A 3-D real-time algorithm based on array processing for the localization of cetaceans

Ludwig Houégnigan, Serge Zaugg, Mike van der Schaar, Michel André

Laboratori d'Aplicacions Bioacústiques, Universitat Politècnica de Catalunya, 08800 Vilanova i la Geltrú, Espanya.
www.lab.upc.es

Abstract. *Hydrophone arrays permit to localize populations of cetaceans in real-time to study their behaviour in a non-invasive way and to work on their conservation. The most commonly used method based on time-difference-of-arrival can be challenged by space-time methods which can locate several simultaneous sources and separate them spatially, but the implementation of the latter can be complex especially when the aperture of the array is large compared to the smallest wavelength in the signal (aliased array). Several space-time methods such as the Capon and Music estimators were implemented, but due to the array design which includes a microphone of lesser sensibility they gave poor results compared to what could be expected from simulations. This paper documents some results obtained qua bearing estimation of vocalizing mammals using a three-dimensional array with 4 hydrophones and a steered-power-response beamformer.*

Keywords: *bioacoustics, ocean observatories, acoustic source localization, beamforming, cetaceans.*

INTRODUCTION

In the frame of the NEMO collaboration (Neutrino Mediterranean Observatory) for neutrino detection, more than 2000 hours of multichannel recordings were gathered. Indeed an underwater station was installed 25 km East of the port of Catania (Sicily) at 2000 m depth. The station is equipped with four hydrophones working in a frequency band which is sufficiently large (from 36 Hz to 43 kHz) not only for the detection of high energy neutrinos but also for the detection, classification and localization of vocalizing cetaceans.

Passive Localization has shown to be a powerful and non-invasive tool for biological and behavioural study and for the conservation of marine mammals. Many sound source localization procedures exist, the most commonly used being either:

- TDOA-based methods which use groups of time differences of arrival (TDOA's) to reconstruct the 2D (angle) or 3D (angle and range) of the radiating source.
- Space-time methods (also called spatial spectral estimation methods) which include beamforming and various high-resolution methods (MuSiC, Capon, Esprit)[]. The common feature of these methods is that they can be expressed in terms of processing of a space-time covariance matrix (STCM).

Passive space-time methods are known to have theoretical robust estimation capabilities but they have been scarcely used for the localization of marine mammals which contrasts with their successful use in the field of digital communication. A key aspect is that they can only be applied under certain measurement constraints: the microphones are supposed identical and hence, after calibration they should have similar amplitude and frequency responses. If some mismatches can be corrected for, others are difficult to compensate. For example in the case of NEMO two microphones are high-pass filtered directly by hardware and the last hydrophone used is from a different series and has a sensitivity 5 dB lower. From that last constraint some space-time methods such as steered-power response beamformer can still be used but it also implies that those relying on the STCM, which in general give more accurate results and have better source separation capabilities seem to be excluded.

THE LOCALIZATION PROBLEM

The source localization problem is usually well-defined in the case of broadband signals for TDOA-based methods and in the narrowband case for Space-time methods. For the latter the broadband case is often decomposed into a series of narrowband sub-problems. This is a rather complex and computationally costly approach which is moreover more subject to the spatial aliasing phenomenon: above the spatial aliasing limit each narrow frequency band will be strongly aliased whereas a truly broadband approach brings a spatial diversity which limits aliasing and permits to achieve a correct localization.

Since the species (mainly sperm whales) expected to be localized with the NEMO antenna emit broadband signals which put in relation with the aperture of the array would cause aliasing for a narrowband analysis, we use and present here a propagation model based on time signals which permits to apprehend broadband signals as such and not as an addition of narrowband signals. Hence we avoid or limit spatial aliasing and reduce the computational costs.

METHODS

1. Propagation Model

We assume a general three-dimensional array of M sensors receiving, due to propagation attenuated, phased and noised versions of the signal *s* emitted from a sound source at spherical position $\underline{\mathbf{r}}_s$ =[r_s Θ_s Φs]. The coordinates of $\underline{\mathbf{r}}_s$ respectively represent range, azimuth and elevation.

