CANCELING SINUSOIDAL INTERFERENCES IN ULTRASONIC APPLICATIONS WITH A BSS ALGORITHM FOR MORE SOURCES THAN SENSORS

Jorge Igual, Andrés Camacho, Luis Vergara

Departamento Comunicaciones
Universidad Politécnica Valencia
Camino de Vera, S/N, Valencia 46020, Spain,
e-mail: jigual@dcom.upv.es

ABSTRACT

Sinusoidal interferences are found in ultrasonic signals when we try to characterize a material, as for example interferences coming from PC cards. We are interested in obtain a robust method that cancels these interferences preserving the waveform of the signal. A Blind Source Separation BSS method to extract these sinusoids is presented in this paper. We will get so many linear mixtures of the backscattering echo of the material and the sinusoids as we need from different pulse responses of the material.[1]

1. INTRODUCTION

The problem of Blind Source Separation (BSS) consists on recovering a set of independent source signals from linear mixtures of them [1]. We will introduce an application of BSS to cancel sinusoidal interferences that are found in measures of ultrasonic signals. In this case we will call the source signals the pulse response of the material and the sinusoids.

We are interested in characterize different kinds of material starting from its ultrasonic scattering [2]. We use a transducer with central frequency 44 kHz. At these frequencies, some interferences are found in the back scattered echo signals. These interferences should be cancelled preserving the waveform of the echo of the material to avoid their effect on different parameter estimates to be used in material characterization, as centroid frequency, probability density parameters...

We are looking for an algorithm that extracts the pulse response of the material without the harmonic interferences. Thus, a BSS method can be applied, supposing that pulse response and harmonics are statistically independent. An additional advantage of BSS is its robustness, so it works very well in low S/N ratios and when the interference is inside the frequency band where most of signal energy is found.

In BSS we need so many mixtures as signals we want to recover. We only have one sensor, so only one signal can be recorded. However, we will obtain the different linear mixtures recording several pulse responses, as we will explain in Section 2, where the problem is mathematically formulated. In Section 3 a BSS solution is presented in a general case and in Section 4 a simulated example is shown.

2. PROBLEM STATEMENT

The mixture model is:

$$y(t) = x(t) + \sum_{i=1}^{N-1} B_i e^{j(w_i t + q_i)} \tag{1}$$

where *y(t)* is the received signal, *x(t)* the backscattering echo and $B_i e^{j(w_i t + q_i)}$ $i = 1 \ldots N-1$ the sinusoidal interferences to be cancelled. BSS definition supposes that there are at least so many linear independent mixtures as sources we want to recover, i.e., we need a mixture model with N mixtures. Usually to register N signals we need N sensors. However in our case we can relax this condition because a prior information about the original sources is available. This fact brings about that we can implement our algorithm with only one sensor, what would reduce highly the cost of a real implementation. In Figure 1, we show graphically how we can overcome the problem of having only one sensor.

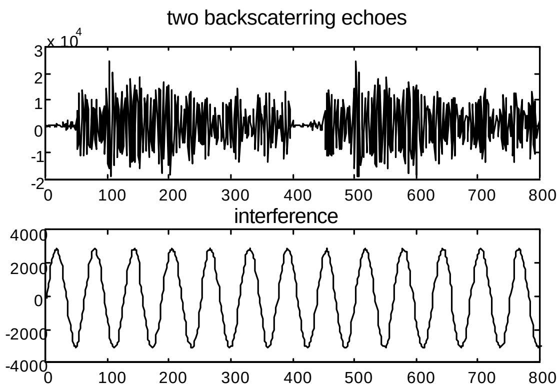

Figure 1. Source signals corresponding to two recordings of 400 samples each one. Up, the backscattering echo; down, the interference signal. The echo is the same in the two recordings because of invariance; the interference is a sinusoid with different phase in each recording. This fact assures that the mixing matrix is not singular.

1 Supported by CICYT under grant DPI2000-0619-C03-01 and Polytechnic University of Valencia

We can record several responses from the material at the same location by taking advantage of the pulse train emitted to the material.

