On the Performance of a HOS-Based ICA Algorithm in BSS of Acoustic Emission Signals

Carlos G. Puntonet[1], Juan-José González de-la-Rosa[2], Isidro Lloret, and Juan-Manuel Górriz

[1] University of Granada, Department of Architecture and Computers Technology, ESII, C/Periodista Daniel Saucedo. 18071, Granada, Spain
carlos@atc.ugr.es
[2] University of Cádiz, Electronics Area, Research Group TIC168 - Computational Electronics Instrumentation, EPSA, Av. Ramón Puyol S/N. 11202, Algeciras-Cádiz, Spain
juanjose.delarosa@uca.es

Abstract. A cumulant-based independent component analysis (Cum-ICA) is applied for blind source separation (BSS) in a synthetic, multi-sensor scenario, within a non-destructive pipeline test. Acoustic Emission (AE) sequences were acquired by a wide frequency range transducer (100-800 kHz) and digitalized by a 2.5 MHz, 8-bit ADC. Four common sources in AE testing are linearly mixed, involving real AE sequences, impulses and parasitic signals from human activity. A digital high-pass filter achieves a SNR up to -40 dB.

1 Introduction

AE signal processing usually deals with the problem of separation multiple events which sequentially occur in several measurement points during a non-destructive test. In most situations, the test involves the study of the behavior of secondary events, or reflections, resulting from an excitation (the main event). These echoes carry information related with the medium through which they propagate, as well as reflecting surfaces [1].

But, in almost every measurement scenario, an acquired sequence contains information regarding not only the AE under study, but also additive noise processes (mainly from the measurement equipment) and other parasitic signals, e.g. originated by human activity or machinery vibrations. As a consequence, in non-favorable SNR cases, BSS should be accomplished before characterization [2], in order to obtain the most reliable *fingerprint* of the AE event.

The main goal of this paper is to show how an ICA algorithm (based in cumulants) can separate signals from a multi-sensor array, which comprises synthetics of AE events and additive signals, widespread used in non-destructive vibration tests. The algorithm have proven success for a SNR=-40 dB situation, and uses the cross-cumulants of the measured time-series to maximize the goal function. These higher-order statistics take advantage from their noise rejection capabilities to extract sources. A high-pass filter completes the post-processing in the cases of low-frequency couplings.

J. Rosca et al. (Eds.): ICA 2006, LNCS 3889, pp. 400–405, 2006.

The paper is structured as follows: in Section 2 we make a brief progress report on AE characterization. Section 3 summarizes the ICA model and outlines its properties. Results are displayed in section 4. Finally, conclusions and achievements are drawn in section 5.

2 Analysis of Acoustic Emission Signals

AE is defined as the class of phenomena whereby transient elastic waves are generated by the rapid (and spontaneous) release of energy from localized sources within a material, or the transient elastic wave(s) so generated. Elastic energy travels through the material as a stress or strain wave and is typically detected using a piezoelectric transducer, which converts the surface displacement (vibrations) to an electrical signal.

AE signal processing is used for the detection and characterization of failures in non-destructive testing and identification of low-level biological signals [2]. Most AE signals are non-stationary and they consist of overlapping bursts with unknown amplitude and arrival time. These characteristics can be described by modelling the signal [1], by means of neural networks, and using wavelet transforms.

The above second-order techniques have been also applied in an automatic analysis context of the estimation of the time of occurrence and amplitude of the bursts. Multiresolution has provided good performance in de-noising (up to SNR=-30 dB) and estimation of time instances, due to the selectivity of the filters banks implemented in the wavelets [3].

Higher order statistics (HOS) have enhanced characterization in analyzing biological signals due to the capability for rejecting noise [4].

3 The ICA Model and Its Properties

3.1 Outline of ICA

BSS by ICA is receiving attention because of its applications in many fields such as speech recognition, medicine and telecommunications [5]. Statistical methods in BSS are based in the probability distributions and the cumulants of the mixtures. The recovered signals (the source estimators) have to satisfy a condition which is modelled by a contrast function. The underlying assumptions are the mutual independence among sources and the non-singularity of the mixing matrix [6],[7],[8].

