a 1st echo is characterized about 0.8 sec. after the direct signal, and other echoes may happen at multiple delays of the 1st (bott.).

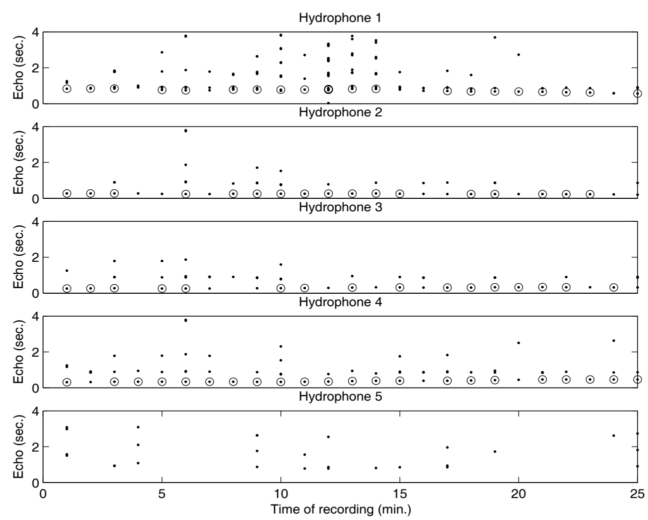

Figure 3: Echo characterization for each hydrophone and each time window. The first echo (circled) corresponds to the highest peak of the cross-correlation (if $R_1/R_0 \geq 0.2$).

to the delay between a direct signal and its echo. In this purpose, for each pair of hydrophone (i,j), all $TDOA_x(i,j)$ verifying one of the following equations are removed:

$$TDOA_x(i,j) - TDOA_1(i,j) = k*E(i), k \in \{1..4\}, x \in \{2..5\},$$
$$TDOA_x(i,j) - TDOA_1(i,j) = - k*E(j), k \in \{1..4\}, x \in \{2..5\}.$$

Furthermore, all TDOA extremely different from those estimated in other 10sec. windows are very unlikely to be correct and are also eliminated by automatic thresholding the TDAO histograms at their mean [5]. TDOA before (resp. after) this selection are shown in Fig 4 (resp. 5).

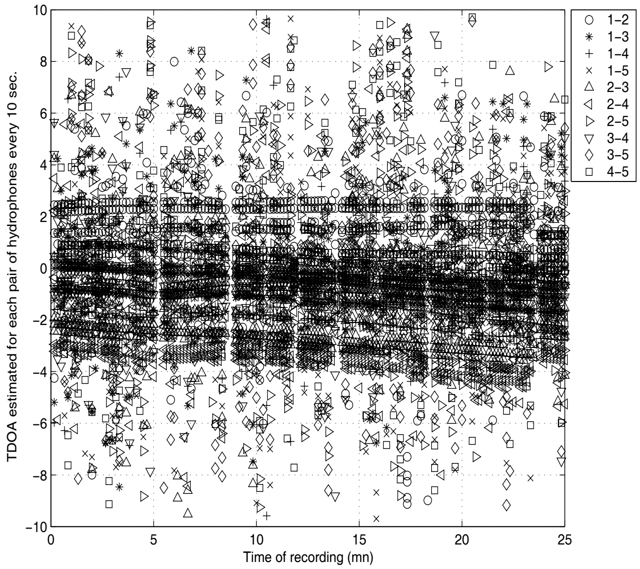

Figure 4: TDOA of each pair of hydrophones calculated every 10sec. by mean of the highest 5 peaks of the cross-correlation of the signals recording during 10sec.

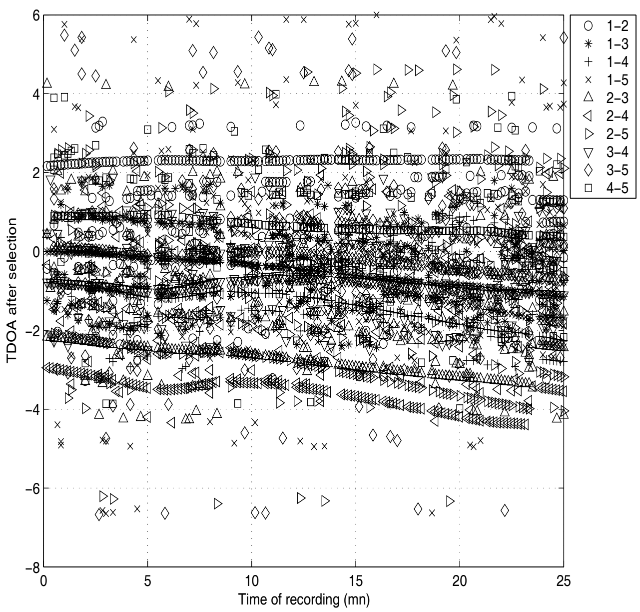

Figure 5: Same TDOA as in Fig. 4 after elimination of the echoes such as explained in subsection 3.4.

