Contributions to passive acoustic oceanic tomography –
Inversion Algorithms based on Time Frequency Space Representation

C. Gervaise*, S. Vallez*, E. Boumansour*, H. Le Floch*, A. Martin*, A. Khenchaf*, Y. Simard**

* E3I2, ENSIETA, EA3876, 2 Rue François Verny, 29200, Brest, France
** Institut des Sciences de la Mer, Université de Québec à Rimouski, 310 Allée des Ursulines, Rimouski, Québec G5L 3A1, Canada,

cedric.gervaise@ensieta.fr

ABSTRACT

This paper addresses the design of signal pre-processors dedicated to passive acoustic tomography. Blind processors of channel's impulse response and spatial response are proposed and adapted with wide band transient signals such as marine mammals vocalizes. Full validation of blind channel's impulse response processor is performed on real world data whereas validation of blind spatial response estimation is carried out on realistic synthetic data.

1. INTRODUCTION

Acoustic tomography is a way to produce a fast, accurate and cheap monitoring of water mass. This monitoring requires an inversion procedure made with two steps. The first one is to estimate acoustic properties (such as the sound speed profile of the water column) from the measurement of a propagated known acoustic waveform between fixed sources and receivers. Then a second step consists in inferring some physical ocean parameters (temperature, bottom nature) from these previous estimated acoustic characteristics. Large scales deep water and small scales shallow water configurations were successfully studied and associated with matched delay, matched field and matched impulse response inversion processing.

Accurate estimate of acoustic properties requires the emission of powerful and recurrent signals in the adapted bandwidth and in agreement with the scale of the monitoring. But we would rather not send these hard active sounds through the water column in a potential military underwater warfare context, or if mammal species health is considered. A recent solution has emerged in the community to tackle this problem with the passive tomography processing. Passive tomography processing consists in estimating acoustic properties by using opportunity sources present in the channel at the time of interest. Some experimentations were recently carried out using ships, marine mammals and surface noises.

Different levels of complexity can be formulated to insure the discreetness of tomography processing. The first one is the Active Discreet Tomography (ADT) where active emission is allowed but with a waveform chosen to insure a low probability of interception using for instance a copy of a noise component or a spread spectrum signal. In that case, Signal to Noise Ratio (SNR) is reduced compared to classic active tomography. The second one is the Aided Passive Tomography (AiPT) where active emission is forbidden but where a cooperate entity of known position can produce an acoustic emission linked to its natural activity. Blind estimation of the impulse response of the channel is performed with the losses of absolute time and magnitude references, and a reduced SNR. The last one is the Autonomous Passive Tomography (AuPT) where active emission is forbidden but where an entity of unknown position can produce an acoustic emission linked to its natural activity. As in the case of the AiPT, blind estimation of the impulse response of the channel is performed with the losses of absolute time and magnitude references, SNR is reduced, and moreover, position of the source considered as a nuisance parameter has to be estimated jointly with the parameters of interest.

For AiPT, preliminary works conducted on performance prediction based on lower Cramer Rao bound calculus have demonstrated that as soon as celerity profile or bottom parameters estimation are concerned, the blind estimation of impulse response of the channel between emitter and receiver in relative time (referenced to the first arrival) carries enough information to inverse the problem. Thus, the first part of the paper addresses the problem of Blind Impulse Response estimation using transient opportunity source with clear time frequency contents such as marine mammals vocalizes. Low-resolution and high-resolution are developed and applied with success to real data obtained from Laurentian channel experiment performed in summer 2003.

For AuPT, the previous preliminary works have demonstrated that as soon as celerity profile or bottom parameters estimation are concerned at the same time that source position estimation, the measurement of direction of arrival associated with blind impulse response channel estimation carried enough information to inverse the problem. Then a second part of the paper deals with the development of a Spatial Time Frequency processor to estimate the temporal and spatial structure of the arrivals. This processor is applied with success to synthetic but realistic data in shallow water environment.

