Plugging an Histogram-Based Contrast Function on a Genetic Algorithm for Solving PostNonLinear-BSS

Fernando Rojas Ruiz[1], Carlos G. Puntonet[1], Ignacio Rojas Ruiz[1],
Manuel Rodríguez-Álvarez[1], and Juan Manuel Górriz[2]

[1] Dpto. Arquitectura. y Tecnología de Computadores, University of Granada, Spain
{frojas,carlos,irojas,mrodriguez}@atc.ugr.es
[2] Area de Electrónica, University of Cádiz, Spain
juanmanuel.gorriz@uca.es

Abstract. This paper proposes a novel Independent Component Analysis algorithm based on the use of a genetic algorithm intended for its application to the problem of blind source separation on post-nonlinear mixtures. We present a simple though effective contrast function which evaluates individuals of each population (candidate solutions) based on estimating the probability densities of the outputs through histogram approximation. Although more sophisticate methods for probability density function approximation exist, such as kernel-based methods or k-nearest-neighbor estimation, the histogram presents the advantage of its simplicity and easy calculation if an appropriate number of samples is available.

1 Introduction

The guiding principle for ICA is statistical independence, meaning that the value of any of the components gives no information on the values of the other components. This method differs from other statistical approaches such as principal component analysis (PCA) and factor analysis precisely in the fact that is not a correlation-based transformation, but also reduces higher-order statistical dependencies. The extensive use of ICA as the statistical technique for solving blind source separation (BSS), may have lead in some situations to the erroneous utilization of both concepts as equivalent. In any case, ICA is just the *technique* which in certain situations can be sufficient to solve a given *problem*, that of blind source separation. In fact, statistical independence insures separation of sources in linear mixtures, up to the known indeterminacies of scale and permutation. However, generalizing to the situation in which mixtures are the result of an unknown transformation (linear or not) of the sources, independence alone is not a sufficient condition in order to accomplish blind source separation successfully. Indeed, in [5] it is formally demonstrated how for nonlinear mixtures, an infinity of mutually independent solutions can be found that have nothing to do with the unknown sources. Thus, in order to successfully separate The observed signals into a wave-preserving estimation of the sources, we need additional information about either the sources or the mixing process.

This paper is structured as follows: Section 2 introduces the post-nonlinear model as an alternative to the unconstrained pure nonlinear model. Afterwards, in Section 3, the basis of the genetic algorithm is described: independence measure, probability density function estimation and evolutionary method depiction. Some experiments are

C.G. Puntonet and A. Prieto (Eds.): ICA 2004, LNCS 3195, pp. 758–765, 2004.

shown in Section 4, using speech and synthetic signals. Finally, a few conclusion remarks and future lines of research terminate this paper.

2 Nonlinear Independent Component Analysis

2.1 Post-non-linear Model

The linear assumption is an approximation of nonlinear phenomena in many real world situations. Thus, the linear assumption may lead to incorrect solutions. Hence, researchers in BSS have started addressing the nonlinear mixing models, however a fundamental difficulty in nonlinear ICA is that it is highly nonunique without some extra constraints, therefore finding independent components does not lead us necessarily to the original sources [5].

Blind source separation in the nonlinear case is, in general, impossible. Taleb and Jutten [11] added some extra constraints to the nonlinear mixture so that the nonlinearities are independently applied in each channel after a linear mixture (see Fig.1). In this way, the indeterminacies are the same as for the basic linear instantaneous mixing model: invertible scaling and permutation.

The mixture model can be described by the following equation:

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

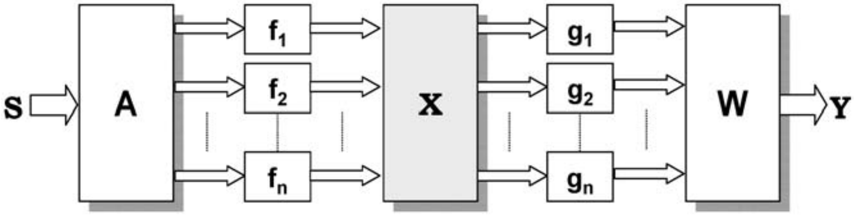

Fig. 1. Post-nonlinear model.

