De-noising procedures for inverting underwater acoustic signals in applications of acoustical oceanography

Michael Taroudakis and Costas Smaragdakis
Department of Mathematics and Applied Mathematics, University of Crete and Foundation for Research and Technology-Hellas Institute of Applied and Computational Mathematics, Voutes University Campus, 70013 Heraklion Crete, Greece.

Summary
Applications of ocean acoustic tomography and sea-bed classification using measurements of acoustic signals emitted from known sources are based on the extraction of specific features of the signal and their subsequent exploitation by post processing. The extraction of these features is sometimes very difficult to be done in a reliable way, due to the fact that the measurements are made in the presence of noise, which in some cases is a severe handicap for the exploitation of the measured signals. Although several inversion methods have been proposed to by-pass the noise problem, the issue is open to further research and de-noising strategies are in the process of being studied. In this paper we focus on a specific inversion method based on the statistical characterization of the acoustic signal [1] and compare a couple of alternative de-noising procedures in order that the statistical features of the signal extracted from the noisy measurements could lead to reliable inversions of the critical parameters of the sea-bed and or the water column. The results presented are based on synthetic signals produced using typical characteristics of tomograpahic sources and adding noise to a level appropriate for simulating realistic experiments in a shallow water environment.

PACS no. 43.30.Pc, 43.60.Pt

1. Introduction

Signals used for underwater acoustic applications due to some specific source are contaminated by noise. The noise may be of physical origin (natural sources) or may be due o the electronics of the measuring system. The exploitation of the signals for applications of acoustical oceanography are based on the identification of signal observables which are independent of noise, or by applying post-processing techniques to the raw signals recorded in the ocean environment to reduce the noise effect. It is obvious that denoising strategies are largely based on the method to be applied for the exploitation of the signal and the anticipated applications. In the present work we will refer to a statistical signal characterization method which has been shown to be efficient for tomogrpahic and geoacoustic inversion applications [1], The method has been used recently with both simulated and real data and the conclusions from the analysis of the results of the studie! s showed that the quality of the inversion results is largely dependent on the noise level of the signal. In this work we compare two alternative denoising strategies using simulated data with added noise. For completeness reasons, Chapter 2 presents shortly the signal characterization method, Chapter 3 presents our first attempt to reduce noise effects of the recorded signal, while Chapter 4 presents a new denoising strategy which, based on the analysis of first results, shows to be efficient in reducing noise effects. Some initial results are presented in Chapter 5.

The whole analysis is based on the following notation :

The noise free signal is x, the noise component is w and the actual signal is S. This is a deterministic treatment of the problem, the statistical character of the noise introduced by assuming Gaussian white noise (AGWN). Thus the signal which is recorded with N samples is given by

$$\mathbf{S}[n] = \mathbf{x}[n] + \mathbf{w}[n], \ n = 1, \ldots, N \tag{1}$$

2. The statistical characterization scheme

2.1. Statistical Characterization

The details of the signal characterization based on the statistical analysis of the section wavelet sub-band coefficients have been presented in other publications and will not be repeated in detail here. The interested reader will find extensive analysis of the method in previous works of Taroudakis et al. [1, 2, 3]

However, for completeness, the outline of the method will be presented here. Consider an acoustic signal S. According to the method, the signal is characterized by means of the statistical parameters of the coefficients resulting from the application of a 1-D Discrete Wavelet Transform (DWT) to the discrete signal $\langle S, \psi_{a,b}\rangle$, where $\psi_{a,b}$ is an appropriately chosen wavelet, with subsequent convolution by a High-Pass and a Low-Pass filter giving two sets of coefficients called detailed $\mathbf{d}_1[n;S]$, and approximate, $\mathbf{a}_1[n;S]$. In our analysis the Daubechies' (db4) wavelet [4] is used. By continuing this process using the detailed coefficients up to the k^{th} level of decomposition the signal is represented as a first step by the vectors of coefficients obtained through this multilevel analysis. The approximate coefficients are kept at the final level only. It has been shown in [1] that the coefficients of the wavelet coefficients of a typical underwater signal emitted from a Gaussian source obey a Symmetric Alpha Stable distribution (SaS) described by its characteristic function:

$$\Phi(t) = \exp(i\delta t - \gamma^{\alpha}|t|^{\alpha}), \quad (2)$$

where $0 \leq \alpha \leq 2$ is the characteristic exponent which controls the marginal behavior of the tails, $-\infty < \delta < +\infty$ is the local parameter, γ is the dispersion of the distribution, which determines the spread of the distribution around the local parameter δ and t is the value of the coefficient.

In our case $\delta = 0$ and the signal S is eventually characterized by a vector $\mathbf{d}$ of dimensions $2L+2$ as following:

$$S \leftrightarrow \mathbf{d} = (\alpha^0, \gamma^0, \alpha^1, \gamma^1, \ldots, \alpha^L, \gamma^L), \quad (3)$$

where L is the total number of levels considered.

2.2. The Kullback Leibler Divergence (KLD) cost function

The Kullback Leibler Divergence (KLD) which expresses the difference (or distance) D_S between two acoustic signals S_1 and S_2, when these signals are characterized by some statistical distribution of selected coefficients [5]. In the case of two signals represented by the parameters of the SaS distributions of the wavelet sub-band coefficients as described above, the KLD is expressed in closed form according to the following formula:

$$D_S(S_1,S_2) = \sum_{k=0}^{L}\left\{\ln\left(\frac{c_2^k}{c_1^k}\right) - \frac{1}{\alpha_1^k} + \left(\frac{\gamma_2^k}{\gamma_1^k}\right)^{\alpha_2^k}\frac{\Gamma(\frac{\alpha_2^k+1}{\alpha_1^k})}{\Gamma(\frac{1}{\alpha_1^k})}\right\}, \quad (4)$$

where $\Gamma(x)$ is the Gamma function and

$$c_i^k = \frac{2\Gamma(\frac{1}{\alpha_i^k})}{\alpha_i^k\gamma_i^k}, i = 1,2 \quad , k = 0,\ldots,L. \quad (5)$$

Formula (4) is based on the assumption that the statistical character of the wavelet coefficients at each level is independent to that of another level.

3. Keeping the efficient energy part of the signals

Consider a discrete acoustic signal represented by N samples and denote them as $\{S[n]\}$ with $n = 1,\ldots,N$. Let a, b be integer numbers, with $2 \leq a < b \leq N-1$ and $A \subset \{1,2,\ldots,N\}$ be an index set. We denote the restriction of the signal S into the above set A as

$$S|_A[n] = \begin{cases} S[n], & \text{if } n \in A \\ 0, & \text{otherwise} \end{cases} \quad (6)$$

We define a signal partition which consisted of three sub-signals as:

$$\begin{aligned} S_L(a) &= S|_{[1,a)\cap\mathbb{Z}} \\ S_C(a,b) &= S|_{[a,b]\cap\mathbb{Z}} \\ S_R(b) &= S|_{(b,N]\cap\mathbb{Z}}. \end{aligned} \quad (7)$$

The initial signal S could be reconstructed from its partitioning, using the simple additive relation

$$S = S_L(a) + S_C(a,b) + S_R(b). \quad (8)$$

The signals $S_L(a), S_C(a,b)$ and $S_R(b)$ have disjoint supports, hence the energy norm obey the following additive law

$$\|S\|_2^2 = \|S_L(a)\|_2^2 + \|S_C(a,b)\|_2^2 + \|S_R(b)\|_2^2. \quad (9)$$

Choosing integer numbers a_*, b_* for any pair $\epsilon_1, \epsilon_2 \in (0,1)$ with restriction $\epsilon_1 + \epsilon_2$ to be also in $(0,1)$ as

$$a_* = \sup_a\left(\|S_L(a)\|_2^2 \leq \frac{\epsilon_1}{\epsilon}(1-\epsilon)\|S\|_2^2\right) \quad (10)$$