2. TIME FREQUENCY PROCESSOR FOR AiTP

When the central frequency of the opportunity source is high enough, acoustic ray paths propagation takes place and signal at the receiver can be seen as a sum of attenuated and delayed versions of emission. Then, if the emission has a clear time frequency content

such as marine mammal vocalizes, time frequency processing can be used advantageously. A theoretical Time Frequency mapping of the received signal m(t) concentrates the tempo-spectral power density around N versions of the instantaneous frequency curve of the source s(t) translated in time. The proposed processor is based on the main characteristics of this Time Frequency mapping. A first stage is dedicated to the instantaneous frequency of s(t) estimation, then a second stage performs a time frequency matched filtering of m(t) with s(t).

To estimate the source's instantaneous frequency function, a local maximum is sought on an optimal time frequency mapping which deletes the interference terms without significantly increasing the spread of the auto-terms (RID: Reduced Interferences Distributions). These signal-dependent time frequency representations are based on the optimal weighting of the ambiguity function by a radially signal-dependent Gaussian kernel (Radial Gaussian Kernel, RGK) or based on the optimal weighting of local ambiguity function (Adaptative Optimal Kernel, AOK) developed by Baraniuk and Jones

[Bar93] . Our approach may be biased but is stable over noise and interferences. The algorithm used to estimate the source's instantaneous frequency function is the following:
- compute the signal-adapted Time Frequency mapping of the received signal, ($RGK_m(t,f)$),
- for each frequency bin f_i, estimate the time of the first local maximum of the function of time $RGK_m(t,f_i)$.

The source's instantaneous frequency function $\widetilde{f}_i(t)$ obtained at this stage is used to estimate the channel impulse response.

For known source s(t), two optimal detectors of s(t) in noise may be used to estimate the channel impulse response: the classical matched-filtering and an equivalent formulation in time frequency domain proposed by Flandrin [Fla88] . If the signal to be detected is considered as a random one and if m(t)=s(t)+b(t) or m(t)=b(t), the optimal detector consists in performing the time frequency correlation Q between the auto Wigner-Ville of s(t) and m(t) :

$$Q = \int\limits_{-\infty}^{\infty}\int\limits_{T} WV_{mm}(t,f) \times WV_{ss}(t,f)\,dt\,df$$

In passive tomography, the source s(t) is unknown, but the instantaneous frequency function $\widetilde{f}_i(t)$ of s(t) in given by the first stage, thus we can define an estimated Wigner-Ville of the source as follows:

$$\overline{WV}_{ss}(t,f) = \delta(f - \widetilde{f}_i(t)) \text{ where } \delta \text{ stands for Impulse}$$

A sub-optimal detector is proposed by computing the time frequency correlation between the auto Wigner-Ville of m(t) and the estimated Wigner-Ville of the source ($\overline{WV}_{ss}(t,f)$). The channel impulse response is estimated by looking for local maxima on $E(t_0)$ computed by:

$$E(t_0) = \int\limits_{-\infty}^{\infty}\int\limits_{T} WV_{mm}(t,f) \times \overline{WV}_{ss}(t-t_0,f)\,dt\,df \ .$$

The temporal resolution of this first algorithm is close to the classical active matched filtering one. As soon as celerity profile estimation is concerned, first acoustic paths carry much of information about it, but are usually not resolved by the matched filter and require high-resolution processing. We are now focused

on the development of a blind high-resolution time frequency processor dedicated to AiPT.