Supposing material response is time invariant, we record the response to a train of N pulses, with period T ,

$$
\begin{aligned}
y(t) &= x(t) + \sum_{i=1}^{N-1} B_i e^{j(w_i t + q_i)} \quad 0 \le t \le T \\
y(t) &= x(t) + \sum_{i=1}^{N-1} B_i e^{j(w_i t + q_i)} \quad T \le t \le 2T \\
&\dots \\
y(t) &= x(t) + \sum_{i=1}^{N-1} B_i e^{j(w_i t + q_i)} \quad (N-1)T \le t \le NT
\end{aligned}
\tag{2}
$$

This set of equations is expressed in matrix notation as:

$$
y_k(t) = x(t) + \sum_{i=1}^{N-1} B_i e^{j w_i (k-1)T} e^{j(w_i t + q_i)} \tag{3}
$$

for $k = 1 \dots N,\quad 0 \le t \le T$, where $y_k(t)$ is the response to the k-th pulse emitted: $y_k(t) = x(t) + \sum_{i=1}^{N-1} B_i e^{j(w_i t + q_i)}$, $(k-1)T \le t \le kT$. As we can see, the interference effect in different pulse responses is a phase change in the harmonics, so we can obtain at least N pulse responses to model the BSS problem, which can be written as (we will drop the time variable)

$$
\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & e^{j w_1 T} & \dots & e^{j w_{N-1} T} \\ & & \ddots & \\ 1 & e^{j w_1 (N-1)T} & \dots & e^{j w_{N-1}(N-1)T} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} \tag{4}
$$

where x_1 is the backscattering echo of the material that we want to characterize, and x_{j+1} $j = 1 \dots N-1$ the j-th signal of frequency $w_j, x_{j+1} = B_j e^{j(w_j t + q_j)}$. We will call **A** the matrix obtained in (4). In BSS terminology this matrix is usually referred as the mixing matrix.

Finally, a BSS problem can be set. The aim is to recover the source signals x_j $j = 1 \dots N$ provided they can be considered statistically independent, starting from the linear mixtures of them y_j $j = 1 \dots N$. Normally, in BSS nothing is supposed about the mixing matrix **A**, only that it is not singular. It is clear from (4) that this condition is satisfied if $w_i T \ne 2pk$ with k integer; i.e. the pulse train and the interferences periods must not be multiple numbers.

3. BSS-MLE SOLUTION

For the sake of simplicity we will suppose in the rest of the paper that we have only an interfering signal.

The problem is reduced to:

$$
\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = [\mathbf{A}] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & e^{j w_1 T} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{5}
$$

Traditional BSS solutions do not suppose anything about the mixing matrix **A** or about the source signals distribution. It has an important advantage: if our model is not correct (for example x_2 is not a pure tone; probably it would have harmonics or any kind of noise) it works well. Obviously, the problem is that we are not using all the information that we have.

The difference among the BSS algorithms is the way they approximate the independence assumption and the way they solve numerically the problem, i.e., the optimization process. What all of them have in common is the use of higher order statistics, some of them explicitly and other via non linear functions. For more details see [3], [4], [5], [6], [7], [8] and the references there indicated.

We will present in this Section the Maximum Likelihood Estimate MLE, using the structure of the matrix **A**, the prior information about one of the sources (the interference) and a Gram-Charlier approximation of the probability density function (pdf) of the backscattered echo of the material [9].

First of all, we observe from (5) that the matrix **A** is not invertible if $w_1 T = 2pk$, so we will suppose that $w_1 T \ne 2pk$. Secondly, T is known (is the time between pulses), so the only remaining unknown parameter is the frequency w_1.

If we decompose the complex exponential source in its real and imaginary part, the problem is expressed as:

$$
\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & \cos w_1 T & \sin w_1 T \\ 1 & \cos 2w_1 T & \sin 2w_1 T \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \tag{6}
$$

where x_1 is the response of the material, $x_2 = \sin(w_0 t + q)$ and $x_3 = \cos(w_0 t + q)$. In (6) we fix the order of the sources, removing the permutation indetermination of the problem (the order the sources are recovered).

The MLE is the value of w_1 that maximizes the (log)-likelihood function of the signals y_1, y_2, y_3. For N independent observations,

$$
w_1^{MLE} = \arg\max_{w_1} \sum_{k=1}^{N} p_{\mathbf{y}(k)}(\mathbf{y}(k); w_1) \tag{7}
$$

The joint pdf $p_{\mathbf{y}}(\mathbf{y})$ is obtained from the pdf of the sources and from the mixing matrix **A**.

The pdf of the observed signals with respect to the pdf of the sources is:

$$
p_{\mathbf{y}}(\mathbf{y}) = \frac{p_{\mathbf{x}}\left(\mathbf{A}^{-1} \cdot \mathbf{y}\right)}{\left|\det \mathbf{A}\right|} \tag{8}
$$

Imposing the independence assumption of the sources,
$$p_{x_1x_2x_3}(x_1,x_2,x_3) = p_{x_1}(x_1)\cdot p_{x_2}(x_2)\cdot p_{x_3}(x_3)\,, \qquad (8)$$
can be factorized in terms of the pdf of the sources and the unknown parameter w_1.

