ICA METHOD FOR SPECKLE SIGNALS

D. Blanco, B. Mulgrew, S. McLaughlin

IDCOM, School of Engineering and Electronics,The University of Edinburgh
The King's Buildings, Mayfield Road, EH9 3JL, Edinburgh, UK
(David.Blanco Bernie.Mulgrew Steve.McLaughlin)@ee.ed.ac.uk

ABSTRACT

Independent component analysis (ICA) has shown success in the separation of sources in lots of applications. Almost all of them assume that a set of recorded signals is the result of a linear mixture of independent sources. Although ICA methods were firstly designed to apply only to free-noise signals, numerous methods have extended it to deal with additive noise, using only higher order statistic. However, in speckle environments signals the noise is multiplicative, so the applicability of ICA is seriously reduced. This paper proposes an ICA method for speckle signals, taking into account the multiplicative nature of the noise and improving the results obtained by standard ICA methods.

1. INTRODUCTION

In the last years, blind source separation (BSS) by ICA has been applied to signal processing problems as diverse as speech enhancement, medical signal processing, image analysis, telecommunications, financial series, etc. The goal of ICA is to find a linear transformation of a set of signals such that the transformed data are as statistically independent as possible. When the recorded signals $\mathbf{x}$ are a linear mixture of independent sources $\mathbf{s}$, i.e., $\mathbf{x} = \mathbf{A}\mathbf{s}$ (which will be called *ICA model*), ICA produces the inverse of the mixing matrix $\mathbf{A}$ and the separation of the sources (column vectors are noted with lower bold letters, while matrices are noted with capital bold letters). In order to attain the separation, ICA uses second- and higher-order statistics in different ways [1][2] and [3], generalizing the second-order technique of principal component analysis (PCA).

Speckle noise is a multiplicative noise that can appear associated to different type of signals. One of the most important are the coherent images, as sonar, laser, ultrasound-B and synthetic aperture radar (SAR) images. It is in this last kind of images where ICA has been most applied in the last years. In [4] [5] and [6], different ICA approaches developed in image processing and remote sensing have been extended to SAR images, but all of them do the separation supposing that the N-dimensional data $\mathbf{x}$ follow the ICA model, possibly with additive noise. However, in speckle signals, the appearance of speckle noise can be seen as a multiplicative noise [7], so each component z_i of an observed N-dimensional data $\mathbf{z}$ can be expressed as $z_i = v_i x_i$, where the vector $\mathbf{x}$ are the data without noise and the vector $\mathbf{v}$ are the speckle noises present in each component of the data vector. Although the data without noise $\mathbf{x}$ can follow an ICA model, the recorded data $\mathbf{z}$ will not due to the presence of the speckle noise. In this paper we propose a method to extract the mixing matrix from signals that follow the ICA model contaminated with speckle noise. In the section 2, the model of a speckle mixture of independent sources is presented and its second- and third-order statistic is studied. In the section 3, the structure of these statistics is used to develop a new method that allows to obtain the mixing matrix in speckle signals. The results of the proposed method are compared with them of standard ICA methods in the section 4. The paper finishes with the conclusions in the section 5.

2. SPECKLE ICA MODEL

It is assumed that the signals follow the *speckle ICA model*, where the underlying signal follow an ICA model and the recorded signal is contaminated with speckle noise. This can be expressed as:

$$z_i = v_i x_i,\ i = 1, \ldots, N \ \text{ with } \ \mathbf{x} = \mathbf{A}\mathbf{s} \tag{1}$$

where the signals are real, $\mathbf{s}$ is the vector of independent sources, the speckle noises $\mathbf{v} = [v_i, \ldots, v_N]$ are random variables with one mean and mutually independent each other and with the signals $\mathbf{x}$ [7], $\mathbf{A}$ is the $N \times N$ mixing matrix (same number of sources and signals is assumed for simplicity) and N is the number of signals.