This tool is dedicated to signals s(t) having a curvilinear distribution of time spectral power density. Each signal of this family can be locally approximated by a Chirp signal and then if the area of validity of this assumption and the chirp parameters (central frequency and bandwidth) are known, MUSIC algorithm can be applied to m(t) in order to estimate each delay τ_i. The critical point of the algorithm is to determine automatically for each time t_0, the optimal neighborhood where the chirp-like assumption is valid. This is achieved by looking for the length L of a rectangular time window ($w_L(t-t_0)$) to apply to m(t) which minimizes the spread of the Fractional Fourier Transform of the windowed signal. The algorithm developed follows the flow chart presents in [Ger01] .
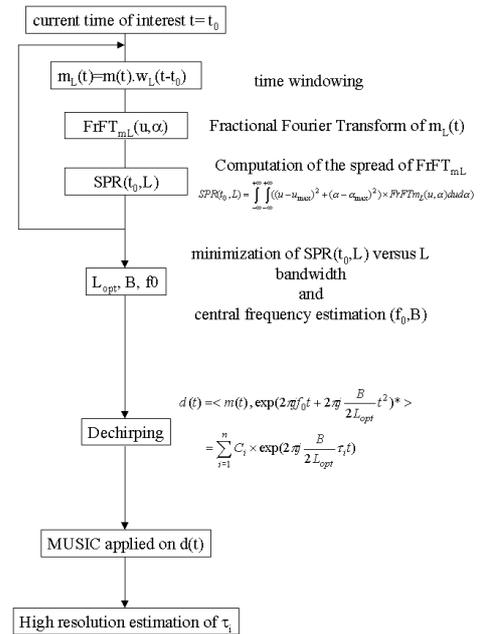

Figure 1 : Time Windowing – Dechirp – MUSIC Flow Chart

3. REAL WORLD APPLICATIONS

To observe the efficiency of the algorithms described in the second paragraph, we applied them on real data.

3.1 Material

To evaluate our algorithms, we used data provided by the 'Institut des Sciences de la Mer' (ISMER) from the University of Quebec at Rimouski (UQAR). Saint-Lawrence channel presents two critical habitat areas of marine mammals and a very dense ship traffic (see fig 2). To understand interaction between marine mammals behavior and ship radiated noise, passive recording of a few marine mammals sounds were performed by a network of hydrophones (5 ocean bottom hydrophones and a 6 coastal array hydrophones) in order to identify and to locate them in summer 2003. At the same time, to understand acoustic propagation, bathymetry and sound speed profile measurements were done (see fig 2). Details on this experimentation can be found in [Sim04] .
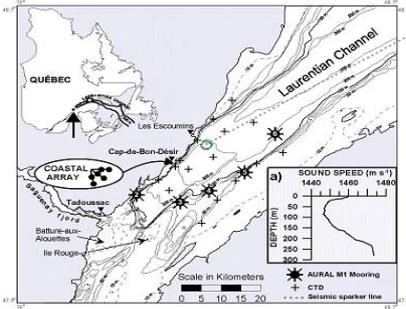
Figure 2. Study area with the location of both the coastal array
and the 5 buoys
a) A typical celerity profile measured in this area.

3.2 Data Analysis

Among the large set of collected data, impulse sound with bandwidth from 0 to 1 kHz, frequency modulation from 500 Hz to 8 kHz, and narrow band sound produce the three major families of marine mammals sounds where multi-paths structure of the measurement can be observed using the first two ones. A Beluga vocalize with decreasing frequency modulation of 2 kHz bandwidth (see fig3) is selected as a test bench for our algorithms.
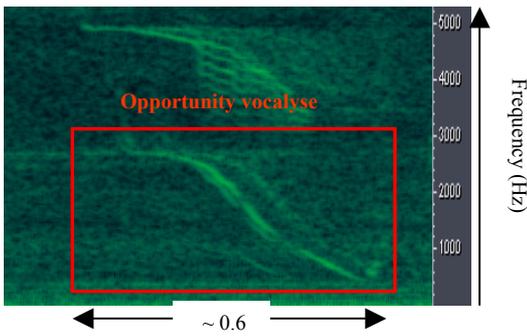

Figure 3. Spectrogram of the opportunity vocalize
(Window used: Hamming 25ms)

The time frequency in figure 3 is the spectrogram of the test sound recorded at the first hydrophone of the coastal array.
Resolved multi-paths can be observed on the spectrogram. It seemed to confirm our hypothesis of acoustic ray paths propagation, then it is proposed to estimate the impulse response of the channel between beluga and the receiver with low and high resolution algorithms presented in section 2 and to compare them with a simulated one given by ray path Bellhop code using the position of Beluga obtained by triangulation between all coastal array receivers.