$$b_* = \inf_b\left(\|S_R(b)\|_2^2 \leq \frac{\epsilon_2}{\epsilon}(1-\epsilon)\|S\|_2^2\right), \quad (11)$$

then the central projected part $S_C(a_*, b_*)$ has about 100ϵ percent of the energy of the whole signal, where

$$\epsilon = \epsilon_1 + \epsilon_2. \tag{12}$$

Typical recordings corresponding to signals from sources utilized in ocean acoustic tomography experiments show concentration of energy around the central time of the recording while noise has important contribution in the whole signal. On the other hand, when considering the statistical signal characterization, the actual signal has no effective energy at the S_L, S_R parts of it and therefore the wavelet coefficients are estimated with narrow SaS distributions and small dispersion parameters γ. These parameters exhibit constant behavior to the small signal perturbations which are involved in inversion processes.

In view of the above observations, we studied a strategy to reduce the noise effect which is based on the use of the the central parts of signal $S_C(a^*, b^*)$ for signal characterization and subsequent inversion procedure. This part will be denoted in the sequel as "cropped" signal. The parameters a^* and b^* are chosen so that the effective energy of the signal is between 90 and 98 % of the total signal energy. It is obvious that the lower limit corresponds to noisy data, as noise is present in the whole signal while for noise free data the upper bound gives a safe limit for the inclusion of the energy significant part of the signal in the characterization process. The values of epsilons (ϵ_1, ϵ_2) are proportional to the choice of ϵ as well as the position of the effective part of signal in the time axis. Denoting by t_c the sample number for which the cumulative energy distribution has value equal or close enough to 0.5, the following relation

$$\frac{\epsilon_1}{\epsilon_2} = \frac{t_c}{N - t_c}, \tag{13}$$

in connection with the relation (12) give an appropriate pair of the coefficients

$$(\epsilon_1, \epsilon_2) = \left(\frac{t_c}{N}\epsilon, \frac{N - t_c}{N}\epsilon \right), \tag{14}$$

whose values are following the energy interpretation as described above.

4. Sparse denoising strategy

4.1. Sparse Decomposition

Given a digitally recorded acoustic signal S with N samples, we divide this into overlapping windows S_k, each of length L using the maximum overlapping rate of $L-1$ samples [6]. Hence, the $k-th$ window S_k is given by :

$$S_k = \{S[i],\ i \in \mathbb{Z} \cap [k, k + L - 1]\}$$

where the index $k \in \mathbb{Z} \cap [1, N - L + 1]$. For the cases subsequently studied and despite the fact that the approach is computationally expensive, we have chosen a large number of windows. With this manner we have achieved a more detailed matching between model and signal. For simplicity reasons, let us define $N - L + 1$ as M.

Nest step is to find a dictionary $\mathbf{D} = \{g_p\}_{p \in M}$ in which, each specific window S_k has been estimated in a proper sparse form as $\tilde{S}_k$:

$$\tilde{S}_k = \sum_{p \in M} a_{kp} g_p,$$

given a quite small approximation error $\sum_k \|S_k - \tilde{S}_k\|_2 \leq \epsilon$ and the same time the coefficients matrix $\mathbf{A} = \{a_{kp}\}$ to be in sparse form as well.

4.2. Inference of sparse codes

Assuming that the dictionary matrix $\mathbf{D} \in \mathbb{R}^{L,M}$ is given, we define $\ell(\mathbf{a}_j; \mathbf{D})$ by means of the following expression:

$$\ell(\mathbf{a}_j; \mathbf{D}, \mathbf{S}_j) = \\ = \min_{\mathbf{a}_j \in \mathbb{R}^M} \left\{ \frac{1}{2} \|\mathbf{S}_j - \mathbf{D}\mathbf{a}_j\|_2^2 + \lambda \|\mathbf{a}_j\|_1 \right\} \tag{15}$$