The covariance between two signals z_i and z_j of $\mathbf{z}$ is defined as:

$$\sigma_{ij}^z = \mathcal{E}\{(z_i - \mu_i^z)(z_j - \mu_j^z)\} \tag{2}$$

where μ_i^z is the mean of the signal z_i and $\mathcal{E}\{\cdot\}$ is the expectation operator. This covariance is easily computed:

$$\sigma_{ij}^z = \sigma_{ij}^x + \sigma_i^v \delta_{ij} \left(\sigma_{ij}^x + \mu_i^x \mu_j^x\right) \tag{3}$$

so σ_{ij}^x is the covariance between x_i and x_j, σ_i^v is the variance of v_i and μ_i^x is the mean of x_i. As the signals $\mathbf{v}$ have one mean, it means $\mu_i^z = \mu_i^x$.

The third-order cumulants of the signals z_i, z_j and z_k can be computed as:

$$\kappa_{ijk}^z = \kappa_{ijk}^x + \lfloor \sigma_j^v \delta_{jk} \left(\kappa_{ijj}^x + 2\mu_j^x \sigma_{ij}^x \right) \rfloor + \kappa_i^v \delta_{ijk} \left(\kappa_{iii}^x + 3\mu_i^x \sigma_{ii}^x + (\mu_i^x)^3 \right) \quad (4)$$

where κ_{ijk}^x and κ_i^v are the third order cumulant of x_i, x_j and x_k and the skewness of v_i, respectively, and $\lfloor f_{ijk} \rfloor = f_{ijk} + f_{jki} + f_{kij}$. These statistical functions will be used in the next section for developing a method to extract the mixing matrix from a speckle ICA modelled signal.

3. SPECKLE ICA METHOD (SICA)

The aim of ICA is to find a linear transformation $\mathbf{W}$ of the data $\mathbf{x}$, so the outputs $\mathbf{u} = \mathbf{W}\mathbf{x}$ are independents. If the data follows an ICA model, the matrix obtained after applying ICA, $\mathbf{W}$, is the inverse of the mixing matrix, and the outputs $\mathbf{u}$ are the originals sources $\mathbf{s}$. But in the case of speckle ICA model, if the inverse of the mixing matrix is applied over the speckle data $\mathbf{z}$, the output signals are not independent, so the matrix obtained by ICA is not the inverse of the mixing matrix and the outputs are not the original sources. However, although it is not possible obtain the original sources by a linear transformation of the speckle data $\mathbf{z}$, it can be seen that the outputs $\mathbf{y} = \mathbf{B}\mathbf{z}$, where $\mathbf{B} = \mathbf{A}^{-1}$, are the original sources plus an additive zero mean noise, which depends on the level of speckle noise, the mixing matrix and its inverse and the original sources. Then, it is interesting obtain the mixing matrix, not because of the information provided by it, which is what is the aim in lots of ICA applications, but because it allows us to obtain the original sources plus a zero mean noise. In order to obtain the mixing matrix, the *speckle ICA* method (SICA) will use some particular structure that the signals $\mathbf{y}$ have in their second- and third-order statistic.

If the matrix $\mathbf{B}$, which is called the *unmixing matrix*, is the inverse of the mixing matrix, its outputs are $\mathbf{y} = \mathbf{B}\mathbf{z}$ and the covariance between y_i and y_j is:

$$\sigma_{ij}^y = \sigma_i^s \delta_{ij} + \sum_r B_{ir} B_{jr} \lambda_r \quad (5)$$

with $\lambda_i = \sigma_i^v \left(\sigma_{ii}^x + (\mu_i^x)^2 \right)$, σ_i^s the variance of the signal s_i and the sum goes from 1 to N, as all the sums will do in the sequel. Since it has been modelled $\mathbf{x} = \mathbf{A}\mathbf{s}$, the arbitrary scaling factor associated with ICA problems appears. To avoid this indetermination that affects to the estimation of $\mathbf{B}$, in this paper the normalization is to take $\mathbf{s}$ such that the variances are $\sigma_i^s = 1$, for all $i = 1, \ldots, N$. With this selection, the first term in (5) is simply Dirac's delta. The third-order cumulant of y_i, y_j and y_k is:

$$\kappa_{ijk}^y = \kappa_i^s \delta_{ijk} + \sum_r B_{ir} B_{jr} B_{kr} \beta_r + \left\lfloor \sum_r B_{jr} B_{kr} \alpha_{ri} \right\rfloor \quad (6)$$

with

$$\beta_i = \kappa_i^v \left(\kappa_{iii}^x + 3\mu_i^x \sigma_{ii}^x + (\mu_i^x)^3 \right)$$
$$\alpha_{ij} = \sigma_i^v \left({A_{ij}}^2 \kappa_j^s + 2A_{ij}\mu_i^x \right)$$

and κ_i^s are the skewness of the signal s_i.