3.3 Results

Low-resolution algorithm provides estimates of instantaneous frequency law:

$$e(t) = \exp\left(2\pi j \int_0^t f_i(u)\,du\right), t \in [0, 0.6s]$$

with $f_i(t) = 10^4(-1.78t^3 + 1.43t^2 - 0.4t + 0.18)$ (Hz)
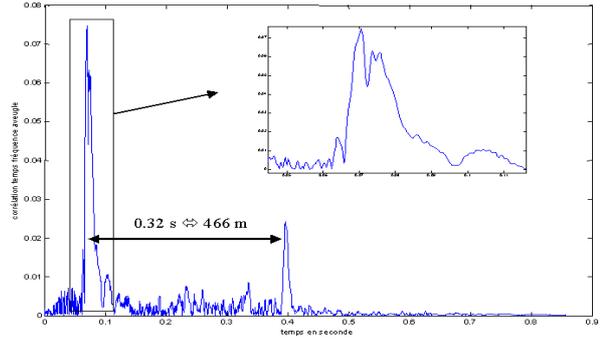and of the channel impulse response (see fig 4).

Figure 4 : Blind impulse response estimation using time-frequency processor applied on vocalize of the figure 3

The estimate impulse response shows two groups of beams separated by 0.32s, which represents a length difference of 466 meters. The order of magnitude of this difference is compatible with the lag due to one vertical go and back scan in the water column and tends to prove that the first beams are direct beams and the second ones are surface or surface-bottom reflected beams (the receivers are close to the bottom). A zoom on the first group of arrivals presents multiples local maxima proving multiple arrivals existence (refracted and bottom reflected ones). Inversion to estimate the celerity profile requires a more resolved estimation of these first beams.

High-resolution algorithm applied on the same test sound provides the time frequency image of the arrival given fig.5 and the estimated impulse response fig.7.
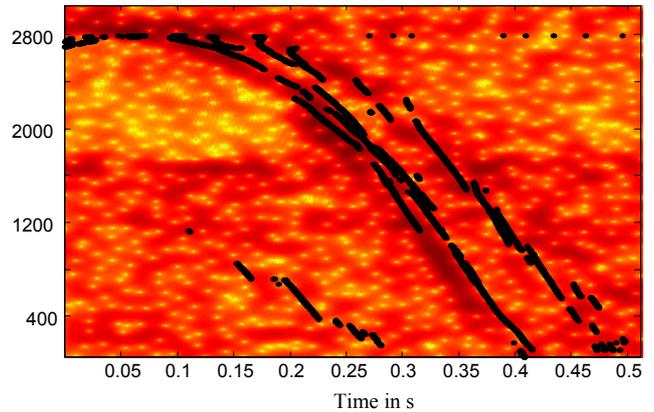

Figure 5. Spectrogram of the opportunity vocalize and the 4th order blind channel IR high-resolution algorithm solution

This algorithm seemed to give good results and succeeded in finding 4 coherent paths. However, to definitely validate our approach, we had to locate the source to be able to simulate the impulse response between beluga and receiver. For this, location is estimated by triangulation with relative time of arrival between all the hydrophones of the coastal array thanks to hyperboling fixing method described in [Spi99] .
Because of the small aperture in deep, we were only able to locate the whale in Latitude and Longitude (48,2657 °N, -69.4640°W). To handle this problem, we simulated the propagation in the channel thanks to Bellhop (parameterized with true bathymetry and sound speed profile) for a source positioned at the Latitude and Longitude found by hyperboling fixing and for all depths between 0 to 240m.

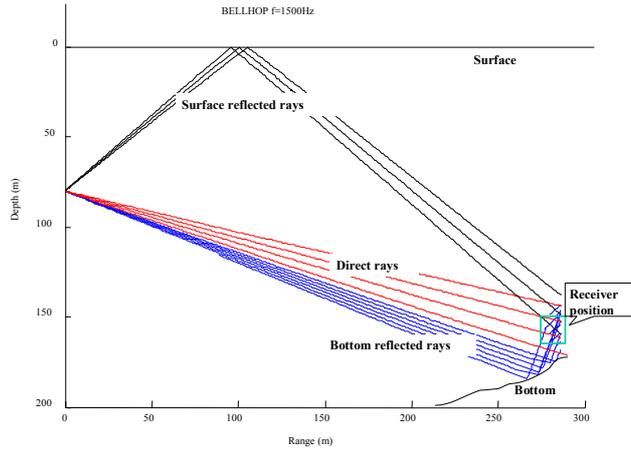On figure 6, an example of plots of the rays generated by BELLHOP is presented.

