A Non-Gaussianity Measure for Blind Source Separation

Cesar F. Caiafa
Laboratorio de Sistemas Complejos
Facultad de Ingeniería
Universidad de Buenos Aires, Argentina
Email: ccaiafa@fi.uba.ar

Araceli N. Proto
Laboratorio de Sistemas Complejos
Facultad de Ingeniería
Universidad de Buenos Aires, Argentina
and
Comisión de Investigaciones Científicas(CIC)
de la Prov. de Buenos Aires
Buenos Aires, Argentina
Email: aproto@fi.uba.ar

***Abstract*— Blind Source Separation (BSS) problem has been extensively studied for signals composed by statistical independent components. As it is well known, the applied methods usually fail when sources exhibit some degree of dependence. Our work points to solve the BBS problem removing the condition of statistical independence looking for sources estimates that maximize a measure of non-Gaussianity. We present a measure of non-gaussianity based on the L^2- Euclidean distance using a non-parametric technique for the estimation of probability densities. These mathematical tools have allowed us to build new algorithms for BSS which may have good performance for dependent as well as independent non-Gaussian real-world sources.**

I. Mathematical treatment of BSS

The mathematical framework is the following: assuming the existence of M non-Gaussian input signals s_0, s_1,.., s_{M-1} with zero mean ($E(s_i) = 0$) and unitary variance ($E\left(s_i^2\right) = 1$), a set of M linear mixtures x_0, x_1,.., x_{M-1} are generated instantaneously, i.e. $x_i(t) = \sum_{j=0}^{M-1} a_{ij} s_j(t)$, for which the matrix representation is: $\mathbf{x}(t) = A\mathbf{s}(t)$ where $\mathbf{s}(t) = [s_0 \ s_1 \ s_{M-1}]^T$ and $\mathbf{x}(t) = [x_0 \ x_1 \ x_{M-1}]^T$ are $M \times 1$ column vectors and A is the $M \times M$ invertible mixing matrix.

Hence, the purpose of any BSS algorithm is to obtain a separating matrix D such that if we define $\mathbf{y}(t) = D\mathbf{x}(t)$, then $\mathbf{y}(t)$ is composed by permuted and/or sign changed versions of $\mathbf{s}(t)$ entries.

If we assume to have non-Gaussian, dependent, and correlated sources ($E[\mathbf{ss}^T] = R_{\mathbf{ss}}$ not diagonal) it appears to be natural to propose the following two main strategies for searching the appropriate separation matrix: 1) Minimize Mutual Information (MinMI) of vector $\mathbf{y}$ (like most of ICA algorithms use) or 2) Maximize Non-Gaussianity (MaxNG) of $\mathbf{y}$ entries. We propose in this work to use MaxNG strategy concentrating our search for those source estimates that are more non-Gaussian in a certain sense.

II. Our measure of non-Gaussianity based on the L^2- Euclidean distance

Considering a continuous random variable y for which $E[y] = 0$ (zero mean) and $E\left[y^2\right] = 1$ (unitary variance), we define our non-Gaussianity measure of a probability density function p_y denoted by $\Gamma(p_y)$, as following: $\Gamma(p_y) = \int \left[\Phi(y) - p_y(y)\right]^2 dy$, where $\Phi(y) = N(0,1) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}y^2\right)$ is the Gauss probability density function.

Based on N samples of the random variable y, it means, time samples of the stochastic process $y(t)$: $y(0)$, $y(1)$, $y(2)$, etc., using the Parzen windows non-parametric density estimation technique with a Gaussian kernel ([1]) we obtain:

$$\Gamma(p_y) = \frac{1}{2\sqrt{\pi}} - \frac{2}{N\sqrt{h^2+1}} \sum_{i=0}^{N-1} \Phi\left(\frac{y(i)}{\sqrt{h^2+1}}\right) + \frac{1}{N^2 h\sqrt{2}} \sum_{\substack{i=0 \\ j=0}}^{\substack{N-1 \\ N-1}} \Phi\left(\frac{y(j)-y(i)}{\sqrt{2}h}\right) \quad (1)$$

where h is a parameter which affects the width and height of the Parzen windows.