We can model $x_i(t)$ the signal received at the ith sensor at instant t as :

$$x_i(t) = \alpha_i(\underline{r}_s) \cdot s\left(t - del_{1,i}(\underline{r}_s)\right) + v_i(t)$$

Where v_i represents the additive noise in microphone **i** which may include background and propagation noise, reverberation, and electronic noise. If sensor *1* is taken as the reference sensor, the ith signal can be expressed by using the propagation delay $del_{1,i}(\underline{r}_s)$ which is related to the path difference between the signals impinging on *1* and *i*. Each x_i is thus modelled as a phased and attenuated by distance – illustrated by the term $\alpha_i(\underline{r}_s)$- and noise-corrupted version of *s*.

2. Importance of calibration and design

In the near-field, TDOA's are dependant on all components of $\underline{r}_s$ whereas in the far-field the range r_s has very little effect on them. The dimensions of large arrays make it possible to determine range because the ratio between the dimensions of the array and the actual range make it a quasi near-field situation; on the contrary, with a relatively small array like on the NEMO station (1m length tetrahedron), range discrimination becomes a difficult task. The advantage of Space-time methods is that they can take into account not only TDOA's but also the received amplitudes through basic modifications of the STCM: the curvature of the wave and hence the range can be better estimated. The requirement of well-calibrated and equally sensitive sensors would result in a better estimation of the amplitudes and TDOA's, and consequently in a better range estimation and localization.

3. Space-time methods

These methods aim at extracting information concerning the direction of the wavenumber $\underline{\mathbf{k}}$. This parameter can be understood as a spatial frequency and its direction corresponds to the direction of propagation of the signal. Hence estimating this spatial frequency is equivalent to estimating the direction of arrival (DOA) of the signal. Instead of extracting $del_{1,i}(\underline{r}_s)$ from x_1 and x_i this method uses all the x_i 's (which contains not only phase but also amplitude variation information) to directly estimate $\underline{r}_s$ or some of its components (DOA, range) without any additional inversion phase.

a) Steered- power response beamformer

This is one of the earliest spatial spectral estimators, yet it is rather robust. It's principle is to virtually steer the array in various direction and to look for a maximum of the average power $P(r_k, \theta_k, \phi_k)$ received at the array. Finally the estimated position (or DOA) $\hat{r}_s$ is given by

$$\hat{r}_s = \underset{k}{\mathbf{argmax}}\left(P(r_k, \theta_k, \phi_k)\right).$$

RESULTS

In this section we present part of the hundreds of results obtained for the steered-power response for different datasets from the NEMO station. We present two very clear localization results, the first one for sperm whale clicks and the second one for boats.

1) Localization of sperm whales

a) Click-by-click localization

In this dataset sperm whale clicks can clearly be heard and were detected automatically. For each of the 701 detected clicks in this 5 minutes sequence the localization procedure was repeated. Figure 1 and 2 present the spatial distribution of power received respectively for a high energy click and for a lower energy click. Higher energy most lightly-coloured regions in space correspond to the located sources and are very clear in figure 1. Bright "spider web-like" lines clearly visible in figure 1b are probably artefacts and aliases appearing due to the fact that the signals are not totally white and due to background noise. However the main source can still be located without doubt at DOA $\left(\hat{\theta}_s,\hat{\phi}_s\right)=\{281^\circ,40^\circ\}$. A one degree resolution was used for the computation of this spatial spectrum. In figure 2, the lower energy of the clicks causes the sharpness of the main localization peak to decrease in figure 1. Even though the main source is still localized without ambiguity at $\left(\hat{\theta}_s,\hat{\phi}_s\right)=\{281^\circ,40^\circ\}$.

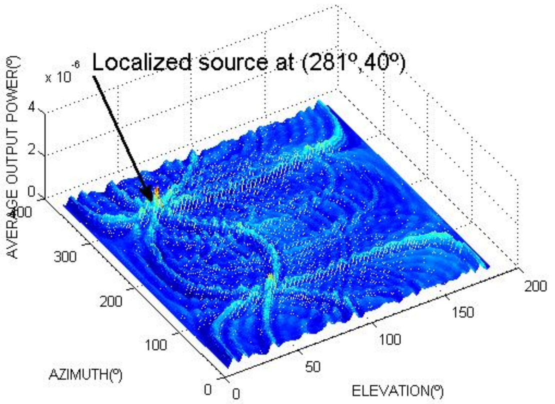

Fig. 1a: Average power received as a function of steered azimuth and elevation (3D view), for a high energy click. A main source is located at azimuth 281º and elevation 40º.