If we do not consider any prior information about the pdf of the sources, a approximation of them is necessary. They can be approximated in different ways [8]. In particular, we will use the Gram-Charlier expansion [9] for the response of the material and the histogram for the sinusoidal interference, so

$$p_{x_1}(x_1) = c\cdot e^{-x_1{}^2/2}\cdot \left(1+\frac{1}{6}(x_1^3-3x_1)m_{x_1}^3+\frac{1}{24}(x_1^4-6x_1^2+3)(m_{x_1}^4-3)\right) \qquad (9)$$

$$p_{x_i}(x_i)=\frac{1}{p\cdot\sqrt{2-x_i^2}} \qquad -\sqrt{2}\le x_i\le\sqrt{2},\, i=2{,}3 \qquad (10)$$

where m^3, m^4 are the third and fourth order and c is a normalizing constant.

Thanks to the multilinearity property of the statistics, we can obtain the moments of **x** starting from the moments of the mixed signals **y**.

$$\begin{aligned} m_{ijk} &= \sum_{pqr} A_{ip}^{-1}A_{jq}^{-1*}A_{kr}^{-1}h_{pqr} \\ m_{ijkl} &= \sum_{pqrs} A_{ip}^{-1}A_{jq}^{-1*}A_{kr}^{-1}A_{ls}^{-*}h_{pqrs} \end{aligned} \qquad (11)$$

where (m_{ijk}, m_{ijkl}), (h_{ijk}, h_{ijkl}) are the third and fourth order moments of **x** and **y** respectively.

We are only interested in the moments of equation (9), so we have to compute just the third and fourth order marginal moments for the echo source.

The resulting algorithm can be summarized:

1. Estimate the moments of the observed signals h_{ijk}, h_{ijkl}
2. Compute the moments (11) of the material-response source as a function of the frequency w_1.
3. Approximate the pdf of the sources (9), (10) with the expression obtained in the previous step and the relation $\mathbf{x}=\mathbf{A}^{-1}\mathbf{y}$.
4. Obtain the w_1 that maximizes the (log)-likelihood function (7).

3.1 General solution

In our mixing model (6), the mixing matrix has been expressed as a function of only one parameter. But, in this model, considering the assumption of unit-variance sources, the mixing matrix only expresses the particular case in which the sources are mixed with the same power. In the general case, the mixing matrix can be defined as:

$$\begin{bmatrix} \sqrt{S/N} & 1 & 0 \\ \sqrt{S/N} & \cos w_1T & \sin w_1T \\ \sqrt{S/N} & \cos 2w_1T & \sin 2w_1T \end{bmatrix} \qquad (12)$$

where the *S/N* ratio is defined as the relation $\dfrac{\text{power echo}}{\text{power sinusoid}}$. In our previous model a *S/N*=0 dB was considered.

In this case, instead of obtain the MLE of the two parameters, the S/N ratio and the frequency w_1, we will obtain directly the sources via a two steps BSS solution. It is known that the mixing matrix can be decomposed as the product of a whitening **L** and an orthogonal matrix **Q** [1]. The whitening matrix **L** can be obtained in several ways [10]. Supposed that the whitening step is carried out, for the *2x2* problem the orthogonal matrix **Q** is a Givens matrix parameterized by the rotation angle:

$$\mathbf{Q}=\begin{bmatrix} \cos a & \sin a \\ -\sin a & \cos a \end{bmatrix} \qquad (13)$$

Because imposing general independence among all three sources is equivalent to pairwise independence for the BSS problem [1], we iteratively obtain the Givens matrix for the *2x2* problem, considering only two of three sources in every sweep. When the algorithm converges, the original sources are recovered. As we mentioned before, we will impose the order of the sources, so the permutation indetermination of the BSS problem is removed. We will assume that the order of the sources is corresponding with the mixing matrix (12), i.e. the first source is the response of the material, the second the sine and the third the cosine function. In every *2x2* problem the sources will be approximated by the Gram-Charlier expansion (9) when the source is the response of the material or (10) when it is the sine or cosine functions.

4. RESULTS

A simple simulation was carried out with a computer generated interference and a real backscattered echo registered in a no interference environment. However we will not use any prior information about the material response in our algorithm.

The original sources are:

Source x_1. The normalized backscattered echo of a material.

Sources x_2 and x_3. Synthetic normalized interference generated with a computer $x_2=\sqrt{2}\sin\left(\frac{4p}{125}n\right)$, $x_3=\sqrt{2}\cos\left(\frac{4p}{125}n\right)$. This interference simulates the case of an interference of frequency 80 KHz, sampled at 5 MHz. This interference is in the same frequency band of the other source.