Figure 6. Plots of the rays generated by Bellhop obtained with a 80m-depth source.

Thus, several channel impulse responses are obtained as a function of depth, and they are compared to the impulse response estimated with the high-resolution algorithm described in the second paragraph. Finally, the more similar channel impulse response was found for an almost 80 meters depth.

On figure 7, a comparison between 5 channel impulse responses estimates is given:

- One simulated with Bellhop with a 80m depth source (fig.7 Curve 5)

- One obtained by an Active-Adapted Filter (we extracted the instantaneous frequency law of the real vocalize, then we constructed the corresponding signal and we made it forward in the simulated channel (fig.7 Curve 4)

- One given by high-resolution passive algorithm (fig.7 Curve 2)

- One given by low-resolution passive algorithm (fig.7 Curve 1)

- One obtained by an Active-Adapted Filter but with a smaller band. This band is equal to the larger chirp one, obtained during the limited development stage, actually 272 Hz. (fig.7 Curve 3)
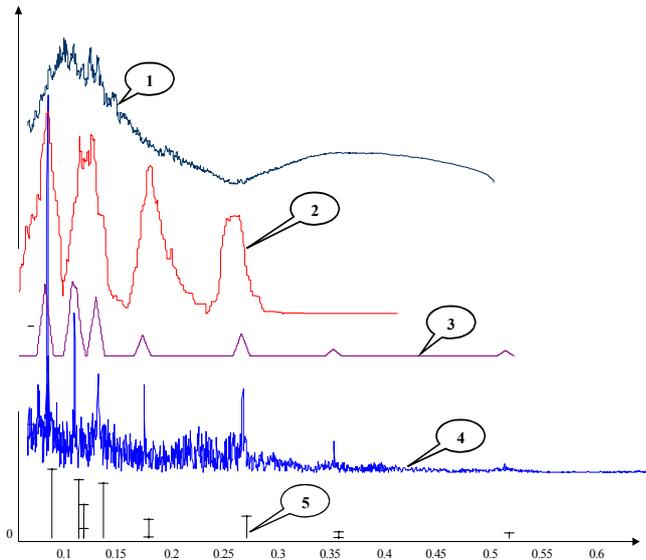

Figure 7. Comparison between Channel IR obtained with different methods

This comparison seemed to prove the efficiency of our high-resolution algorithm because of the good fit observed between theoretical time of arrival given by Bellhop and passive estimation of the impulse response given by the algorithms. The level of resolution offered by the high-resolution passive algorithm is still far from the active matched filter one but is very close to the one obtained by active matched filter with local time windowing. This level of resolution is satisfying enough in our case to explain the temporal structure of the arrival.

4. ESTIMATION OF SPATIAL AND TEMPORAL STRUCTURE OF MULTIPATHS ARRIVAL DEDICATED TO AuPT

Time-frequency-space representation (TFSR) has been recently introduced [Ami00] [Zha01] to estimate the direction of arrival of source signals that are localizable in the time frequency domain thanks to array of sensors. From the measurement of the acoustic field by an array of sensors, a square matrix TFSR is formed with the auto and cross time frequency representation of any couple of sensors. Under narrow-band assumption it is shown that TFSR, based on Cohen's class of time frequency distribution, follows the same model that the spatial autocorrelation matrix.

The application of the TFSR to high-resolution direction finding algorithm MUSIC outperforms classical MUSIC in the following critical cases :
-     close direction of arrival,
-     a number of sources higher than the number of sensors,
as soon as the components are well separated in time frequency space.

