Automatic and passive whale localization in shallow water using gunshots

Julien Bonnel, Grégoire Le Touzé, Barbara Nicolas, Jerome I. Mars, Cedric Gervaise

Automatic and passive whale localization in shallow water using gunshots

Julien Bonnel
GIPSA-Lab/DIS
Grenoble INP, France
Email: julien.bonnel@gipsa-lab.inpg.fr

Gregoire Le Touzé,
Barbara Nicolas,
and Jérôme I. Mars
GIPSA-Lab/DIS
Grenoble INP, France

Cedric Gervaise
E3I2
ENSIETA (Brest), France

***Abstract*— This paper presents an automatic and passive localization algorithm for low frequency impulsive sources in shallow water. This algorithm is based on the normal mode theory which characterizes propagation in this configuration. It uses specific signal processing tools and time-frequency representations to automatically extract features of the propagation. Then, it uses the dispersive properties of the oceanic waveguide as an advantage to perform the localization. Only few hydrophones are needed and neither knowledge of the oceanic environment nor simulation of the propagation is required. The proposed method is successfully applied on North Atlantic Whale gunshots in the Bay of Fundy recorded with a network of three hydrophones.**

I. Introduction

The study of marine mammals is a difficult task as most of the visual observations are closely restricted by weather, daytime and environment. However, passive acoustics which only consists in ”listening” the acoustical environment could provide another tool to obtain spatial and temporal distribution of marine mammals [15]. This could be helpful to understand and protect these animals, especially when endangered species are concerned. Indeed with passive acoustics, animals are located thanks to their own calls, in opposition to classical SONAR methods where a signal has to be emitted. This allows a more autonomous system, requiring less energy, and save the sea fauna being disturbed by active acoustic signals. Moreover, it can be used when visual observations failed.

When a marine mammal emits a call, the sound travels from the animal position to one or several hydrophones. The received signal depends on the oceanic environment, and on the positions of both source and receiver. Using signal processing techniques, it is possible to extract features from this signal and use them to estimate the source localization [8] [17]. Localization algorithms are usually based on an acoustic propagation model and require the knowledge of oceanic environment.

Right whales calls have been described in [16]. They are often low frequency calls, but have various waveforms: constant low-frequency, moan, up and down sweeping modulations, and gunshot. Gunshots are loud impulsive sounds from 10Hz to 20kHz lasting approximately 2ms. They are produced by lone males (or small groups), mainly in the bay of Fundy (Canada) and are probably implied in reproduction [13]. As they are emitted near the surface, they could be used for an automatic alert system to avoid whales and ship collisions. Moreover, the bay of Fundy is a shallow water area with internal tides producing large and quick variation of sound speed profile [4]. This implies that a robust localization algorithm has to be developped.

This paper presents a passive localization method (in a 2D horizontal plane) for low-frequency transient signals (such as gunshots) in shallow water environment, using a sparse network of hydrophones. The first part of the paper will introduce the experimental data used for this study. The main ideas of the modal propagation model on which relies our scheme are presented in a second part. A third part will describe the algorithm itself, including the necessary signal processing and time-frequency tools. Finally, the method will be applied on real data and discussion will be done.

II. The data set

The dataset used in this paper comes from the 2003 Workshop on detection, localization and classification of marine mammals using passive acoustics. The acquisition system is composed by five Ocean Bottom moored Hydrophones. Their localization is given in Table I. They have a flat sensibility from 50Hz to 700Hz and the data were digitized using a 12-bits A/D converter with a sampling frequency of 1200Hz.