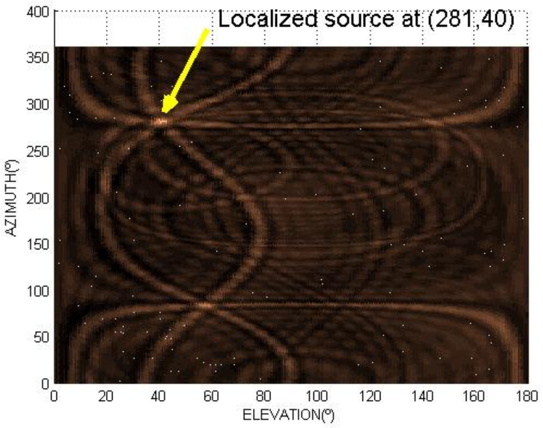

Fig. 1b: Average power received as a function of steered azimuth and elevation (2D view), for a high energy click

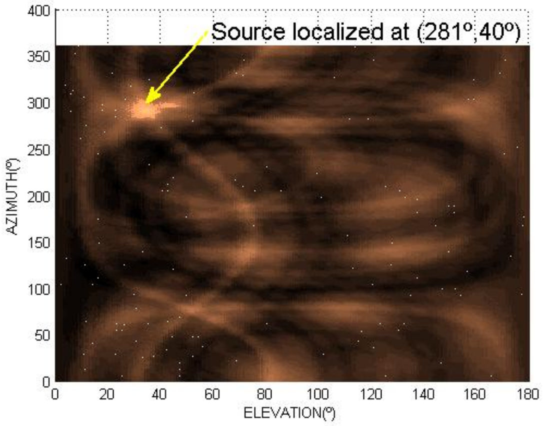

Fig. 2: Average power received as a function of steered azimuth and elevation (2D view), for a low energy click. The source can still be well localize

b) Tracking results for sperm whales

By repeating the localization procedure for each of the clicks occurring in a 5 minutes window we are able to track the movement of the emitting sources be they sperm whales or boats. Figure 3a shows the localization of some isolated clicks but more interestingly shows two main clusters of points. The first one indicates a sound source moving around $\left(\hat{\theta}_1,\hat{\phi}_1\right)=\{281^\circ,40^\circ\}$ and the second one a sound source moving around $\left(\hat{\theta}_2,\hat{\phi}_2\right)=\{81^\circ,49^\circ\}$. Hence we can number with confidence at least two vocalizing mammals in this sequence. The mammal corresponding to the $\left(\hat{\theta}_2,\hat{\phi}_2\right)$ cluster has a very clear pattern of decreasing elevation and azimuth in time. The $\left(\hat{\theta}_1,\hat{\phi}_1\right)$ cluster is less obvious, there could as well be two close animals. It is difficult to tell whether the algorithm is instable due to a poor signal-to-noise ratio or whether there are two animals. However, given the width of this cluster as seen in b) and c), the presence of at least two sperm whales is likely.

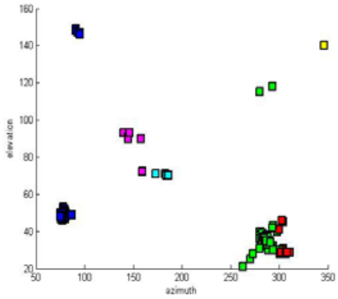

Fig. 3a: elevation vs. azimuth

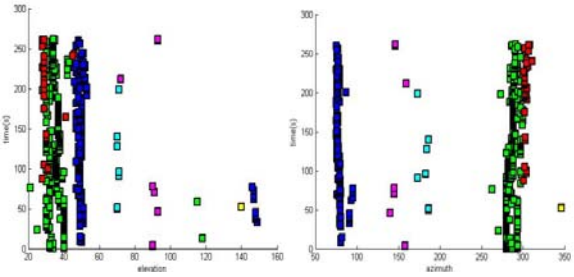

Fig. 3b: elevation vs. time (left) and azimuth vs. time (right)