The unmixing stage, which will be performed by the algorithm here proposed is expressed by Equation (2):

$$\mathbf{y}(t) = \mathbf{W}\cdot\mathbf{G}\left(\mathbf{x}(t)\right) \tag{2}$$

The post-nonlinearity assumption is reasonable in many signal processing applications where the nonlinearities are introduced by sensors and preamplifiers, as usually happens in speech processing. In this case, the nonlinearity is assumed to be introduced by the signal acquisition system.

3 Genetic Algorithm for Source Separation

3.1 Mutual Information Approximation

The proposed algorithm will be based on the estimation of mutual information, value which cancels out when the signals involved are independent. Mutual information I between the elements of a multidimensional variable $\boldsymbol{y}$ is defined as:

$$I(y_1, y_2, \dots, y_n) = \sum_{i=1}^{n} H(y_i) - H(y_1, y_2, \dots, y_n)\,. \tag{3}$$

where $H(x)$ is the entropy measure of the random variable or variable set x.

For Eq. 5, in the case that all components y_1 ..... y_n are independent, the joint entropy is equal to the sum of the marginal entropies. Therefore, mutual information will be zero. In the rest of the cases (not independent components), the sum of marginal entropies will be higher than the joint entropy, leading thus to a positive value of mutual information.

In order to exactly compute mutual information, we need also to calculate entropies, which likewise require knowing the analytical expression of the probability density function (PDF) which is generally not available in practical applications of speech processing. Thus, we propose to approximate densities through the discretization of the estimated signals building histograms and then calculate their joint and marginal entropies. In this way, we define a number of bins m that covers the selected estimation space and then we calculate how many points of the signal fall in each of the bins ($B_i \; i = 1, \dots, m$). Finally, we easily approximate marginal entropies using the following formula:

$$H(y) = -\sum_{i=1}^{n} p(y_i)\, \log_2 p(y_i) \approx \; -\sum_{j=1}^{m} \frac{Card(B_j(y))}{n}\, \log_2 \frac{Card(B_j(y))}{n}. \tag{4}$$

where $Card(B)$ denotes cardinality of set B, n is the number of points of estimation y, and B_j is the set of points which fall in the j^{th} bin.

The same method can be applied for computing the joint entropies of all the estimated signals:

$$\begin{aligned} H(y_1, \dots, y_p) &= \sum_{i=1}^{p} H(y_i \,|\, y_{i-1}, \dots, y_1) \\ &\approx -\sum_{i_1=1}^{m} \sum_{i_2=1}^{m} \dots \sum_{i_n=1}^{m} \frac{Card(B_{i_1 i_2 \dots i_p}(y))}{n}\, \log_2 \frac{Card(B_{i_1 i_2 \dots i_p}(y))}{n}. \end{aligned} \tag{5}$$

where p is the number of components which need to be approximated and m is the number of bins in each dimension.

Therefore, substituting entropies in Eq.5 by approximations of Eqs.6 and 7, we obtain an approximation of mutual information (Eq. 8) which will reach its minimum value when the estimations are independent:

$$\begin{aligned} Est(I(\mathbf{y})) &= \sum_{i=1}^{p} Est(H(y_i)) - Est(H(\mathbf{y})) = \\ &= -\sum_{i=1}^{p} \left[\sum_{j=1}^{m} \frac{Card(B_j(y_i))}{n}\, \log_2 \frac{Card(B_j(y_i))}{n} \right] + \dots \\ &\dots + \sum_{i_1=1}^{m} \sum_{i_2=1}^{m} \dots \sum_{i_n=1}^{m} \frac{Card(B_{i_1 i_2 \dots i_p}(y))}{n}\, \log_2 \frac{Card(B_{i_1 i_2 \dots i_p}(y))}{n}. \end{aligned} \tag{6}$$

where $Est(X)$ stands for “estimation of x”.

Next section describes an evolution based algorithm that minimizes the contrast function defined in Eq. 8, escaping from local minima.

3.2 Proposed Genetic Algorithm

A genetic algorithm (GA) evaluates a population of possible solutions and generates a new one iteratively, with each successive population referred to as a generation. Given the current generation at iteration t, G(t), the GA generates a new generation, G(t+1), based on the previous generation, applying a set of genetic operations. Aside from other aspects regarding genetic algorithms, the key features that characterize a genetic algorithm are the encoding scheme and the evaluation or fitness function.