Given a set of speckle signals $\mathbf{z}$, the SICA method will consists in searching the unmixing matrix such that its outputs have a structure equal to the ones shown in (5) and (6). In order to do this, the covariance and third-order cumulants of $\mathbf{y}$ are estimated from those ones of $\mathbf{z}$. More formally,

$$\begin{aligned} \hat{\sigma}_{ij}^y &= \sum_{lm} B_{il} B_{jm} \hat{\sigma}_{lm}^z \\ \hat{\kappa}_{ijk}^y &= \sum_{lmn} B_{il} B_{jm} B_{kn} \hat{\kappa}_{lmn}^z \end{aligned} \quad (7)$$

where $\hat{\sigma}_{ij}^z$ and $\hat{\kappa}_{ijk}^z$ are the covariance and the third-order cumulant estimated with the correspondent components of $\mathbf{z}$, and all the indexes in the double and triple sums go from 1 to N. A cost function J can then be constructed:

$$J = \sum_{i \geq j} \left(\sigma_{ij}^y - \hat{\sigma}_{ij}^y \right)^2 + \sum_{i \geq j \geq k} \left(\kappa_{ijk}^y - \hat{\kappa}_{ijk}^y \right)^2 (1 - \delta_{ijk}) \quad (8)$$

with the definitions in (5), (6) and (7). All the indexes in the sums go from one 1 to N with the order constraints indicated in the sums' lower limit. The function J is function of several parameters, concretely λ_i, β_j, α_{nm} and B_{rs}, with all the indexes between 1 and N, and the correct solution will be the minimum of this cost function.

In order to carry out the minimization, the steepest descendent gradient method is used, because of its simplicity and common use in ICA problems. In this method the parameters are computed iteratively, using the gradient of the cost function J. Concretely, whatever set of parameters $\vec{\theta}$ of J in the step k is computed by:

$$\vec{\theta}(k) = \vec{\theta}(k-1) - \mu \nabla_{\vec{\theta}} J \quad (9)$$

where $\nabla_{\vec{\theta}} J = [\partial J/\partial\theta_1, \ldots, \partial J/\partial\theta_N]^{\mathrm{T}}$ is the gradient of J respect to the set of parameters and μ is the *learning ratio* that takes account of the size of the steps.

The gradients of J respect to the different set of parameters are:

$$\frac{\partial J}{\partial \lambda_i} = 2\sum_{j\geq k}(\sigma^y_{jk} - \hat{\sigma}^y_{jk})B_{ji}B_{ki}$$
$$\frac{\partial J}{\partial \alpha_{ij}} = 2\sum_{l\geq m}(\kappa^y_{jlm} - \hat{\kappa}^y_{jlm})(1+\delta_{jl}+\delta_{jm} - $$
$$- 3\delta_{jlm})B_{li}B_{mi}$$
$$\frac{\partial J}{\partial \beta_i} = 2\sum_{l\geq m\geq n}(\kappa^y_{lmn} - \hat{\kappa}^y_{lmn})(1-\delta_{lmn})B_{li}B_{mi}B_{ni}$$
$$\frac{\partial J}{\partial B_{ij}} = 2\sum_{l}(1+\delta_{li})(\sigma^y_{il} - \hat{\sigma}^y_{il})(B_{lj}\lambda_j -$$
$$-\sum_m B_{lm}\sigma^z_{jm}) + 2\sum_{l\geq m}(\kappa^y_{ilm} - \hat{\kappa}^y_{ilm})(1+\delta_{il} +$$
$$+\delta_{im} - 3\delta_{ilm})(B_{lj}\alpha_{jm} + B_{mj}\alpha_{jl} +$$
$$+ B_{lj}B_{mj}\beta_j - \sum_{np} B_{ln}B_{mp}\hat{\kappa}^z_{jnp}) \quad (10)$$

With the gradients, the method is complete with the addition of initial values for the parameters. If the method is applied with random initial values it seldom reaches the correct solution. In practice, most of the methods of minimizing nonlinear functions only converge to the correct solution if the starting point is close enough of it. If it is not so, the method can diverge, converge to a spurious minimum or wander without reach any solution. So, suitable close starting points will be specified in the next section in order to permit the convergence to proceed correctly.