TABLE I
Dataset OBH Position

OBH	Latitude (N)	Longitude (W)	Water depth (m)
C	44.60073	66.49723	210
E	44.60237	66.31591	134
L	44.66203	66.40453	183
H	44.73051	66.31556	123
J	44.73038	66.49619	170

The area around the OBH is shallow water, with bathymetry from 100 to 200 meters. Sound speed profiles were measured during this experiment: they were downward refractive or had a local minimum. As said before, they also presented quick temporal variations. The bottom structure is mainly composed of a first Lahave clay layer over a thick layer of Scotian drift [11]. The Lahave clay layer is characteristic as its compression sound speed is lower than the sound speed in water. It implies a high dispersion for normal mode propagation. The dataset contains several right whale sounds recorded in the bay of Fundy between 2000 and 2002 [4], including gunshots. The Fig. 1 presents an example of a recorded gunshot (S035-2 on hydrophone H) in the time and the time-frequency domails. The latter presents a multicomponent pattern, which is typical of a dispersive normal mode propagation: each component has its own time of arrival which depends on frequency. We give some details on this propagation in the following section.

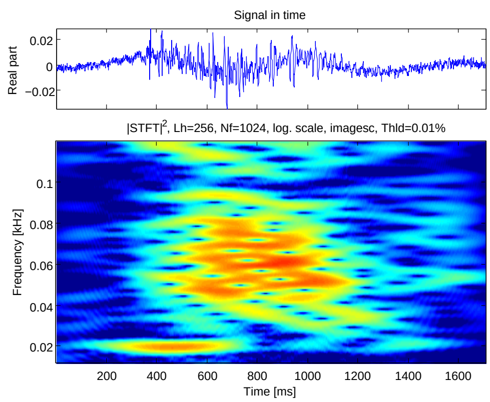

Fig. 1. Time representation and time-frequency representation of the recorded gunshot on OBH #H

III. Normal mode theory

In our configuration (shallow water and low frequency signals), the most suitable propagation model is normal mode theory. In this case, in a range independent environment, for a frequency f, the transfer function H between a source at depth z_s and a receiver at a depth z_r separeted by a radial distance r is [7] :

$$H(f) \approx Q \sum_{m=1}^{\infty} g_m(z_s) g_m(z_r) \frac{e^{jk_r(m,f)r}}{\sqrt{k_r(m,f)r}} \tag{1}$$

where g_m is the m^{th} modal function, $k_r(m,f)$ the radial wavenumber of mode m (which is supposed to be real as the evanescent modes are not taken in account), and $Q = \frac{e^{j\pi/4}}{\rho(z_s)}$ (with $\rho(z_s)$ the water density at the source depth). Thus, the propagation is multicomponent. For each component of index m, phase speed v_ϕ and group speed v_g can be defined by :

$$v_\phi(m,f) = \frac{2\pi f}{k_r} \tag{2}$$

$$v_g(m,f) = 2\pi \frac{\partial f}{\partial k_r} \tag{3}$$

The group velocity describes the propagation speed of energy. We can note that v_ϕ and v_g depend both on frequency f and mode index m. Consequently, each frequency of each mode will travel with its own speed, which is the definition of a dispersive propagation.

If a transient signal is emitted with a time frequency modulation $t_e(f)$ (t_e is the time of emission of the frequency f), the time-frequency structure TFR of the receive signal after modal propagation is:

$$TFR(t,f) = \sum_{m=1}^{\infty} A(m,f,r,z_s,z_r)\delta(t - t_e(f) - \frac{r}{v_g(m,f)}) \tag{4}$$

where $\delta(t)$ is the dirac distribution describing the localization of the time-frequencies structures and A is the attenuation term describing their amplitude.

IV. The localization algorithm

The main idea is to take advantage of the dispersive behaviour of the oceanic waveguide to localize a transient emission. For a source s emitting a transient signal with an unknown time-frequency modulation $t_e(f)$, the arrival time of the frequency f of the mode m measured after propagation at a hydrophone n is given by:

$$t_r(m,n,f) = t_e(f) + \frac{r(s,n)}{v_g(m,f)} \tag{5}$$

A. Estimation of the arrival times

The first step of the algorithm is to estimate all the t_r. It is impossible in the time domain as the modes are overlapped. It is neither direct in the time-frequency domain because the modes are broadband and close from each others. Consequently, each mode has to be first filtered.