III. The MaxNG algorithm

Given the vector $\mathbf{x}$ which contains the linear mixtures of sources $\mathbf{s}$ according to equation $\mathbf{x} = A\mathbf{s}$, we first we apply a whitening filter using the Karhunen-Loeve transformation: $\widetilde{\mathbf{x}} = W\mathbf{x}$. Then we derive a new criteria that we called MaxNG (Maximum Non-Gaussianity), which means to search for the appropriate separating matrix $\widetilde{D}$ such that entries of vector: $\widehat{s} = \mathbf{y} = \widetilde{D}\widetilde{\mathbf{x}}$ are as more non-Gaussian as possible. Each entry of $\mathbf{y}$ is obtained as a linear combination of whitened mixtures so we need to obtain M local maxima of the non-Gaussianity measure of a linear combination of mixtures.

Is easy to see that, in case of having independent sources (ICA), after the data is whitened, the task to be performed by ICA algorithms is to find an orthogonal matrix ($\widetilde{D}\widetilde{D}^T = I$) to achieve the separation. When sources are not independent, orthogonality is not still valid, but the structure of separating matrix is conditioned by the covariance structure of sources by the equation: $R_{\widehat{s}\widehat{s}} = E[\widehat{\mathbf{s}}\widehat{\mathbf{s}}^T] = \widetilde{D}\widetilde{D}^T$. If source covariance matrix is known, this a priori information can be used to restrict the search to a valid subset matrices $\widetilde{D}$.

IV. EXAMPLES OF REAL WORLD SIGNALS SEPARATION AND CONCLUSIONS

In this section we present two additional examples of real world signals separation. We have compared the results of our MaxNG algorithm against the results obtained through the application of some classical BSS/ICA methods like: AMUSE, EVD2, SOBI, JADE-opt, FPICA, Pearson-opt (all these methods are fully reviewed in the book [2]). For the application of these algorithms we have used the ICALAB Mathlab software package [3]. Our MaxNG algorithm was implemented in IDL 6.1 language. The following cases were analyse, where a mixing matrix A was arbitrarly selected:

Example 1: Speech signals. Two speakers say the same sentence. These signals were extracted from the ICALAB benchmark example named halo10.mat ([3]). These signals exhibit a slight level of correlation, in our case was: $\rho = E[s_0 s_1] = -0.05$. The number of used data samples was $N = 6000$.

Example 2: Satellite signals. Two pixel columns were extracted from an optical satellite image. These two columns were 2 pixels apart one from the other in the original image, therefore they are highly correlated, the coefficient correlation was: $\rho = E[s_0 s_1] = 0.81$ which is a very high value. The number of used data samples was $N = 5960$.

In order to measure the performance of separation the SIR Signal To Interference Ratio was used ([4]). In general a SIR value above 12dB is indicative of a successful signal estimate and SIR values over 20dB means an excellent estimation. In Table 1, corresponding values of SIR for sources estimates, are shown for both examples. Note that for the Example 1 (speech signals), our MaxNG algorithm gives excellent results (Mean SIR=41.24dB) which is comparable with some other classical methods. On the other hand, for the Example 2 (satellite signals), our MaxNG is the only able to perfectly recover both source signals. In Figure 1, a comparison of sources and their estimates using Pearson and MaxNG algorithms is shown (we have choose a subset of 200 samples). Note that for the Pearson case, only one signal was successfully estimated.

We have shown that, when original sources are dependent, traditional BSS/ICA algorithms may fail in separating signals and a better strategy is the MaxNG which search for sources which are more non-Gaussian. We have provided a new way to measure non-Gaussianity and we used it to build a new BSS algorithm, the MaxNG, for non-Gaussian sources (dependent or independent) like speech signals, remote sensed image signals and others.