We have to infer the best matching $\mathbf{a}_j$ by solving the previous problem using a simple gradient descent method. The initial step is to calculate the gradient of $\ell(\mathbf{a}_j; \mathbf{D}, \mathbf{S}_j)$ with respect to matrix $\mathbf{D}$.

$$\bigtriangledown_{\mathbf{a}_j} \ell(\mathbf{a}_j; \mathbf{D}, \mathbf{S}_j) = \mathbf{D}^T(\mathbf{D}\mathbf{a}_j - \mathbf{S}_j) + \lambda\, \mathrm{sign}(\mathbf{a}_j) \tag{16}$$

The gradient has two components, the first one is related to the reconstruction term whereas the second one comes from the ℓ_1 norm which appears in the sparsity term.

One critical issue is the fact that the above gradient function is not differentiable at zero. So, the usage of a classical gradient descent method is not acceptable in this case.

4.3. Iterative Shrinkage and Thresholding Algorithm (ISTA)

The proximal mapping of a convex function h is

$$\mathrm{prox}_h(\mathbf{x}) = \arg\min_{\mathbf{u}} \left(h(\mathbf{u}) + \frac{1}{2}\|\mathbf{u} - \mathbf{x}\|_2^2 \right). \tag{17}$$

When $h(\mathbf{x}) = \lambda \|\mathbf{x}\|_1$ the i-th element of the solution has been proven to be [7]

$$\mathrm{prox}_h(\mathbf{x})_i = \\ = \begin{cases} \mathbf{x}_i - \lambda\, \mathrm{sign}(\mathbf{x}_i - \lambda) & \text{if } |\mathbf{x}_i| > \lambda \\ 0 & \text{otherwise.} \end{cases} \tag{18}$$

Looking up again at $\ell(\mathbf{a}_j; \mathbf{D}, \mathbf{S}_j)$ we can easily observe that this function can be decomposed into two convex functions g and h as:

$$g(\mathbf{a}_j; \mathbf{D}, \mathbf{S}_j) = \frac{1}{2}\|\mathbf{S}_j - \mathbf{D}\mathbf{a}_j\|_2^2 \tag{19}$$

and

$$h(\mathbf{a}_j) = \lambda\|\mathbf{a}_j\|_1 \tag{20}$$

Note that the first function is quadratic, whereas the second is a l_1 function.

Each iteration of the ISTA generalized gradient descent algorithm [8] includes two phases. First, it performs a single step gradient descent algorithm to the reconstruction error term and then does an update based on the sparsity term using the proximal mapping theory. The whole procedure is expressed in **Algorithm** 1 bellow.

Algorithm 1 Iterative Shrinkage and Thresholding Algorithm

procedure ISTA($\mathbf{D}$,$\mathbf{S}_j$) ▷ Required the dictionary.
  Initialize $\mathbf{a}_j^0$ ▷ Randomly chosen
  $k \leftarrow 0$
  repeat
    $\tilde{\mathbf{a}}_j^{k+1} \leftarrow \mathbf{a}_j^k - t\ \mathbf{D}^T(\mathbf{D}\mathbf{a}_j^k - \mathbf{S}_j)$ ▷ update from reconstruction term
    $\mathbf{a}_j^{k+1} \leftarrow \text{prox}_h(\tilde{\mathbf{a}}_j^{k+1})$ ▷ update from the sparsity term
    $k \leftarrow k+1$
  until $\|\mathbf{a}_j^{k+1} - \mathbf{a}_j^k\|_2 < \text{tol}$
  $\hat{\mathbf{a}}_j \leftarrow \mathbf{a}_j^{k+1}$
  return $\hat{\mathbf{a}}_j$

4.4. Online dictionary learning algorithm

We see our initial problem of total optimizing again that was previously presented. The purpose of this section is to choose the best possible dictionary $\mathbf{D}$ for the best description of the signal characteristics. For this end, we shall employ the algorithm of online dictionary learning with mini-batch extension proposed by J. Mairal et al [8].