Let $\mathbf{s}(t) = [s_1(t), s_2(t), \ldots, s_m(t)]^T$ be the transposed vector of sources (statistically independent). The mixture of the sources is modelled by

$$\mathbf{x}(t) = \mathbf{A} \cdot \mathbf{s}(t) \tag{1}$$

where $\mathbf{x}(t) = [x_1(t), x_2(t), \ldots, x_m(t)]^T$ is the available vector of observations and $\mathbf{A} = [a_{ij}] \in \Re^{m \times n}$ is the unknown mixing matrix, modelling the environment in which signals are mixed, transmitted and measured [9]. We assume that $\mathbf{A}$ is a non-singular n×n square matrix. The goal of ICA is to find a non-singular n×m separating matrix $\mathbf{B}$ such that extracts sources via

$$\hat{\mathbf{s}}(t) = \mathbf{y}(t) = \mathbf{B} \cdot \mathbf{x}(t) = \mathbf{B} \cdot \mathbf{A} \cdot \mathbf{s}(t) \tag{2}$$

where vector $\mathbf{y}(t) = [y_1(t), y_2(t), \ldots, y_m(t)]^T$ is an estimator of the sources. The separating matrix has a scaling freedom on each of its rows because the relative amplitudes of sources in $\mathbf{s}(t)$ and columns of $\mathbf{A}$ are unknown [7]. The transfer matrix $\mathbf{G} \equiv \mathbf{BA}$ relates the vector of independent original signals to its estimators.

3.2 CumICA

High order statistics, known as cumulants, are used to infer new properties about the data of non-Gaussian processes. Before cumulants, such processes had to be treated as if they were Gaussian. Cumulants and polyspectra reveal information about amplitude and phase, whereas second order statistics are phase-blind. The relationship among the cumulant of r stochastic signals and their moments of order $p, p \leq r$, can be calculated by using the *Leonov-Shiryayev* formula [10]:

$$\begin{aligned} Cum(x_1, ..., x_r) = \sum (-1)^k \cdot (k-1)! \cdot E\{\prod_{i \in v_1} x_i\} \\ \cdot E\{\prod_{j \in v_2} x_j\} \cdots E\{\prod_{k \in v_p} x_k\} \end{aligned} \tag{3}$$

where the addition operator is extended over all the set of v_i $(1 \leq i \leq p \leq r)$ and v_i compose a partition of *1,. . . ,r*.

It has been proved that a set of random variables are statistically independent if their cross-cumulants are zero. This property is used to define a contrast function, by minimizing the distance between the cumulants of the sources $\mathbf{s}$(t) and the outputs $\mathbf{y}$(t). As sources are unknown, it is necessary to involve the observed signals. Separation can be developed using the following contrast function based on the entropy of the outputs [6]:

$$H(\mathbf{z}) = H(\mathbf{s}) + log[det(\mathbf{G})] - \sum \frac{\mathbf{C}_{1+\beta, y_i}}{1+\beta} \tag{4}$$

where $\mathbf{C}_{1+\beta, y_i}$ is the $1+\beta$th-order cumulant of the ith output, $\mathbf{z}$ is a non-linear function of the outputs y_i, $\mathbf{s}$ is the source vector, $\mathbf{G}$ is the global transfer matrix of the ICA model and $\beta > 1$ is an integer verifying that $\beta + 1$-order cumulants are non-zero.