The period of the train of pulses is T=0.2 ms, satisfying the condition $w_1 T \neq 2k\pi$. In order to reduce the computational time we will use only N=801 samples from the original record length (samples between 2200 and 3000). In this case, the restriction is also satisfied $w_1 N = 80.52 = 12.82 \cdot 2\pi \neq 2k\pi$.

The mixing matrix is (12) with S/N=1/3. In the Figure 2, the mixtures are shown and, in the Figure 3, we can observe the recovered signals. As we can see the response of the material and the sinusoidal signals are separated.

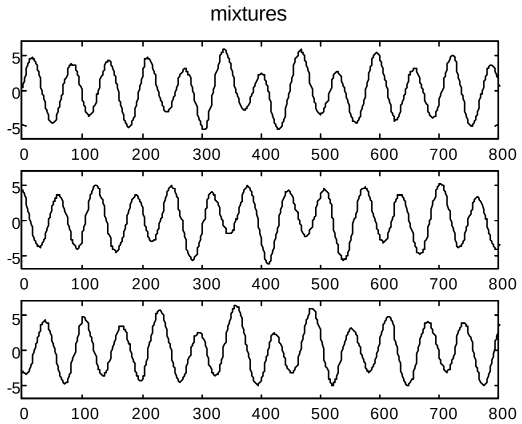

Figure 2. Mixed signals.

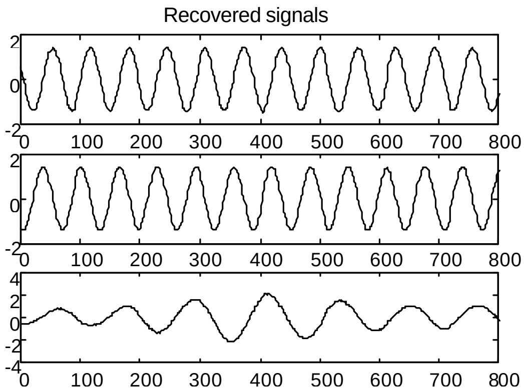

Figure 3. Recovered signals.

5. CONCLUSIONS

A new application of BSS has been presented. We have shown how sinusoidal interferences can be eliminated from the backscattering echo of a material that we want to characterize. One advantage of BSS method is that we do not need to impose any special condition on the sources nor on the mixing process. In particular, we have obtained the ML solution via a Gram-Charlier approximation of the pdf of the sources.

In order to obtain a general solution a two step approach has been carried out, considering the prior information about the interference signals.

Thanks to this prior we have shown how the problem of having more sources than sensors can be overcome, so a cheaper real implementation can be developed.

In fact, our approach can be generalized for all pulsed systems, like those in radar and sonar areas, where using the characteristics of the signals involved and their periodicity, BSS techniques can be applied easily where the general restriction of "as many mixtures as sources" is not a problem even if the number of sensors and mixtures are different, because we can obtain as many mixtures as we need with only one sensor if some conditions are satisfied (in our case, that the phase change of the sinusoidal signal between emitted pulses is not a multiple of 2π).

References

[1] P. Comon, "Independent component analysis, a new concept?", *Signal Processing*, Vol. 36, No. 3, April 1994, pp 287-314.

[2] V.M. Narayanan, P.M. Shankar, L. Vergara, J.M. Reid. "Studies on ultrasonic scattering from quasi-periodic structures", *IEEE Trans. on Ultrasonics, Ferroelectrics and Frequency Control,* Vol. 44, No. 1, January 1997.

[3] J.F. Cardoso. "Blind Signal Separation: statistical principles", *Proc.of the IEEE,* Vol. 86, No. 10, pp. 2009-2025, Oct. 1998.

[4] S. Amari, A. Cichocki, Adaptive blind signal processing, neural network approaches. *Proceedings of the IEEE*, vol. 86, No. 10, pp 2026-2048, Oct 1998.

[5] E. Oja, A. Hyvarinen, J. Karhunen, *Independent Component Analysis*, John Wiley & Sons, 2001.

[6] T-W. Lee, Independent Component Analysis. Kluwer academic publishers 1998.

[7] Advances in Independent Component Analysis. Mark. Girolami Ed. Springer. 2000.

[8] Independent Component Analysis. Principles and practice. Edited by S. Roberts, R. Everson. Cambridge University Press. 2001.

[9] A. Stuart, J.K. Ord, *Kendall's Advanced Theory of Statistics.* Vol. 1, Sixth Edition, London, Edward Arnold, 1994.

[10] C. Therrien. *Discrete random signals and statistical signal processing*, Englewood Cliffs, NJ. Prentice-Hall, 1992.