4.1 Signal Model

Let an Uniform Linear Array of M sensors receives L narrow-band transient signals coming from L unknown directions.
The M×1 vector of sensor outputs is modelled as :
$$\mathbf{x}(t) = [\mathbf{A}(\theta)]\mathbf{s}(t) + \mathbf{n}(t) \qquad (1)$$
where :

- $[\mathbf{A}(\theta)] = [\mathbf{a}(\theta_1).....\mathbf{a}(\theta_L)]$ is the M×1 matrix stacking the steering vectors,

- $\mathbf{a}(\theta_i)$ is the steering vector of the *i*th source,

- $\boldsymbol{\theta} = [\theta_1........, \theta_L]$ is the L×1 direction of arrival vector,

- $\mathbf{s}(t) = [\mathbf{s}_1(t),...,\mathbf{s}_L(t)]^T$ is the L×1 vector of sources waveforms

- $\mathbf{n}(t)$ is the M×1 vector of white sensor noise,

and $(.)^T$ stands for the transpose.

4.2 Spatial Time-Frequency analysis and DOA Estimation using transient signals

The discrete form of the spatial pseudo-Wigner-Wille distribution matrix (SPWVD) for any time frequency point is given by **:**
$$[\mathbf{TFSR}_{x_i x_j}(t_0, f_0)]_{NB} = WV_{x_i x_j}(t_0, f_0)$$
$$= \sum_{\tau=-(N-1)/.2}^{(N-1)/2} \mathbf{x}(t_0 + \tau)\mathbf{x}^H(t_0 - \tau)e^{-j4\pi f \tau} \qquad (2)$$
inserting (1) into (2) and taking the expectation, the TFSR model becomes :
$$E[\mathbf{TFSR}(t,f)]_{NB} = [\mathbf{A}(\theta)][\mathbf{WV}_{ss}(t,f)][\mathbf{A}^H(\theta)] + \sigma^2 \mathbf{I} \quad (3)$$

then the directions of arrival can be estimated by eigen-decomposition of the mean TFSR matrix and signal and noise subspace estimations.

When applications on real data are concerned, exact expectation cannot be achieved and an estimate is obtained by an average of 'instantaneous' TFSR matrices on a pre-selected time-frequency area Ω.

$$\overline{\textbf{TFSR}}_\Omega = \int_\Omega \textbf{TFSR}(t,f)dtdf \approx E[\textbf{TFSR}(t,f)] \qquad (4)$$

Main characteristics of TFSR depend on careful choice of area Ω. Choosing Ω where SNR is high improves detection, choosing Ω where only one single source takes place allows to obtain high resolution estimation of DOA and to deal with more sources than sensors.

4.3 Time-Frequency-Space Processor

The following scheme provides a synoptic implementation of the time-frequency-space processor applied to blind autonomous passive tomography.
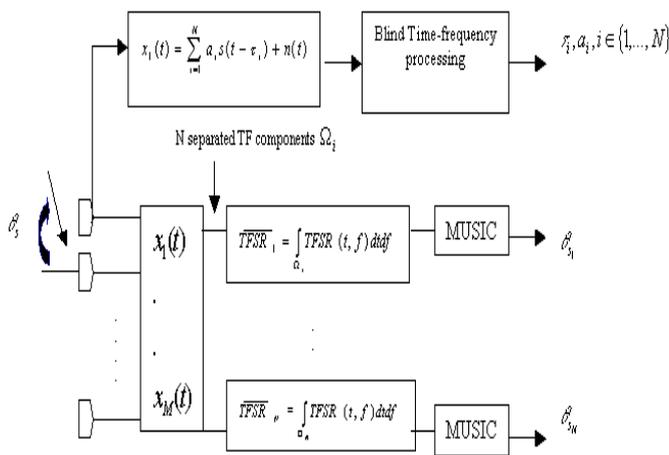

Figure 8. Time-Frequency-Space Processor

Assuming an ULA of M sensors and L arrivals coming from different DOA, by performing a single-sensor processing describes in section 2, the temporal structure of arrival (time of arrival and magnitude of arrival) is estimated. In addition, analyses conducted on time frequency content of the received signals allow to identify energetic time frequency area Ωi associated with a particular source and path. For each time frequency area Ω, a time-frequency-spatial processing is applied and is able to provide all DOA of the sources.