1) Warping operators: As proposed in [10], the pressure signal will be warped in order to have a better representation of the modal information. The warping is based on a model of the environment and is computed with an unitary equivalence approach [1]. Here, the used model is the isovelocity one. It is made of a homogeneous layer of fluid between perfectly reflecting boundaries. Of course, this modelisation is simplistic, and does not match to the real oceanic environment. However, it is useful as it does not require information of the environment and is enough efficient for our goal.

In the isovelocity case, the pressure signal is given by [7]:

$$p(t) = \sum_m g_m(t) e^{j2\pi\nu_c(m)\xi(t)} \tag{6}$$

with $g_m(t)$ describing the envelop of the m^{th} mode, $\nu_c(m)$ the cutoff frequency of the m^{th} mode (depending only on m, on the constant velocity V of the water and of the depth D

of the waveguide) and $\xi(t)$ the general dispersivity function which is:

$$\xi(t) = \sqrt{t^2 - \frac{r^2}{V^2}} \tag{7}$$

The warping operator is based on a deformation function $\omega(t)$ and is noted W_ω. Its aim is to linearize the pressure signal. Consequently, $W_\omega p(t)$ must be a sum of linear structures and follows the equations :

$$W_\omega p(t) = \sum_m \sqrt{|\omega'(t)|} C_m e^{j2\pi\nu_c(m)\xi[\omega(t)]} \tag{8}$$

$$W_\omega p(t) = \sum_m \sqrt{|\omega'(t)|} C_m e^{j2\pi\nu_c(m)t} \tag{9}$$

The corresponding deformation function ω is [10]:

$$\omega(t) = \xi^{-1}(t) = \sqrt{t^2 + \frac{r^2}{V^2}} \tag{10}$$

If this operator is applied on a pressure signal (from an isovelocity waveguide), the modes of the warped signal will be sinusoids. There is also an inverse operator W_ω^{-1}. It is defined by $W_\omega^{-1} W_\omega p(t) = p(t)$ and is linked to the deformation function $w^{-1}(t)$.

2) Modal filtering: This operator could be applied on a real pressure signal. As the real waveguide is not an isovelocity one and as r is unknown, the warped modes are not perfect sinusoids. However, they are quite well separated in the time-frequency domain and can be easily filtered. Here, a simple threshold is sufficient to create masks on time-frequency domain and filtered them. Consequently, the filtering scheme is as follow :

1) Using the isovelocity model, the recorded signal is warped: modes become nearly sinusoids (in the time-frequency domain, they are nearly horizontal lines).
2) A time-frequency representation of the warped signal is computed: Short Time Fourier Transform (STFT) is chosen as it is easily computable and allows filtering in the time-frequency domain using masks.
3) Masks are created to filter warped modes (in the time-frequency domain).
4) Each filtered (warped) mode expressed in the time domain with an inverse STFT.
5) Each warped mode are unwarped (in the time domain).

3) Final estimation of arrival times: As each mode is filtered, it is easy to obtain the times of arrival $t_r(m, n, f)$ (given by equation 5). For a given hydrophone n, the reallocated spectrogram of the mode m is computed [5]. Indeed, a reallocated spectrogram reassigned position of time-frequency points in order to obtain thin structures. But even after reallocation, the time-frequency representation of the mode is not perfectly thin. Consequently, the instantaneous frequency of the mode is obtained by computing the skeleton of the shape of the mode. This is done for each mode on each hydrophone and gives all the t_r (to have a better understanding of the whole procedure, figures will illustrate it on section V-A). Once the times of arrival are known, they could be used to localize the source.