$$\min_{\mathbf{D}} \frac{1}{M}\sum_{j=1}^{M} \min_{\mathbf{a}_j} \ell(\mathbf{a}_j; \mathbf{D}, \mathbf{S}_j) \tag{21}$$

Let us assume that $\mathbf{a}_j$ does not depend at all on dictionary choice. This is generally not true but if we consider that $\mathbf{D}$ changes slowly enough we can use this simplification without issues. So, the minimization problem is written in the form

$$\min_{\mathbf{D}} \frac{1}{M}\sum_{j=1}^{M} \frac{1}{2}\|\mathbf{S}_j - D\hat{\mathbf{a}}_j\|_2^2 \tag{22}$$

Note that only the reconstruction term contributes to the dictionary learning procedure.

It is crucial that we choose matrix $\mathbf{D}$ to have all columns with unit Euclidean norm. We remind that the columns of $\mathbf{D}$ need not be orthogonal (linear independency is not required).

Having assumed that the training set composed of independent and identically distributed (i.i.d) samples of a distribution function $p(x)$, here p is simply an typical uniform distribution, dictionary training algorithm draws η elements (a mini-batch) and adapts the dictionary in order to achieve a better matching with the members of the mini-batch. This procedure is being iterated for a quite big number of steps (in our case $T = 1000$).

Algorithm 2 Online Dictionary Learning with mini-batch extension

procedure DictLearning
  $\mathbf{A} \leftarrow 0^{M,M},\ \mathbf{B} \leftarrow 0^{L,M}$ ▷ No prior information
  $k \leftarrow 0$
  for $t = 1,\ T$ **do**
    Drawn $\{\mathbf{S}_{t_i}\}_{i \in [1,\ \eta] \cup \mathbb{Z}}$ from $p(x)$
    for $i = 1,\ \eta$ **do**
      $\hat{\mathbf{a}_{t_i}} = \text{ISTA}(\mathbf{D}, \mathbf{S}_{t_i})$, ▷ using **Algorithm 1**
      if $t < \eta$ **then**
        $\theta \leftarrow t\eta$
      else
        $\theta \leftarrow \eta^2 - \eta + t$
      $\beta \leftarrow (\theta - \eta + 1)/(\theta + 1)$
      $\mathbf{A} \leftarrow \beta\mathbf{A} + \hat{\mathbf{a}}_{t_i}\hat{\mathbf{a}}_{t_i}^T$,
      $\mathbf{B} \leftarrow \beta\mathbf{B} + \mathbf{S}_{t_i}\hat{\mathbf{a}}_{t_i}^T$,
    for $j = 1,\ N$ **do**
      $\mathbf{u}_j \leftarrow \frac{1}{\mathbf{A}_{jj}}(\mathbf{b}_j - \mathbf{D}\mathbf{a}_j) + \mathbf{d}_j$
      $\mathbf{d}_j \leftarrow \frac{1}{\max(\|\mathbf{u}_j\|_2, 1)}\mathbf{u}_j$ ▷ Update the j column of dictionary
  return D ▷ Return the trained dictionary

4.5. Signal reconstruction

Once dictionary $\mathbf{D}$ has been trained and the coefficient vectors $\hat{\mathbf{a}}_j$ has been calculated for all windows $\mathbf{S}_j$, we are capable of approximating a denoised version $\hat{\mathbf{S}}_j$ for each window frame $\mathbf{S}_j$ as

$$\hat{\mathbf{S}}_j = \mathbf{D}\ \hat{\mathbf{a}}_j, \quad j = 1, \ldots, M. \tag{23}$$

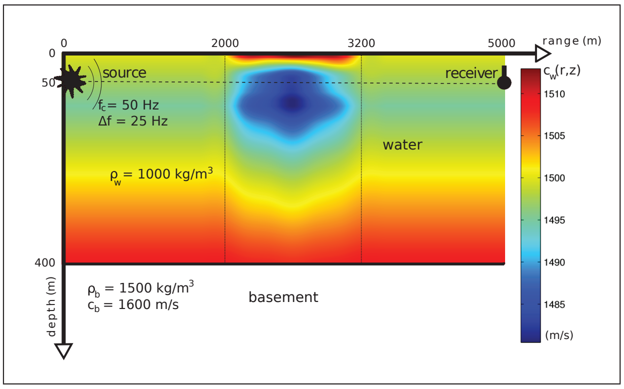

Figure 1. The simulated shallow water environment with a cold eddy.