First of all, it should be recalled that the proposed algorithm needs to estimate two different mixtures (see Eq. 4): a family of nonlinearities g which approximates the inverse of the nonlinear mixtures f and a linear unmixing matrix W which approximates the inverse of the linear mixture A [9,10]. This linear demixing stage will be performed by the well-known FastICA algorithm by Hyvärinen and Oja [4]. To be precise, FastICA will be embedded into the genetic algorithm in order to approximate the linear mixture.

Therefore, the encoding scheme for the chromosome in the post-nonlinear mixture will be the coefficients of the odd polynomials which approximate the family of nonlinearities g . Fig. 2 shows an example of polynomial approximation and encoding of the inverse non-linearities.

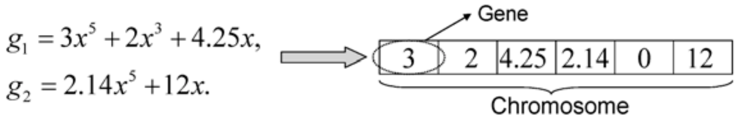

Fig. 2. Encoding example for p=2 signals and polynomials up to grade 5.

The fitness function is easily derived from Eq. 8 which is precisely the inverse of the approximation of mutual information, so that the genetic algorithm maximizes the fitness function, which is more usual in evolution programs literature.

$$Fitness(\mathbf{y}) = \frac{1}{Est(I(\mathbf{y}))}. \tag{7}$$

Expression (7) obeys to the desired properties of a contrast function [2], that is, a mapping ψ from the set of probability densities $\left\{p_x, x \in \mathbb{R}^N\right\}$ *to* $\mathbb{R}$ satisfying the following requirements:

i. $\psi(p_x)$ does not change if the components of x_i are permuted.
ii. $\psi(p_x)$ is invariant to invertible scaling.
iii. If x has independent components, then $\psi(p_{Ax}) \leq \psi(p_x), \quad \forall A$ invertible.

Regarding other aspects of the genetic algorithm, the population (i.e. set of chromosomes) was initialized randomly within a known interval of search for the polynomial coefficients. The genetic operators involved were "Simple One-point Crossover" and "Non-Uniform Mutation" [8]. Selection strategy is elitist, keeping the best individual of a generation for the next one.

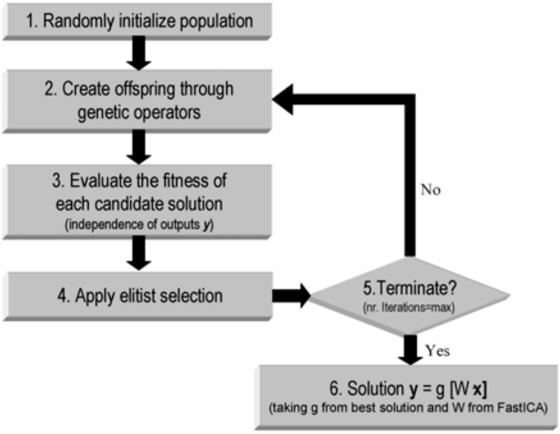

Fig. 3. Genetic algorithm scheme for post-nonlinear blind separation of sources.

4 Simulations

This section illustrates the validity of the genetic algorithm here proposed and investigates the accuracy of the method. We combined voice signals and noise nonlinearly and then try to recover the original sources. In order to measure the accuracy of the algorithm, we evaluate it using the Mean Square Error (MSE) and the Crosstalk in decibels (Ct):

$$MSE_i = \frac{\sum_{t=1}^{N}(s_i(t)-y_i(t))^2}{N} \qquad \mathrm{Ct_i} = 10\log\left(\frac{\sum_{t=1}^{N}(s_i(t)-y_i(t))^2}{\sum_{t=1}^{N}(s_i(t))^2}\right) \tag{8}$$

4.1 Two Voice Signals

This experiment corresponds to a “Cocktail Problem” [] situation, that is, separating one voice from another. Two voice signals corresponding to two persons saying the numbers from one to ten in English and Spanish were non-linearly according to the following matrix and functions:

$$\mathbf{A} = \begin{bmatrix} 1 & 0.87 \\ -0.9 & 0.14 \end{bmatrix}, \mathbf{F} = \left[f_1(x) = f_2(x) = \tanh(x) \right]. \tag{9}$$

Then the genetic algorithm was applied (population size=40, number of iterations=60). Polynomials of fifth order were used as the approximators for $g=f^{-1}$. Performance results and a plot of the original and estimated signals are briefly depicted below (Figure 5.b, right).