It is to be noted that our time frequency processor is efficient in the case of a well-solved arrivals in the time frequency plane, which means that the arrivals should be wideband signals. But this fact introduces a bias and a loss of resolution in DOA measurements after spatial processing [Del02] . For this reason the narrowband MUSIC algorithm is replaced with the wideband-MUSIC one [Hsi88] [Wan85] and in the future we will deal with the WIDEBAND-TF-MUSIC [Ger00] to rectify the loss of resolution during spatial-time-frequency processing.

4.4 Simulation Results

In all simulations, the case of a realistic propagation in a horizontal uniform celerity profile shallow water channel (200m depth and 1000m range). We assume that the source and the receiver are close to the surface and the distance in-between is equal to the channel's range ; the source and the receiver positions correspond to 20m and 30 m depth, respectively.

Unfortunately, in that case of critical geometric source and receiver configurations, the arrivals are not well-solved spatially and in the time-frequency plane, it is also important to note that the arrivals are classified by echoes of four. Figure 9 displays the impulse and the spatio-temporal responses of the channel.
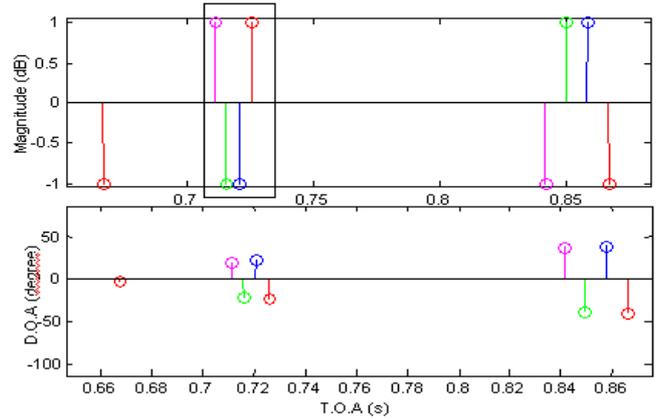

Figure 9. Impulse and spatio-temporal responses of the channel

In what follows the first four echoes squared in will be studied and respectively correspond to :
$\theta = [-21.3, -23.2, 20.3, 22.3]; \tau = [0.71, 0.715, 0.72, 0.725]$.

The receiver is a ULA of M=36 sensors spaced of $\lambda/3$ apart and the sources are considered first as narrowband ($B = 85Hz$) and then as wideband LFM ($B = 2Khz$) around a central frequency $f_0 = 9Khz$.

In these figures the fixed $SNR=15dB$ has been assumed and only the positive angles are drawn for better illustration.
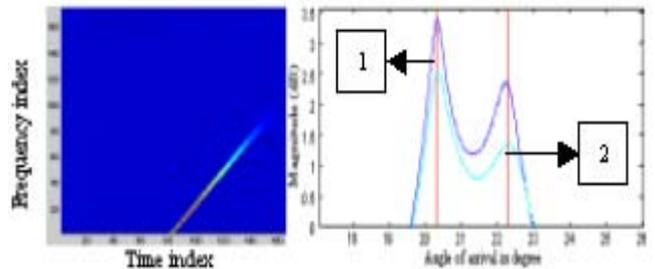

Figure.10
The left one displays a WV time-frequency representation behind the first sensor of the array.
The right one displays :
1 – narrowband MUSIC applied to narrowband signals
2 – narrowband TF MUSIC applied to narrowband signals

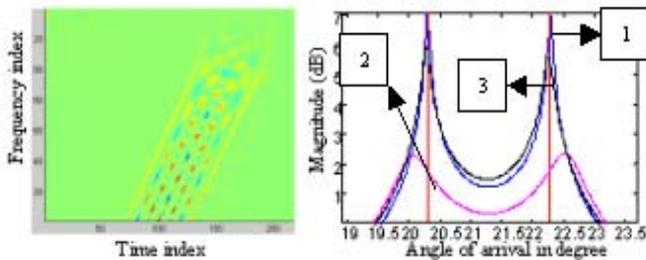Figure.11
The left one displays a WV time-frequency representation behind
the first sensor of the array.
The right one displays :
1 – narrowband MUSIC applied to narrowband signals
2 – narrowband TF MUSIC applied to wideband signals
3 – wideband MUSIC applied to wideband signals

Figure 10 displays spatial spectra, time-frequency-space spatial spectra and time frequency representation of signal measure at first sensor of the array, in the case of narrowband signal. In this configuration, spatial narrow band processors perform well and are able to resolve DOA of each path (MUSIC narrowband and TFSR-MUSIC) but it is important to note that the time frequency representation does not resolve the four echoes and so complete spatial and temporal characterisation fails.