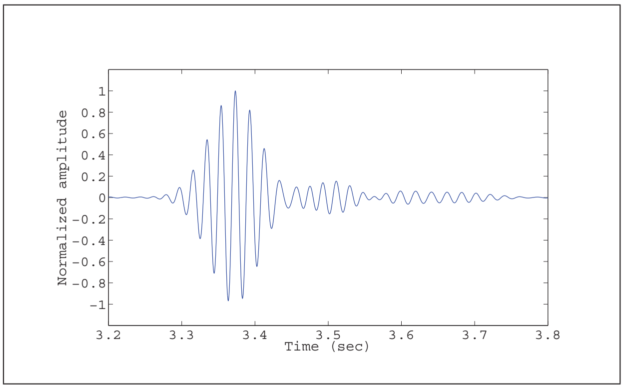

Figure 2. The simulated noise-free signal

Our final goal is to construct a denoised version of the signal using the approximations $\hat{\mathbf{S}}_j$ of overlapping windows. We have got $c(n)$ number of segments each of them includes an estimation of $n-$th sample of the signal, where $c(n)$ is given by the function :

$$c(n) = \begin{cases} n & \text{if } n \in [1, L) \\ L & \text{if } n \in [L, N-L] \\ N-n+1 & \text{if } n \in (N-L, N]. \end{cases} \quad (24)$$

We are able to define an approximation of actual signal $\mathbf{x}$ as the expected value of all mentioned approximations. We would like to make clear that we have been working under the assumption that every approximation has the same impact to the final estimated value. Therefore, we get the approximation as a typical mean value of the observations, hence

$$\hat{\mathbf{x}}[n] = \frac{1}{c(n)} \sum_{k=\max(1,n-L+1)}^{\min(n,M)} \hat{\mathbf{S}}_k[n-k+1] \quad (25)$$

5. Testing with a synthetic signal

We are now testing the denoising strategies presented above, using synthetic signals.

As a test case, we are considering a simulated shallow water environment with a cold eddy. This environment has been used in a previous work [9]. A tomographic experiment is considered using a Gaussian source and a single receiver. The environment is shown in Figure 1 and the simulated recorded noise-free signal is shown in Figure 2.

We add noise to this signal so that the noisy signal has SNR, 5, 10 and 17 dB and we apply the new denoising strategy with whole or cropped signals.

5.1. Improvement measure of the characterization scheme

In order to assess the contribution of the denoising schemes to the quality of the statistical characterization, we define a measure in the logarithmic scale, expressed in dB employing the KLD as following :

$$I(\hat{\mathbf{S}}_1, \hat{\mathbf{S}}_2; \mathbf{S}) = -10 \log_{10} \left(\frac{D_s(\hat{\mathbf{S}}_1, \mathbf{S})}{D_s(\hat{\mathbf{S}}_2, \mathbf{S})} \right) \quad (26)$$

where $\hat{\mathbf{S}}_i$ are the signals after the denoising processing and $\mathbf{S}$ the noise free version of the simulated signal (whole or cropped).

Using this measure of course we assume that we know the noise-free signal, which is possible as we are dealing with simulated (synthetic) signals. When two signals are identical the KLD is 0 as their statistical characterization is also identical. When KLD has a low value, the denoised version of the signal is close to the noise-free version. Therefore, comparing two KLDs the lowest value indicates better correspondence between actual and noisy signal. Thus, by means of this formula we have a first indication on the efficiency of the denoising strategy. Larger value of I indicates better denoising effect.

