Speckle Reduction of Polarimetric SAR Images Based on Neural ICA

Jian Ji[1] and Zheng Tian[1,2,3]

[1] Department of Computer Science & Technology, Northwestern Polytechnical University, Xi'an, 710072, China
jijiangao@gmail.com
[2] Department of Applied Mathematics, Northwestern Polytechnical University, Xi'an, 710072, China
[3] Key Laboratory of Education Ministry for Image Processing and Intelligent Control, Huazhong University of Science & Technology, Wuhan 430074, China

Abstract. The polarimetric synthetic aperture radar (PSAR) images are modeled by a mixture model that results from the product of two independent models, one characterizes the target response and the other characterizes the speckle phenomenon. For the scene interpretation, it is desirable to separate between the target response and the speckle. For this purpose, we proposed a new speckle reduction approach using independent component analysis (ICA) based on statistical formulation of PSAR image. In addition, we apply four ICA algorithms on real PSAR images and compare their performances. The comparison reveals characteristic differences between the studied neural ICA algorithms, complementing the results obtained earlier.

1 Introduction

Recent advances in the remote sensing polarimetric synthetic aperture radar (PSAR) systems provide a rich set of data for the same scene. This set of data brings knowledge on the nature of targets [1]. However, the PSAR images are corrupted by speckle that appears as a granular signal-dependent noise. It has the characteristics of a non-Gaussian multiplicative noise [2]. Due to its granular appearance in an image, speckle noise makes it very difficult to visually and automatically interpret SAR data. Therefore, speckle filtering is a critical preprocessing step for many SAR image processing tasks, such as segmentation and classification [3].

Independent component analysis (ICA) is an unsupervised technique that tries to represent the data in terms of statistically independent variables [4]. ICA has lately drawn a lot of attention both in unsupervised neural learning and statistical signal processing. ICA is suitable for neural network implementation and different theories recently proposed for that purpose lead to the same iterative learning algorithm. Different neural-based blind source separation algorithms are reviewed in [5-9]. The potential application of ICA in remote sensing has been validated, especially in SAR image processing. It can improve the image quality and enhance the performance of pixel classification. In short, ICA algorithm will be a useful method for remote sensing research [10].

I. King et al. (Eds.): ICONIP 2006, Part II, LNCS 4233, pp. 387 – 393, 2006.

For the same scenario, polarimetric SAR can provide a group of different polarimetric image data, and the characters of target are separated in the images polluted by speckle and are independent to the speckle noise. Thus ICA can be applied to this model and a new method is put forward to reduce speckle. In addition, it is important to know the computational properties of available algorithms in remote sensing applications. This calls for an experimental comparison of the ICA algorithms. In a companion paper [11], it had presented a first comparison of neural ICA algorithms using artificially generated data related blind source separation (BSS) problem. In this paper, we complementing the results obtained earlier by apply the four ICA algorithms to reduce speckle and compare their performance.

2 Model and Statistics of PSAR Image

Let x_i be the content of the pixel in the ith SAR image, s_i the noise-free signal response of the target, and n_i the speckle. Then, we have the following multiplicative model [2]:

$$x_i = s_i \cdot n_i \tag{1}$$

By supposing that the speckle has unity mean, standard deviation of σ_i, and is statistically independent from the observed signal x_i, the multiplicative model in (1) can be rewritten as:

$$x_i = s_i + s_i \cdot (n_i - 1) \tag{2}$$

The term $s_i \cdot (n_i - 1)$ represents the zero mean signal-dependent noise and characterizes the speckle noise variation. Thus, we have converted the multiplicative model into the additive model. The speckle filtering can be considered as the estimation of the unobservable image s_i from the noisy observation x_i.

3 ICA Formulation

The concept of ICA was first proposed by Common [4] in 1994, which has undergone a rapid development. By transforming the input signals, ICA algorithms make the mutual dependency among different signal components minimum. When the mutual dependency among signal components is measured by the different criteria, the different ICA algorithms can be derived.

Let us assume that an array of sensors provides a vector of m observed signals $\mathbf{x} = [x_1, x_2, \cdots, x_m]^T$ that are linear mixtures of $n \le m$ unobserved random processes $\mathbf{s} = [\mathrm{s}_1, \mathrm{s}_2, \cdots, \mathrm{s}_n]^T$ sources. The ICA problem is typically formulated as follows [4].

$$\mathbf{x} = \mathbf{A}\mathbf{s} + \mathbf{e} \tag{3}$$

where $\mathbf{A}$ is an unknown $m \times n$ full-column rank matrix that represents the mixing system, and $\mathbf{e} = [e_1, e_2, \cdots, e_m]^T$ is the vector of noise components which are assumed in this paper to be Gaussian and statistically independent of the sources.