3.5. Quadruplets selection and source localization

Once the TDOA of each pair of hydrophones are selected, the remaining TDOA are combined every 10sec. in order to find the 4 hydrophones whose TDOA correspond to the same source. We consider that any quadruplet of hydrophones (i,j,k,h) can be used for localization with the TDOA (u,v,w,x,y,z) if the following equations are verified :

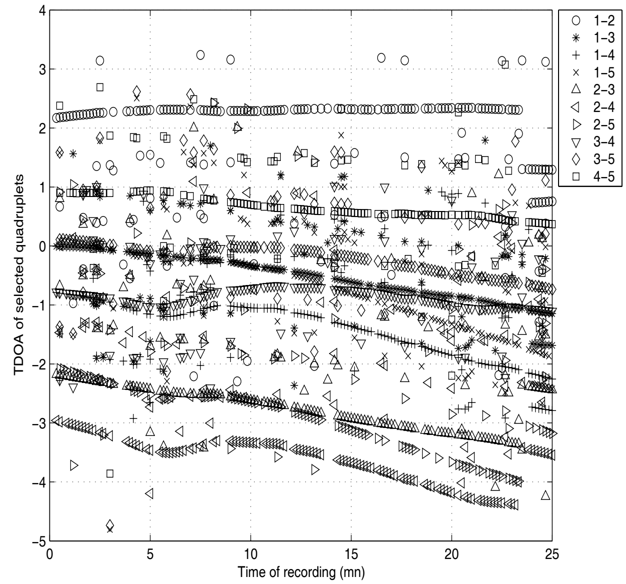

Figure 6: TDOA of the pairs of hydrophones belonging to the quadruplets selected by the method described in subsection 3.5.

$$TDOA_u(i,j)+TDOA_v(j,k) = TDOA_w(i,k) \pm 4 \text{ ms},$$
$$TDOA_u(i,j)+TDOA_x(j,h) = TDOA_y(i,h) \pm 4 \text{ ms},$$
$$TDOA_w(i,k)+TDOA_z(k,h) = TDOA_y(i,h) \pm 4 \text{ ms},$$
$$TDOA_v(j,k)+TDOA_z(k,h) = TDOA_x(j,h) \pm 4 \text{ ms}.$$

Then, in order to localize the signal source S, the TDOA $\{TDOA_u(i,j) \quad TDOA_w(i,k) \quad TDOA_y(i,h)\}$ of each selected quadruplet of hydrophones (i,j,k,h) are used to solve the following equations, by non-linear least squares method [11] :

$$Q(S,i)-Q(S,j) = TDOA_u(i,j) * 1500,$$
$$Q(S,i)-Q(S,k) = TDOA_w(i,k) * 1500,$$
$$Q(S,i)-Q(S,h) = TDOA_y(i,h) * 1500,$$

where Q(a,b) is the euclidian distance between a and b in 3D space. TDOA of these quadruplets are shown in Fig. 6.

4. RESULTS

4.1. Tracking estimates in 3 dimensions

Results of the whale localization are plotted in Fig. 7. We only keep position estimates for which LMS error < 100. A simple analysis of position continuity generates one trajectory similar to a sperm whale trajectory: regular speed (around 9 km/h), diving to -500m in 3’30’’ (9 km/h).

4.2. Comparison to original signal analysis

For each selected positions (Fig. 7), we can rerun the TDAO of each corresponding quadruplet under original signal (48KHz). But this is not useful because it doe not significantly change position estimates (see histograms of position differences in Fig. 8).

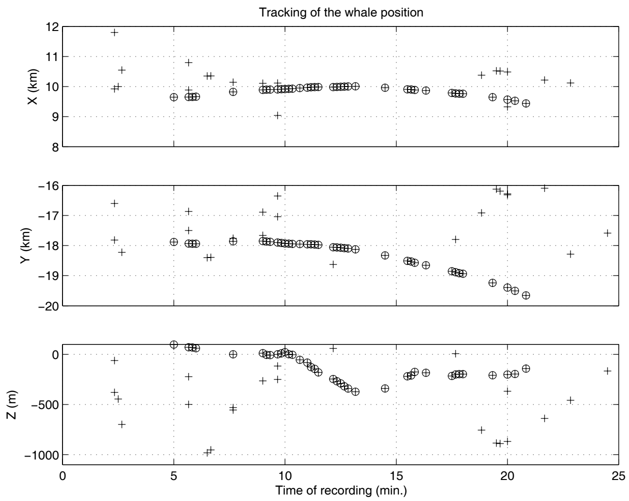

Figure 7: Results of the whale localization. Each cross indicates a position estimated with LMS error < 100. Circles show contiguous space positions similar to a sperm whale trajectory.