2) Localization and tracking results for boats

Contrary to the tracking of sperm whale, the tracking of boats features a clear evolution of the DOA during the five minutes (figure 4b). This seems to confirm the fact that we are localizing boats since their speed is usually much faster than that of sperm whales. The first cluster (figure 4a) around $\left(\hat{\theta}_3,\hat{\phi}_3\right)=\{100^{o},65^{o}\}$ corresponds to a source which starts to radiate around 150s. It features a slow but clear increase of both azimuth and elevation. The second cluster around $\left(\hat{\theta}_4,\hat{\phi}_4\right)=\{275^{o},45^{o}\}$ corresponds to a source which radiates regularly during the 5 minutes of recording. It features a fast decrease of azimuth and a fast increase of elevation.

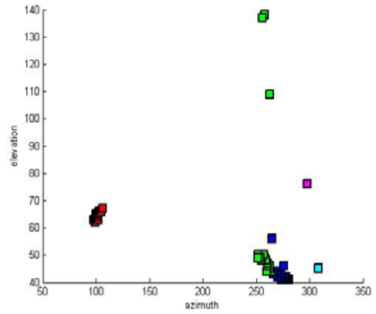

Fig. 4a: elevation vs. azimuth

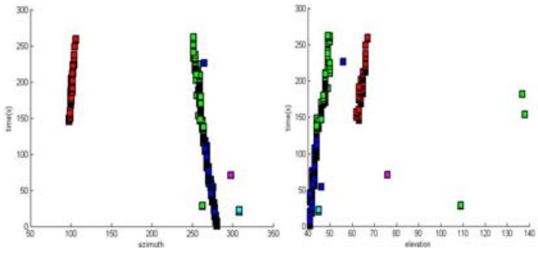

Fig. 4b: elevation vs. time (left) and azimuth vs. time (right)

CONCLUSIONS

The developed localization algorithm was tested both on simulated data and on real sequences of recordings from the LIDO project containing sperm whale clicks where the position of the sperm whale is unknown. Simulations showed that applying a broadband version of the MuSiC algorithm could provide consistent estimates of bearing and range which can then be used for the spatial separation of several sources and their tracking based on the evolution of the estimated location. However for the real data this method could not give good results due to engineering limitations (mainly sensitivity differences in the hydrophones). Hence we presented satisfying results obtained with a steered-power-response beamformer which localized quite clearly several radiating broadband sources, both boats and sperm whales. The engineering design is a limiting factor which is highly constrained by technical costs, however it should be thought that these costs could be balanced by the gain of more precise results in localization and enhanced capabilities for de-noising and sound enhancement which can be used to improve detection and classification.

REFERENCES

[1] D.H. Johnson. The application of spectral estimation methods to bearing estimation problems (invited paper). *Proc. IEEE*, **70**(9): 1018–1028, 1982.

[2] Krim, H. and M. Viberg, “Two Decades of Array Signal Processing Research,” *IEEE Signal Processing Magazine,* pp. 67-94, July 1996

[3] Knapp, C.H., Carter, G.C.: The generalized correlation method for estimation of
time delay. IEEE Trans. Acoust. Speech Signal Process. 24(4), 320–327 (1976)

[4] Schmidt, R. O., “Multiple Emitter Location and Signal Parameter Estimation,” *IEEE Transactions on Antennas and Propagation,* 34(3), pp. 276-280, March 1986

[5] J. Dmochowski, J. Benesty, and S. Affes, “On spatial aliasing in microphone arrays,” *IEEE Trans. Signal Processing*, vol. 57, pp. 1383-1395, Apr. 2009.

ACKNOWLEDGEMENTS

Recordings from Sicily provided by Gianni Pavan (CIBRA, ITALY). This study is funded by the European Commision under contract ESONET 036851, EC FP6.