$\mathrm{MSE}(y_1, s_1) = 0.0012$ Crosstalk(y_1, s_1) (dB) = -17.32 dB,
$\mathrm{MSE}(y_2, s_2) = 0.0009$ Crosstalk(y_2, s_2) (dB) = -19.33 dB.

As can be seen, estimations (**y**) are approximately equivalent to the original sources (**s**) up to invertible scalings and permutations. E.g. estimation y_1 is scaled and inverted in relation to s_1.

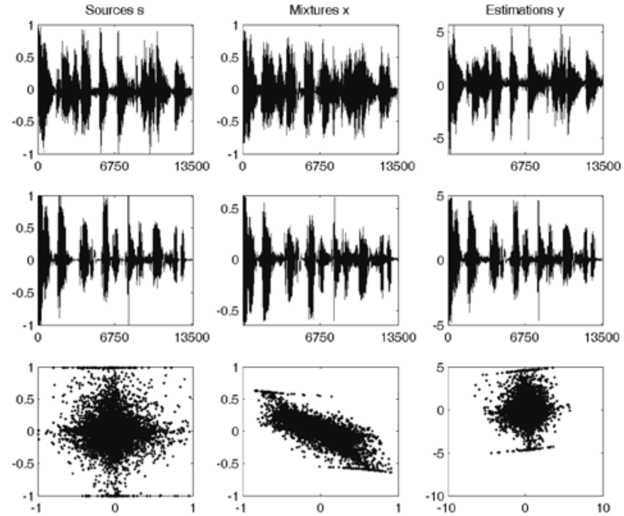

Fig. 4. Sources, mixtures and estimations (along time and scatter plots in the bottom line) for two voice signals.

4.2 Three Image Signals

In this experiment, source signals correspond to three image signals ("Lena", "Cameraman", "University of Granada emblem"). They were mixed according to the PNL scheme (Fig. 1) with the following values:

$$A = \begin{bmatrix} 0.9 & 0.3 & 0.6 \\ 0.6 & -0.7 & 0.1 \\ -0.2 & 0.9 & 0.7 \end{bmatrix}, f = \begin{bmatrix} f_1(x) = \tanh(x), \\ f_2(x) = \tanh(0.8 \cdot x), \\ f_3(x) = \tanh(0.5 \cdot x) \end{bmatrix}. \tag{10}$$

In this case, the simulation results draw a slightly worse performance than the former case, due to the increase of the dimensionality from to two to three sources:

$\mathrm{MSE}(y_1, s_2) = 0.0006$ Crosstalk(y_1, s_2) (dB) = -16.32 dB,
$\mathrm{MSE}(y_2, s_1) = 0.0010$ Crosstalk(y_2, s_1) (dB) = -12.12 dB.
$\mathrm{MSE}(y_3, s_3) = 0.0011$ Crosstalk(y_3, s_3) (dB) = -11.58 dB.

Original images can be clearly distinguished through the estimations, although some remains of the other images interfere. Also note that an inversion in the signal results obviously in the negative estimation of the source (e.g. the cameraman).