In Table I we illustrate the comparison among possible choices of denoising strategies. We are applying the sparse denoising strategy before or after the cropping of the signal as described in Section 3, thus obtaining four different procedures. We observe an important improvement of signal characterization when we apply the sparse denoising procedure to the measured signal, with the improvement reaching its absolute maximum when this procedure is applied to the whole signal, as expected. Moreover we observe that the improvement is more pronounced in cases where we have signals with low SNR.

6. CONCLUSIONS

We have presented the impact of a new sparse denoising scheme to the quality of characterization of underwater acoustic signals under different denoising scenaria involving or not cropping of the recorded signal to isolate the energy efficient part of it. The new

Table I. Comparison between signals as regards their characterization quality. The measure I express the amount (in dB) of characterization improvement when applying alternative denoising strategies

comparison between	I (dB), SNR 17 dB	I (dB), SNR 10 dB	I (dB), SNR 5 dB
denoised (whole) and noisy (whole)	62.83	59.50	68.88
denoised (whole) and noisy (cropped)	15.21	12.77	22.44
denoised (cropped) and noisy (cropped)	28.31	36.24	48.64
denoised (cropped) and denoised (whole)	13.11	23.47	26.21

method is considered among the most efficient techniques of denoising and feature extraction of waveforms and images. As expected, the characterization quality was significantly improved at least for the case studied.

However, in order to get more conclusive results, we have to validate these schemes by applying the signal characterization scheme to inverse problems of ocean acoustic tomography and sea-bed classification. The efficiency of the denoising strategy will be assesed on the basis of the inversion results which will correspond to the estimation of the critical geoacoustic parameters of the environment.

Acknowledgement

Work supported by PEFYKA project within the KRIPIS action of the GSRT. The project is funded by Greece and the European Regional Development Fund of the European Union under the NSRF and the O.P. Competitiveness and Entrepreneurship.

References

[1] M. Taroudakis, G. Tzagkarakis, and P. Tsakalides: Classification of shallow water acoustic signals via alpha stable modeling of the one dimensional wavelet coefficients, J. Acoust. Soc. Am. Vol 119, 1396–1405, 2006.

[2] G. Tzagkarakis, M. Taroudakis, and P. Tsakalides: A statistical geoacoustic inversion scheme based on a modified radial basis functions neural network, J. Acoust. Soc. Am. Vol 122, 1959-1968, 2007.

[3] M. Taroudakis and C. Smaragdakis: Tomographic and bottom geoacoustic inversions using genetic algorithms and a statistical characterization of the acoustic signal, Acta Acoustica united with Acoustica, Vol 95(5), 814-822, 2009.

[4] C. Chui: An Introduction to Wavelets, Academic Press, San Diego, 1992, pp 215-243.

[5] S. Kullback: Information Theory and Statistics, Dover Publications, N.Y, 1968, pp. 6-7.

[6] Maria Jafari, Mark D. Plumbley: Fast dictionary learning for sparse representations of speech signals. IEEE journal of selected topics in Signal Processing, special issue on Adaptive Sparse Representation of Data and Applications in Signal and Image Processing, IEEE Signal Processing Society, 2011, 5 (5), <10.1109/JSTSP.2011.2157892>. <inria-00599051>.

[7] N. Parikh, S. Boyd: Proximal algorithms, Foundations and trends in optimization, Vol 1, No 3, 2014, 123-231.

[8] J. Marial, J. Ponce, G. Sapiro: Online dictionary learning for sparse coding, Appearing in Proceedings of the 26-th International Conference on Machine Learning, Montreal, Canada, 2009.

[9] M. Taroudakis and C. Smaragdakis: Inversions of Statistical Parameters of an Acoustic Signal in Range-Dependent Environments with Applications in Ocean Acoustic Tomography, J. Acoust. Soc. Am. Vol 134(4), 2814-2822, 2013.