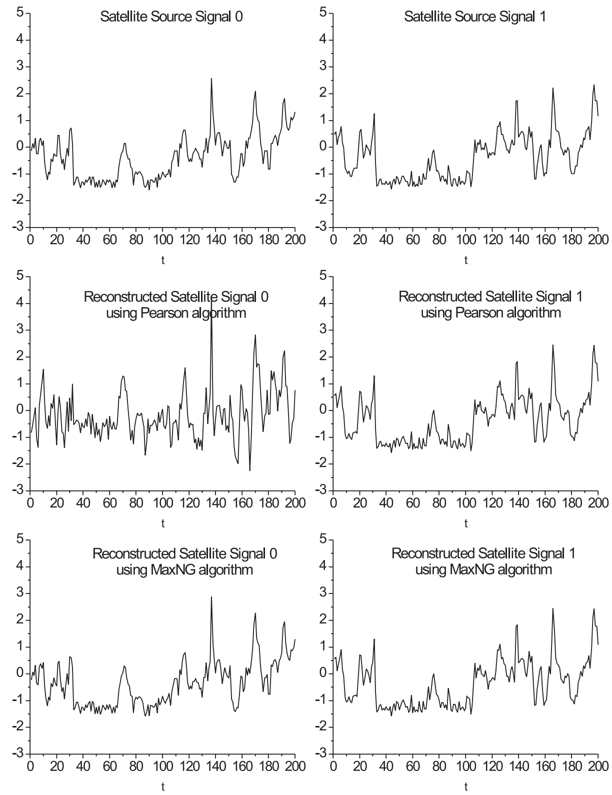

Figure 1: A comparison between original sources estimates for Pearson and MaxNG algorithm results

Example 1: speech signals			
BSS/ICA Algorithms	**SIR signal 0**	**SIR signal 1**	**MeanSIR**
AMUSE	39.59	28.36	33.97
EVD2	49.45	32.27	40.86
SOBI	63.97	31.34	**47.66**
JADE	11.41	10.57	10.99
FPICA Gauss	31.42	61.11	46.27
Pearson	25.83	22.07	23.95
MaxNG	25.20	57.27	**41.24**

Table 1: Separation performance in Example 1 (speech signals) for some classical BSS/ICA methods and MaxNG algorithm (the best result is in bold text).

Example 2: satellite signals			
BSS/ICA Algorithms	**SIR signal 0**	**SIR signal 1**	**MeanSIR**
AMUSE	9.80	10.42	10.11
EVD2	9.92	10.30	10.11
SOBI	0.11	0.11	0.11
JADE	9.83	10.39	10.11
FPICA Gauss	9.57	10.68	10.12
Pearson	2.95	19.92	**11.43**
MaxNG	20.29	20.40	**20.34**

Table 2: Separation performance in Example 2 (satellite signals) for some classical BSS/ICA methods and MaxNG algorithm (the best result is in bold text).

V. Acknowledgements

C. Caiafa acknowledges financial support from Facultad de Ingenieria, Universidad de Buenos Aires, Argentina (Beca Peruilh) and thanks the hospitality from Institute of Information Science and Technologies ISTI, Consiglio Nazionale di Ricerca (CNR), Pisa, Italy. This research is supported by Grant BID-ANPCyT PICT 02-13533 .

References

[1] R. O. Duda, P. E. Hart and D. G. Stork, Pattern Classification, 2nd Ed., J. Wiley, 2000, Chapter 4, pp. 164-174.

[2] A. Cichocki and S. Amari. Adaptive Blind Signal and Image Processing: Learning Algorithms and Applications. J. Wiley & Sons, Chichester UK, 2002.

[3] A. Cichocki, S. Amari, K. Siwek, T. Tanaka et al., ICALAB Toolboxes, http://www.bsp.brain.riken.jp/ICALAB.

[4] R. Boscolo, H. Pan and V. P. Roychowdhury, Independent Component Analysis Based on Nonparametric Density Estimation, IEEE Trans. on Neural Networks, 15 (1) (2004) 55-65