B. A ratio only depending on source-receiver distance

By combining different arrival times, it is now possible to define a ratio which depends only on source-receivers distance [6]. First, let us considere the arrival times of a frequency f of two modes m and m' on a single hydrophone n. By substrating them, the influence of the time of emmission $t_e(f)$ disappears:

$$d(m, m', n, f) = t_r(f, m, n) - t_r(f, m', n) \tag{11}$$

$$d(m, m', n, f) = r_{sn}\left(\frac{1}{v_g(m, f)} - \frac{1}{v_g(m', f)}\right) \tag{12}$$

As the group velocity is the same for all hydrophones, it is possible to combine the quantities d for two different hydrophones n and n' to get rid of them :

$$R(m, m', n, n', f) = \frac{d(m, m', n, f)}{d(m, m', n', f)} = \frac{r_{sn}}{r_{sn'}} \tag{13}$$

The quantity R is computable only with the received measures. It does not depend on the waveguide properties nor on the emitted signal. Consequently, it is a suitable quantity to robustly estimate the localization of the source. Thus, in a horizontal plane, it restrains the position of the source to lie on a circle (or a line if $R = 1$).

C. A cost function to localize the source

Instead of considering geometrical properties, the localization of the source is estimated by minimising a cost function J, as in [6]. With the extracted t_r, all the possible 5-uplets (m, m', n, n', f) are formed (with $m \neq m'$ and $n \neq n'$). Then, the source localization $\hat{S}(x_s, y_s)$ is estimated by solving the following estimation problem, which reprensents the sum of the quadratic distance to the constraint:

$$\hat{S}(x_s, y_s) = \arg \min_{M(x,y)} J[M(x, y)] \tag{14}$$

with

$$J[M(x, y)] = \sum_{i=1}^{N} [d^2_{M(x,y),n} - R^2(m_i, m'_i, n_i, n'_i, f_i)\, d^2_{M(x,y),n'}]^2 \tag{15}$$

where $M(x, y)$ is the point where the cost function J is computed, $d_{M(x,y),n}$ is the distance between the hydrophone n and the point $M(x, y)$, and N is the number of 5-uplets (m, m', n, n', f) previously selected.

V. Application on real data

A. Presentation of the results

The localization scheme described on the previous section was applied to the recording 'S035-2' from the data set of the 2003 Workshop on detection, localization and classification of marine mammals using passive acoustics. It contains a gunshot of right whale recorded in the Bay of Fundy. For this work, only hydrophones H, E, and L were used. Thus, the gunshot was not properly recorded on hydrophone J, and the environment changes a lot around hydrophone C. As an example, the signal recorded on hydrophone H is given in Fig. 1. Its warped version is presented in Fig. 2. For all the hydrophones, the warping parameters have been arbitrary chosen with the values $r = 8500\ m$ and $V = 1500\ m/s$. Even if it does not correspond to reality, the modes are nearly horizontal and well separated. On Fig. 2 it is noticeable that time and frequency scales have changed since it is in the warped domain. Modes are easily filtered. The mask created to filter the second one could be seen in Fig. 3. Once it is done, each mode is unwarped. A time-frequency representation (reallocated spectrogram) of the second mode is given on Fig. 4. Then, their instantaneous frequency is computed. It allows to obtain all the $t_r(f)$ for this mode. On Fig. 5, the red lines (which are the instantaneous frequencies of each modes) are constitued by all the extracted $t_r(f)$.

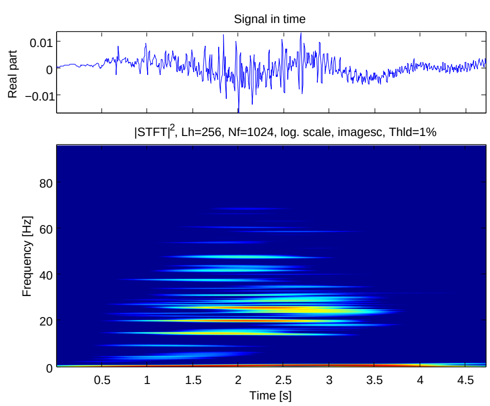

Fig. 2. Time representation and time-frequency representation of the warped version of the gunshot recorded on OBH #H

Then, all the possible 5-uplets $R(m, m', n, n', f)$ are computed. As told before, only hydrophones H, E and L were considered. Different tests were done using only the first two modes, the first three modes, or the first four modes. For each 4-uplet (m, m', n, n'), 10 different frequencies were chosen linearly spaced in the biggest mutual frequency band of m and m'. Criterium J was computed with a step of 5 meters. Table II summarizes the localization's results depending on the number of modes used, while table III allows to compare the proposed method with other techniques on the same data [3] [6] [9]. Fig. 6 presents the J criterium and the corresponding estimated position when only the first two modes were considered. Fig. 8 is a zoom of Fig. 6 allowing comparison of the result with other methods. Both table III, Fig. 6 and Fig. 8 are presented