To sum up, the SICA method consists of the equations (9), with the gradients (10) and some appropriate starting points.

4. RESULTS

In this section, the behaviour of the SICA method is tested to see if it actually improves the result of the standard ICA methods and how it depends on the level of noise and the number of signal. The sources used have unit variance and truncated rational *pdf*. They are obtained by the exponentiation of N uniform distributed signals in the interval $[0, a_i)$, with different a_i in each one of the N sources. These sources are mixed with a full rank square matrix, where all its elements have been generated randomly in the interval $[0, 1)$. After this, independent speckle noises are multiplied with each one of mixed signals. The speckle noises are generated as uniform distributed signal with one mean, to complete the model (1). The standard deviation of the speckle noise will change in order to test the behaviour of the SICA method in different levels of noise.

As it has been pointed out before, if the SICA method starts with random values in their parameters, the method seldom reaches an acceptable solution. A way to reach the convergence is to select starting values close enough of the correct solution. Not too much a priori information is available possible about the values of the parameters β_i and α_{ij}, except they are strongly dependent on the noise statistic, and if the noise is small, they are expected to be small. So, they can be initialized to zero. On the other hand, the parameters λ_i are defined as positive and it is easily to obtain that $\lambda_i = \sigma^v_i(\sigma^z_{ii} + (\mu^z_i)^2)/(1+\sigma^v_i)$, so if the variances of the noises can be estimated, this will be the starting point. The variance of the speckle noise can be obtained in a speckle image using some uniform patch in the image. Although in this section the variances of the noises are supposed to be known, this is not a real limitation, since a value of zero as starting point for the parameters λ_i also provides good performance.

Further, there is no initial information about the parameters B_{ij} and resort has to be made to any standard ICA method in order to find a starting point. This is not an unhappy fact since this paper does not try to design an independent method to obtain the mixing matrix, rather it attemps to improve the results of ICA methods. If the signal is free noise, ICA methods obtain the correct solution, so if the noise is small it is expected that the ICA solution be close to the correct solution. Therefore, a standard ICA solution will be the starting point for the parameters B_{ij}. It will be tested if the standard ICA methods provide solutions close enough to allow SICA to converge in a noisier environment. The ICA method that is used as a starting point is the *FastICA* method [3], which has proved to be an useful and fast method in the application of ICA to different fields.

In order to characterize the behaviour of the method, a parameter d is defined to measure how close the unmixing matrix and the inverse of the mixing matrix are. In the development of the SICA method, the arbitrary scale factor has been eliminated with a condition over the variances of the sources, but this does not remove a sign indetermination. Furthermore, there is a permutation indetermination in the model used. So, if the speckle ICA model is perfectly followed by the data and the functions in (7) are perfectly estimated, the family of linear transformations which fulfil the conditions (5) and (6) are the inverse of the mixing matrix or a permuted sign-switched version of this, i.e. $\mathbf{BA} = \mathbf{PS}$ where $\mathbf{P}$ is a permutation matrix and $\mathbf{S}$ is an diagonal matrix with its diagonal elements equal to one or minus one. The parameter d is defined as the minimum distance of the transformation $\mathbf{BA}$ to the identity or whatever permuted sign-switched version of the identity [2]. The distance is measured as the Frobenius norm of the difference, so:

$$d = \min\{\sum_{ij}((\mathbf{BA} - \mathbf{PS})_{ij})^2\} \quad (11)$$

where the minimum is taken over all the possible permutations and sign's changes. In the figures 1 and 2, the values

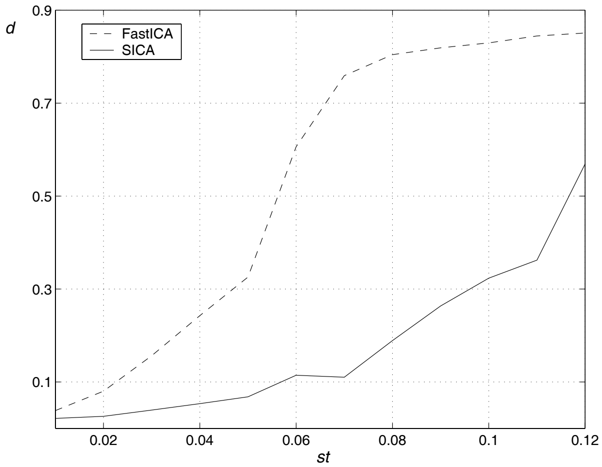

Fig. 1. Mean of *d* as funtion of the *st* for eight signals.