5 Concluding Remarks

In this work, an specific case of nonlinear source separation problem has been tackled by an ICA algorithm based on the use of genetic algorithms. As the separation of sources through the independence basis only is impossible in nonlinear mixtures, we assumed a linear mixture followed by a nonlinear distortion in each channel (Post-Non-Linear model) which constraints the solution space. Experimental results showed

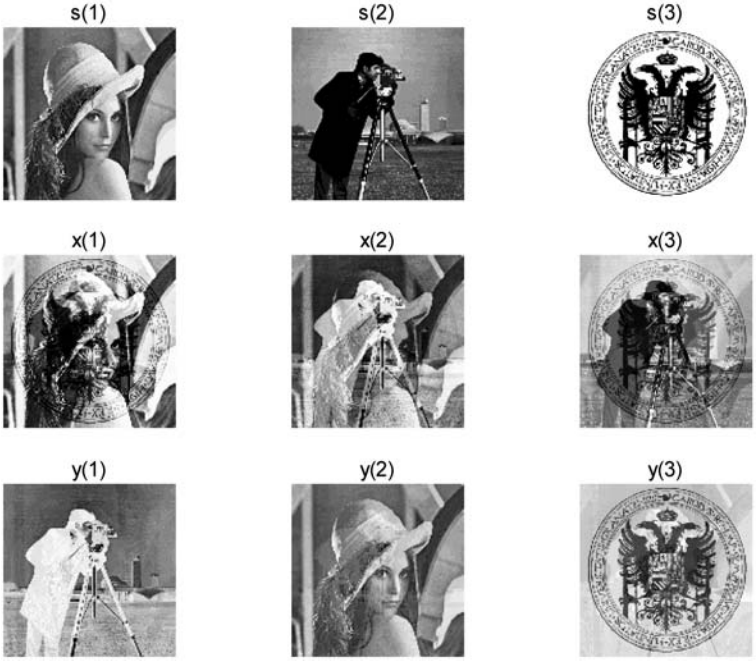

Fig. 5. Original images (s), PNL-mixtures (x) and estimations (y) applying the genetic algorithm.

promising results, although future research will focus on the adaptation of the algorithm for higher dimensionality and stronger nonlinearities.

Acknowledgement

This work has been supported by the CICYT Spanish Project TIC2001-2845.

References

1. G.Burel, Blind separation of sources: A nonlinear neural algorithm, Neural Networks, vol.5, pp.937-947, 1992.
2. P. Comon, Independent component analysis, a new concept?, Signal Processing, vol. 36, no. 3, pp. 287--314, 1994.
3. D.E. Goldberg, Genetic Algorithms in Search, Optimization and Machine Learning,AddisonWesley, Reading, MA, 1989.
4. A.Hyvärinen and E.Oja, A fast fixed-point algorithm for independent component analysis. Neural Computation, 9 (7), pp.1483-1492, 1997.
5. A. Hyvärinen and P. Pajunen. Nonlinear Independent Component Analysis: Existence and Uniqueness results. Neural Networks 12(3): pp. 429-439, 1999.
6. C. Jutten, J. Karhunen, Advances in Nonlinear Source Separation, Proceedings of the 3rd International Conference On Independent Component Analysis and Signal Separation (ICA2003). pp. 245-256, April 1-4, Nara (Japan), 2003.

7. T-W.Lee, B.Koehler, R.Orglmeister, Blind separation of nonlinear mixing models, In IEEE NNSP, pp.406-415, Florida, USA, 1997.
8. Z. Michalewicz, Genetic Algorithms + Data Structures = Evolution Programs, Springer-Verlag, New York USA, Third Edition, 1999.
9. F. Rojas, I. Rojas, R.M. Clemente, C.G. Puntonet. Nonlinear Blind Source Separation using Genetic Algorithms, in Proceedings of the 3rd International Conference On Independent Component Analysis and Signal Separation (ICA2001). pp. 400-405, December 9-13, San Diego, CA, (USA) , 2001.
10. F. Rojas, C.G., Puntonet, M. Rodríguez-Álvarez, I. Rojas, Evolutionary Algorithm Using Mutual Information for Independent Component Analysis. Lecture Notes in Computer Science, LNCS, ISSN 0302-9743, Vol. 2687, pp. 233-240, 2003.
11. A.Taleb, C.Jutten, Source Separation in Post-Nonlinear Mixtures, IEEE Transactions on Signal Processing, vol.47 no.10, pp.2807-2820, 1999.
12. A. Ziehe, M. Kawanabe, S. Harmeling, K.R. Müller, Blind Separation of Post-nonlinear Mixtures using Linearizing Transformations and Temporal Decorrelation, J. of Machine Learning Research Special Issue on Independent Components Analysis, pp.1319-1338, 2003.