Using equation 4, the separating matrix can be obtained by means of the following recurrent equation [9]

$$\mathbf{B}^{(h+1)} = [\mathbf{I} + \mu^{(h)}(\mathbf{C}^{1,\beta}_{y,y}\mathbf{S}^{\beta}_{y} - I)]\mathbf{B}^{(h)} \tag{5}$$

where $\mathbf{S}^{\beta}_{y}$ is the matrix of the signs of the output cumulants. Equation 5 can be interpreted as a quasi-Newton algorithm of the cumulant matrix $\mathbf{C}^{1,\beta}_{y,y}$. The learning rate parameters $\mu^{(h)}$ and η are related by

$$\mu^{(h)} = \min\left(\frac{2\eta}{1+\eta\beta}, \frac{\eta}{1+\eta\|\mathbf{C}^{1,\beta}_{y,y}\|_p}\right) \quad (6)$$

with $\eta < 1$ to avoid $\mathbf{B}^{(h+1)}$ being singular; $\|.\|_p$ denotes de p-norm of a matrix. The adaptative equation 5 converges, if the matrix $\mathbf{C}^{1,\beta}_{y,y}\mathbf{S}^{\beta}_{y}$ tends to the identity.

Provided with the mathematical foundations the experimental results are outlined.

4 Experimental Results

The sensor is attached to the outer surface of the pipeline, which is under mechanical excitation. Each sequence comprises 2502 points (sampling frequency of 2.5 MHz and 8 bits of resolution), and assembles the main AE event and the subsequent reflections (echoes).

Four sources have been considered and linearly mixed. The real AE event, an uniform white noise ($SNR = -40\ dB$), a damped sine wave and an impulse-like event. The damping sine wave models a mechanical vibration which may occur, i.e. as a consequence of a maintenance action. It has a damping factor of 2000 and a frequency of 8000 Hz. Finally, the impulse is included as a very common signal registered in vibration monitoring.

The results of the algorithm are depicted in figure 1. The damping sinusoid is considered as a frequency component of the impulse-like event because IC3 and IC4 are almost the same. The final independent components are obtained filter-

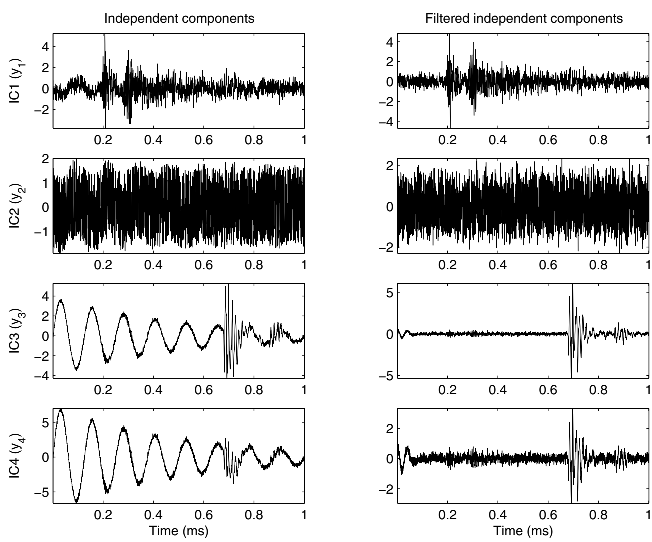

Fig. 1. Estimated sources (ICs; Independent Components) and filtered estimated sources

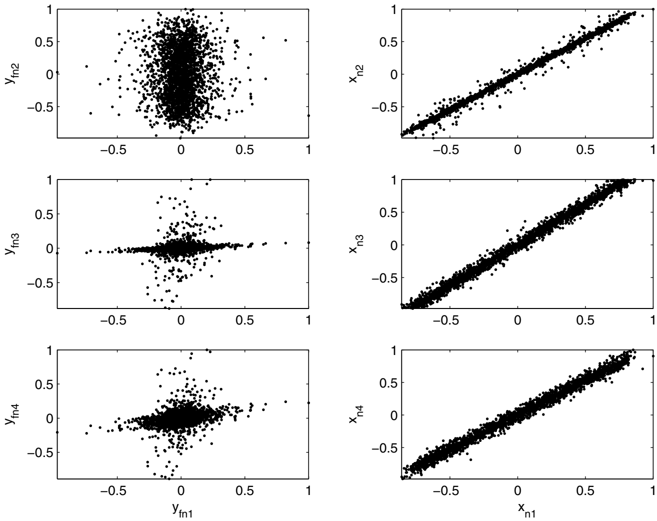

Fig. 2. Joint distributions of the mixtures and the independent components

ing the independent components by a 5th-order *Butterworth* high-pass digital filter (20000 kHz).