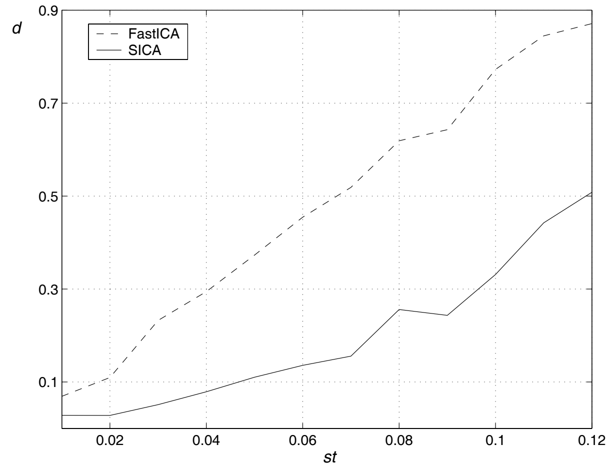

Fig. 2. Mean of *d* as funtion of the *st* for nine signals.

of the parameter d for the FastICA and SICA methods as function of the standard deviation of the speckle noise, st, are presented. The signals has been defined previously, with the number of signals equal to eight in the figure 1 and to nine in the figure 2, and the number of data in the signals are 100000. The showed parameters have computed as the mean of 100 realizations.

In the figures it can be seen how the FastICA results are clearly improved by the SICA method, for both number of signals and for a range of noise level. It also can be seen that a relatively bad value in the FastICA solution allows SICA to converge, as it happens for $st = 0.07$ and eight signals, where a big value of d for FastICA (bigger than 0.07) provides a quite small value of d for SICA (near to 0.01). Theoretically, the SICA method should obtain the correct solution for whatever level of noise. In the practice this is not so, since the bigger the speckle noise the worse the estimation of the statistical functions associated with (5) and (6). Thus, the SICA method also gets worse when the noise increases, but its results continue to be better than the FastICA's. The parameter d has also been computed for all numbers of signals between three and seven and in all the situations it has been smaller for the SICA method than for FastICA.

5. CONCLUSIONS

In this paper a new method, SICA, to obtain the mixing matrix from a linear mixture of independent sources in the presence of speckle noise has been developed. The method tries to overcome the limitations that the ICA methods have in this kind of signals. In order to do this, the SICA method does not find a non correlation and a non higher-order correlation between the outputs, as ICA does, but it finds a specific structure in the second- and third-order statistic of the output, which takes into account the multiplicative model that the speckle noise fulfil. The SICA method has been tested and its results improve clearly the ones obtain by standard ICA methods for speckle mixture of independent sources.

6. REFERENCES

[1] A. J. Bell and T. J. Sejnowski, "An information-maximization approach to blind separation and blind deconvolution," *Neural Computation*, vol. 7, pp. 1129–1159, 1995.

[2] P. Comon, "Independent component analysis—a new concept?," *Signal Processing*, vol. 36, pp. 287–314, 1994.

[3] A. Hayvärinen and E. Oja, "A fast fixed-point algorithm for independent component analysis," *Neural Computation*, vol. 9, pp. 1483–1492, 1997.

[4] X. Zhang and C. H. Chen, "A new independent component analyis (ICA) method and its application to SAR images," in *Proc. of NNSP*, 2001, pp. 283–292.

[5] S. Chitroub and B. Sansal, "Unsupervised learning rules for POLISAR images analysis," in *Proc. of NNSP*, 2002, pp. 567–576.

[6] J. Karvonen and M. Similä, "Independent component analysis for ice SAR image classification," in *Proc. of IGARSS*, 2001, pp. 1255–1257.

[7] S. Chitroub and B. Sansal, "Statistical characterisation and modelling of SAR images," *Signal Processing*, vol. 82, pp. 69–92, 2002.