Figure 11 displays spatial and time-frequency-space spatial spectra and time-frequency representation of signal measure at first sensor of the array in the case of wideband signal. In this case, narrow band TFSR spatial processing failed (bias and loss of resolution) (MUSIC narrowband applied to wide-band signals) and spatial wideband MUSIC processing succeeds in nulling bias. Improvement in angular resolution will be offered as a perspective by wideband spatial time frequency processing. In this case, the time frequency representation resolves the 4 echoes and so, a complete spatial and temporal characterisation is achieved.

5. CONCLUSION

In this communication, we have presented two blind channel impulse response estimation algorithms. We have demonstrated their capabilities to perform the identification of the impulse response channel without using the knowledge of the emitted source signal in the case of a single hydrophone.
We have succeeded in applying them to real data obtained from Laurentian channel experiment performed in summer 2003.
Performances obtained with the high-resolution algorithm are close to classical active matched filtering methods. We will definitely validate our algorithms with the remained Laurentian channel data (data from the five ocean bottom hydrophones deployed in the centred square configuration cf. fig.1), which will allow us to locate more precisely the emitted whales.
The time-frequency-space processor has been validated in a realistic context of underwater acoustic propagation.
The bias and the loss of resolution introduced by the wideband sources were rectified by applying specific wideband spatial treatment.

Additional work will be carried out applying WIDEBAND-TF-MUSIC [6] to deal with the time frequency-space of the same problem.

6. REFERENCES

[Fla88]   P.Flandrin, *A time-frequency formulation of optimum detection*, IEEE Transactions on acoustics, speech and signal processing, vol 36, N°9, pp.1377-1384, september 1988
[Bar93] R.G.Baraniuk, D.L.Jones, *Signal-dependent time-frequency analysis using a radially Gaussian kernel*, Signal processing, 32, pp. 263-284, 1993
[Ger01] . Gervaise, A. Quinquis, I. Luzin, High resolution of the underwater channel from unknown transient stimuli, Gretsi 2001, September 2001, Toulouse, France
[Sim04] Y. Simard, M. Bouhara, N. Roy, *Acoustic detection and localization of whales in Bay of Fundy and St Lawrence Estuary critical habitats*, Canadian Acoustics 32(2): 107-116. NMML Periodicals Collection - TD893.5 .C16 v.32 no.2, 2004.
[Spi99] Spiesberger, J.L, *Locating animals from their sound s and tomography of the atmosphere: Experimental demonstration*. J. Acoust.Soc. Am. 106: 837-846,1999.
[Ami00] Moeness G.Amin, Yimin Zhang, *Direction finding based on spatial time-frequency distribution matrices*, Villanova University Pennsylvania, 10, 325-339, 2000.
[Zha01] Yimin Zhang, Moeness G. Amin, *Subspace Analysis of Spatial Time-Frequency Distribution Matrices*, IEEE Signal Processing, Vol 49, n°4, April 2001.
[Del02] J-P. Delmas, Y. Meurisse, *Robustness of narrowband DOA algorithms with respect to signal bandwidth*, 15 February 2002.
[Wan85] H. Wang, M. Kaveh, *Coherent Signal Subspace processing for the Detection and Estimation of angles of Arrival of Multiple Wide-Band Sources*, August 1985.
[Hsi88]   Hsiensen Hung, Mostafa Kaveh, *Focussing Matrices for Coherent Signals-subspace Processing,* August 1988.
[Ger00] Gershman and M. Amin, "*Coherent wideband DOA estimation of multiple FM signals using spatial time-frequency distributions*," Proceedings of the IEEE International Conf. on Acoustics, Speech, and Signal Processing, Istanbul, Turkey, June2000.