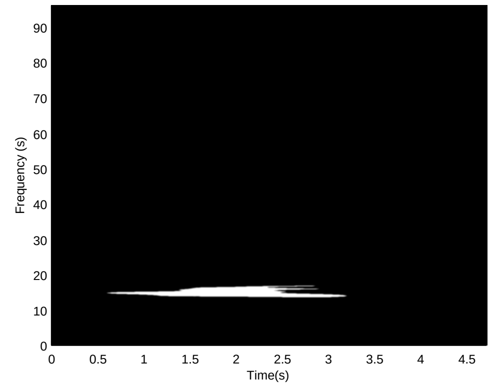

Fig. 3. Mask created to extract the second mode of the warped version of the gunshot recorded on OBH #H

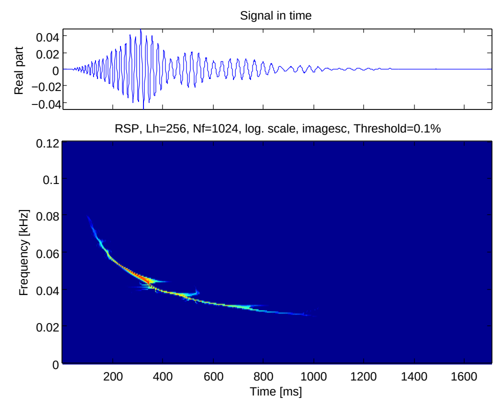

Fig. 4. Time representation and reallocated spectrogram of the second mode (after filtering and unwarping) of the gunshot recorded on the OBH #H

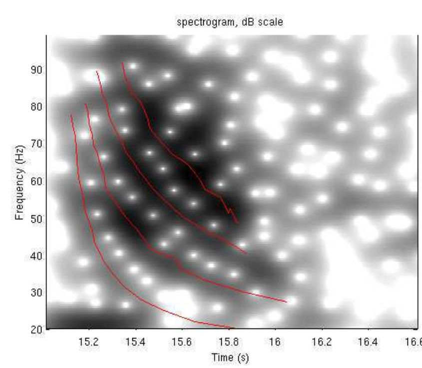

Fig. 5. The tracked modes of the gunshot recorded on the OBH #H

when only the first two modes are considered so the results can be compared with the one obtained by Gervaise et al. in [6] where only this two modes were used. We will discuss this point in section V-B. In the other cases, our method give the same kind of results (localization and shape of J).

B. Discussion

Our gunshot localization is compatible with the solutions obtained in the litterature, and similar results were obtained on S070-3 and 2013-1 recordings. Even if the method was

TABLE II

S035-2, GUNSHOT LOCALIZATION'S RESULTS

Number of modes used	x_{GS} (m)	y_{GS} (m)
2	7886	-1008
3	7483	-535
4	7255	-557

designed for the localization of right whale in the Bay of Fundy, it could be used whenever one wants to localize low frequency transient sounds in shallow water, for example :

- North Pacific right whales and Humpback whales in the Bering sear [12]
- eastern North Pacific blue whales [14]
- blue whale in the Saint Lawrence [2]

1) Analysis of the results: By looking at the shape of the criterium J in Fig. 6, it is noticeable that the precision of the estimated position of the gunshot is lower for direction x than for y. It is normal as the localisation of the gunshot is nearly outside the hydrophone network for the direction x. The results presented in table II show that x_{GS} decreases as the number of modes used in the algorithm increases. It may be explained because the first modes are more energetic and easier to track, so the precision is higher. This is confirmed by Fig. 7, where the shape of J is sharper for only two modes.