Finally, to test the independence of the independent components, some relevant joint distributions have been included in figure 2. The left column shows how for any IC, the values are quite random. This means that for a value (point) of an IC, almost all the values of the another IC are allowed. On the other hand, the joint distributions of the mixtures are linearly shaped, which leads us to infer a dependency before separating by ICA.

The above results lead us to some conclusions on the use of HOS as a characterizing and separating tool to be considered in a non-destructive measurement system.

5 Conclusions and Future Work

ICA is far different from traditional methods, as power spectrum, which obtain an energy diagram of the different frequency components, with the risk that low-level sounds could be masked. This experience shows that the algorithm is able to separate the sources with small energy levels in comparison to the background noise. This is explained away by statistical independence basis of ICA, regardless of the energy associated to each frequency component. The post filtering action let us work with very low SNR signals. The next step is oriented in a double direction. First, a stage involving four real mixtures will be developed. Second, and simultaneously, the computational complexity of the algorithms have to be reduced to perform an implementation.

Acknowledgement

The authors would like to thank the *Spanish Ministry of Education and Science* for funding the projects DPI2003-00878, TEC2004-06096 and PTR1995-0824-OP.

References

1. Piotrkowski, R., Gallego, A., Castro, E., García-Hernández, M., Ruzzante, J.: Ti and Cr nitride coating/steel adherence assessed by acoustic emission wavelet analysis. Non Destructive Testing and Evaluation (NDT and E) International (Ed. Elsevier) **8** (2005) 260–267
2. de la Rosa, J.J.G., Puntonet, C.G., Lloret, I.: An application of the independent component analysis to monitor acoustic emission signals generated by termite activity in wood. Measurement (Ed. Elsevier) **37** (2005) 63–76 Available online 12 October 2004.
3. de la Rosa, J.J.G., Puntonet, C.G., Lloret, I., Górriz, J.M.: Wavelets and wavelet packets applied to termite detection. Lecture Notes in Computer Science (LNCS) **3514** (2005) 900–907 Computational Science - ICCS 2005: 5th International Conference, GA Atlanta, USA, May 22-25, 2005, Proceedings, Part I.
4. Puntonet, C.G., de la Rosa, J.J.G., Lloret, I., Górriz, J.M.: Recognition of insect emissions applying the discrete wavelet transform. Lecture Notes in Computer Science (LNCS) **3686** (2005) 505–513 Third International Conference on Advances in Pattern Recognition, ICAPR 2005 Bath, UK, August 22-25, 2005, Proceedings, Part I.
5. Mansour, A., Barros, A.K., Onishi, N.: Comparison among three estimators for higher-order statistics. In: The Fifth International Conference on Neural Information Processing, Kitakyushu, Japan (1998)
6. Puntonet, C.G., Mansour, A.: Blind separation of sources using density estimation and simulated annealing. IEICE Transactions on Fundamental of Electronics Communications and Computer Sciences **E84-A** (2001)
7. Hyvärinen, A., Oja, E.: Independent Components Analysis: A Tutorial. Helsinki University of Technology, Laboratory of Computer and Information Science (1999)
8. Lee, T.W., Girolami, M., Bell, A.J.: A unifying information-theoretic framework for independent component analysis. Computers and Mathematics with Applications **39** (2000) 1–21
9. de la Rosa, J.J.G., Puntonet, C.G., Górriz, J.M., Lloret, I.: An application of ICA to identify vibratory low-level signals generated by termites. Lecture Notes in Computer Science (LNCS) **3195** (2004) 1126–1133 Proceedings of the Fifth International Conference, ICA 2004, Granada, Spain.
10. Swami, A., Mendel, J.M., Nikias, C.L.: Higher-Order Spectral Analysis Toolbox User's Guide. (2001)