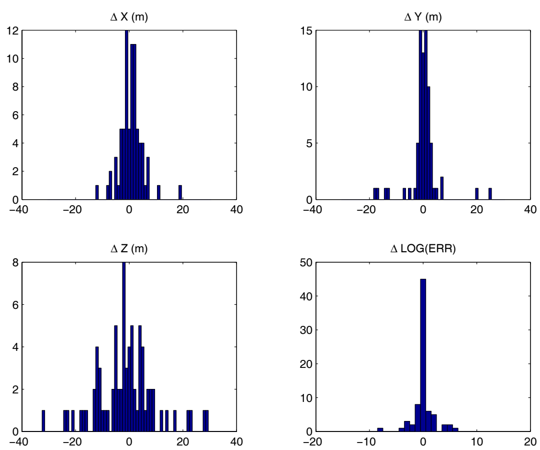

Figure 8: Histograms of the differences between position estimates on decimated filtered signal and original one. Delta X,Y,Z < 20m even if Delta log(LMS error) are large.

Referred steps by the section number of the paper	CPU time (laptop Dell D600)	Ratio to real Time
3.1 to 3.4 (without LMS)	5'25''	0.21
LMS position solving	6'12''	0.25
TOTAL of the full Algorithm	**11'37''**	**0.46**
Optional rerun of 3.2, 3.4 & 3.5 steps on original signal	46'83''	1.90
TOTAL with option	59'00''	2.36

Tab 2: Real time ratio of the various methods. The optional rerun is only for comparison to high sampling frequency data.

5. DISCUSSION AND CONCLUSION

Each step running times are given in Tab.2 using *Windows matlab* on a laptop computer *DELL Latitude* D600 *Intel Pentium* processor 1.7GHz, 1Go RAM**.** The total algorithm is two times faster than real time. Building a noise removal algorithm was necessary, like in other systems [7]; we defined a very simple and robust one which is echo-robust, and general for any transient signal, without any species dependency contrary to some other [7]. Even without taking account of uncertainty of underwater sound speed (fixed in our paper to 1500m/s), our results are competitive to other methods presented at the 2nd Internat. Workshop on MM localization using passive acoustics [15]: positions estimated by other participants, like SOEST-University of Hawaii-USA team, are within few km of ours [16]. In order to improve our system we are currently working on recursive LMS, forward clicks labeling after localization estimates, and further TDOA filtering [16].

6. REFERENCES

[1] W.J. Richardson, C.R. Greene, C.I. Malme and D.H. Thomson, *Marine Mammals and Noise,* Acad. Press, California, 1995.
[2] R.A. Altes, "Detection, estimation and classification with spectrograms", In *JASA* Vol. 67, no. 4, pp. 1232-1246, Apr. 1980.
[3] L. Cohen, *Time-Frequency Analysis*, Prentice Hall, 1993.
[4] C. Ioana and A. Quinquis, "On the use of time-frequency warping operators for analysis of marine-mammal signals", in *IEEE ICASSP*, pp.605-609, 2004.
[5] D. Moretti and al., "Open Ocean Marine Mammals Monitoring using widely spaced bottom mounted hydrophones", in *Journal of Underwater Acoustics,* 2002.
[6] D. Gillespie "Detection and Classification of right whale class using an edge detector operating on smoothed spectrogram", in *Journal of Canadian Acoustics*, Vol. 32, N2, pp.39-47, 2004.
[7] F. Desharnais and A. Hay, Overview ot the 2003 workshop on detection and localization of marine mammals using passive acoustics", *RR Department of Oceano. Dalhousie univ. CA*, 2003.
[8] S. Choi, H. Hong, H. Glotin and F. Berthommier, "Multichannel Signal Separation for cocktail party speech recognition: a dynamic recurrent network", in *Neurocomputing*, Vol 49, pp. 299-314, 2002.
[9] K. Torkkola. "Blind separation of convolved sources based on info. maximization", in *Proc IEEE NNSP, pp. 423-432*, 1996.
[10] T. Lee, A. Bell and R. Lambert, "Blind separation of delayed and convolved sources", in *Advances in Neural Information Processing Systems*, Vol 9 pp 758-764, 1997.
[11] D.W. Marquardt, "An algorithm for least squares estimation of non linear parameters", in *J. SIAM*, Vol 11, pp.431-441, 1963.
[12] E. Tessier, F. Berthommier, H. Glotin and S. Choi, "A CASA front-end using the localisation cue for segregation & then cocktail-party speech recognition", in *IEEE Proc. Int. Conf. on Speech Process. (ICSP)*, pp. 97-102, Seoul, 1999.
[13] H. Glotin, D. Vergyri, C. Neti, G. Potamianos and J. Luettin, "Weighting schemes for audio-visual fusion in speech recognition", in *IEEE ICASSP*, 4p, 2001.
[14] H. Glotin, "Dominant speaker detection based on harmonicity for adaptive weighting in audio-visual cocktail party ASR". in *Int. ISCA wkp Adapt. methods in speech recognition*, Nice, 4p., 2001.
[15] P. Giraudet and H. Glotin, "Real-time localization of sperm whales using click analysis", in *Proc. of 2nd Int. wkp on detection and localization of marine mammals using passive acoustics*, Oceanographic Museum, Monaco, 2005.
[16] P. Giraudet and H. Glotin, "Towards a Real Time Tracking Algorithm for Marine Mammals Survey using Passive Acoustics" LSIS.RR.2006.001, February 2006.