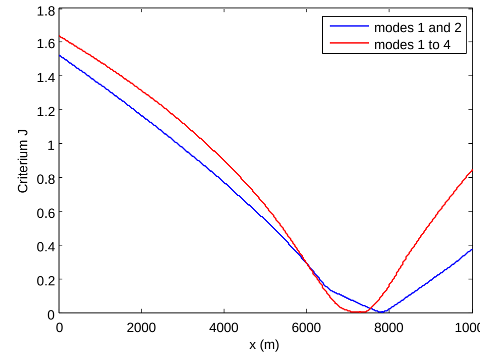

Fig. 7. Criterium J for $y = y_S$ considering the two or the four first modes

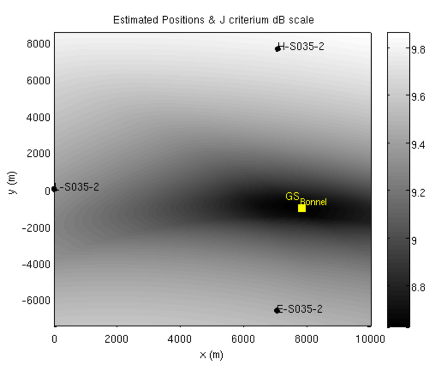

Fig. 6. Global view of the networks of OBH and results of the localization scheme using only two modes

2) Comparisons with other techniques: When the method is compared with the litterature, four points should be noted:

- Our method is an improvment of the one presented by Gervaise and al. in [6] as it is fully automatic. Moreover, the frequency band of interest is now higher as it depends on the couple mode/receiver instead of being fixed. This allows to consider more modes.
- Desharnais et al. [3] and Laurinolli et al. [9] use the direct ray path to model the propagation. Our method is based on the true propagation model, so it should be more precise. However, they use the whole frequency band of the gunshot (from 20Hz to 20kHz) whereas we use only the lower part of it. Consequently, it is difficult to know which method is the more accurate. We assume that our method is better when the environment is unknown or varying, but Desharnais' and Laurinolli's methods are probably more accurate when the environment is constant and well known. Fig. 8 allows a graphical comparison of our results with the three other methods discussed in this subsection.
- Our method is based on modal propagation, but does not require to run a normal mode propagation code [17]. It is a major advantage in such an environment which cannot be modelised with a simple Pekeris waveguide as there is a poorly compact first layer in the bottom and a time variable celerity profile.

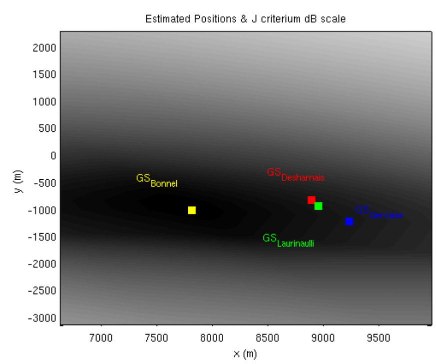

Fig. 8. Estimated position and J criterium using only the first two modes, and estimated positions with the methods presented in [3], [6] and [9]

TABLE III

S035-2, GUNSHOT LOCALIZATION WITH THE FIRST TWO MODES AND COMPARISON WITH OTHER METHODS

Method	x_{GS} (m)	y_{GS} (m)
Bonnel et al.	7886	-1008
Gervaise et al. [6]	9225	-1248
Desharnais et al. [3]	8884	-848
Laurinolli et al. [9]	8950	-970

VI. Conclusion

This paper presents a passive localization scheme for low frequency transient sounds in shallow water and its application to right whale gunshot in the Bay of Fundy, Canada. This method is fully automatic. Moreover, it does not require information of the environment properties, nor need to run a propagation code (whereas it is based on normal mode theory). As it uses relative time of arrival, it is not sensitive to clock's drift. All this properties make it suitable to obtain a real-time localization system, provided that a real time implementation of the time-frequency tool is possible. As perspective, the author woud like to apply the proposed method on a larger dataset to test the proposed approach, and also on simulations to determine its precision and its variability to the different parameters.

References

[1] R. G. Baraniuk and D. L. Jones,
Unitary equivalence: a new twist on signal processing,
IEEE Trans. Signal Processing, 43(16): 2269-2282, 1995

[2] C. L. Berchock, D. L. Bradley, T. B. Gabrielson,
St Laurence blue whale vocalization revisited: characterization of calls detected from 1998 to 2001,
J. acoust. Soc. Am. 120(4), October 2006 p 2340-2354

[3] F. Desharnais, M. Côté, C. J. Calnan, G. R. Ebberson, D.J. Thomson, N.E.B Collison, C. A. Gillard Right whale localisation using a downhill simplex inversion scheme,
Canadian Acoustics, vol32, no 2, 2004, p 137-145

[4] F. Desharnais, M.H. Laurinelli, D.J. Schillinger, A.E. Hay, A description of the workshop dataset,
Canadian Acoustics, vol.32 No 2 (2003), p 33-38

[5] F. Auger, P. Flandrin,
Improving the Readability of Time-Frequency and Time-Scale Representations by the Reassignment Method,
IEEE Transactions on Signal Processing, Vol. 43, No. 5, pp. 1068-89, 1995

[6] C. Gervaise, S. Vallez, Y. Stepahn, and Y. Simard ,
Robust 2D Localization of low-frequency calls in shallow waters using modal propagation modelling,
Canadian Acoustics, Vol. 36, no 1, p. 153-159, Mars 2008

[7] F.B. Jensen, W.A. Kuperman, M.B. Porter, H. Schmidt Computational ocean acoustics,
AIP Series in Modern acoustcs and signal processing, AIP Press, Springer, 1993

[8] M. H. Laurinolli, F. Desharnais, C. T. Taggart,
Localization of North Atlantic Right whale sounds in the bay of Fundy using a sunobuoy array,
Marine Mammal Science, 19(4):708-723, October 2003

[9] M. H. Laurinolli, A. Hay Localization of right whale sounds in the workshop bay of fundy dataset by spectrogram cross-correlation and hyperbolic fixing,
Canadian Acoustics, vol32, no 2, 2004, p 132-136

[10] G. Le Touzé, J. I. Mars, and J-L. Lacoume,
Matched time-frequency representations and warping operator for modal filtering,
EURASIP Signal and Image Processing Conference (EUSIPCO 06), Florence, sept. 2006

[11] Surficial geology – Eastern Gulf of Maine and Bay of Fundy, Map 4011-G *Canadian HydrographicService, 1977*

[12] M.A. Mc Donald, S. E. Moore,
Calls recorded from North Pacific right whales (Eubalaena Japonica) in the eastern Bering sea,
J. Cetacean Res. Manage. 4(3):261-266, 2002

[13] S. E. Parks, P. K. Hamilton, S. D. Kraus, P. L. Tyack The gunshot sounds produced by male north atlantic right whales (Eubalaena glacialis) and its potential function in reproductive advertisement,
Marine Mammal Science, 21(3):458-475 (july 2005)

[14] E. M. Oleson, J. Calambokidis, W. C. Burgess, M. A. McDonald, C. A. LeDuc, J. A. Hildebrand,
Behavioral context of call production by eastern North Pacific blue whales,
Marine Ecology progress series, vol.330:269-284, 2007

[15] J. L. Spiesberger, K. M. Fristrup,
Passive localization of calling animals and sensing of their acoustic environment using acoustic tomography,
The american naturalist, vol. 135, No. 1, January 1990, p 107-153

[16] A. S. M. Vanderlaan, A. E. Hay, C. T. Taggart Characterization of North Atlantic right whale (Eubalaena glacialis) sounds in the Bay of Fundy,
IEEE JOE, VOL. 28, No. 2, April 2003, p 164 – 173

[17] S. M. Wiggins, M. A. McDonald, L. M. Munger, S. E. Moore, J. A. Hildebrand,
Waveguide propagation allows range estimates for north pacific right whales in the Bering sea,
Canadian acoustics, vol. 32 No.2 (2004), p 146-154