In order to recover the sources, the observations are processed by a $n \times m$ separating matrix $\mathbf{B}$ to produce the vector of outputs or sources estimation

$$\mathbf{u} = \mathbf{B}\mathbf{x} \tag{4}$$

When the separation is obtained the overall mixing and separating transfer matrix $\mathbf{G} = \mathbf{BA}$ contains a single nonzero element per row and per column.

In several ICA algorithms, the data vectors $\mathbf{x}$ are preprocessed by whitening (sphering) them: $\mathbf{v} = \mathbf{Vx}$. Here $\mathbf{v}$ denotes the whitened vector satisfying $E[\mathbf{v}\mathbf{v}^T] = \mathbf{I}$, where $\mathbf{I}$ is the unit matrix, and $\mathbf{V}$ is a $m \times n$ whitening matrix. After prewhitening the subsequent $m \times m$ separating matrix $\mathbf{W}$ can be taken orthogonal, which often improves the convergence. Thus in whitening approaches the total separating matrix is $\mathbf{B} = \mathbf{WV}$.

4 Neural ICA Algorithms

In this paper we concentrate on reduction speckle for PASR image, describing the algorithms included in our study only briefly. For more details, see the references [5-9].

4.1 Natural Gradient Algorithm (NG)

Originally proposed on heuristic grounds [5], this popular and simple neural gradient algorithm was later on derived from information-theoretic criteria [6]. The update rule for the separating matrix $\mathbf{B}$ is

$$\Delta\mathbf{B} = \mu_k[\mathbf{I} - \mathbf{g}(\mathbf{u})\mathbf{u}^T]\mathbf{B} \tag{5}$$

The notation $\mathbf{g}(\mathbf{u})$ means that the nonlinearity $g(t)$ is applied to each component of the vector $\mathbf{u} = \mathbf{Bx}$. The learning parameter μ_k is usually a small constant. The basic algorithm (5) does not use prewhitening, which leads in many cases to a poor convergence. Therefore whitening is often applied to improve the convergence properties.

4.2 Equivariant (EICA) Algorithm

This algorithm is a quasi-Newton iteration that will converge to a saddle point with locally isotropic convergence, regardless of the distributions of sources. It has the following equivariant and robust in respect to Gaussian noise algorithm [7]:

$$\Delta\mathbf{B}(l) = \mathbf{B}(l+1) - \mathbf{B}(l) = \eta_l[\mathbf{I} - \mathbf{C}_{1,q}(y,y)\mathbf{S}_{q+1}(y)]\mathbf{B}(l) \tag{6}$$

where $\mathbf{S}_{q+1}(y) = \mathrm{sign}(\mathrm{diag}(\mathbf{C}_{1,q}(y,y)))$ an $\mathbf{C}_{p,q}(y,y)$ denotes the cross-cumulant matrix whose elements are $[\mathbf{C}_{p,q}(y,y)]_{ij} = Cum(\underbrace{y_i \ \cdots \ y_i}_{p}, \underbrace{y_j \ \cdots \ y_j}_{q})$.

4.3 Extended Information Maximization (Infomax) Algorithm

The purpose of extended information maximization algorithm [8] is, to provide a learning rule with a fixed nonlinearity that can separate sources with sub- and super-Gaussian p.d.f.'s. Employing a strictly symmetric bimodal univariate distribution, obtained by a weighted sum of two Gaussian distributions, given as,

$$p(\mathbf{u}) = \frac{1}{2}(N(\mu,\sigma^2) + N(-\mu,\sigma^2)) \tag{7}$$

leads to the learning rule [4] for strictly sub-Gaussian sources,

$$\Delta\mathbf{W} \propto [\mathbf{I} + \tanh(\mathbf{u})\mathbf{u}^T - \mathbf{u}\mathbf{u}^T]\mathbf{W} \tag{8}$$

For unimodal super-Gaussian sources, the following density model is adopted,

$$p(\mathbf{u}) \propto N(0,1)\mathrm{sech}^2(\mathbf{u}) \tag{9}$$

which leads to the following learning rule for strictly super-Gaussian sources,

$$\Delta\mathbf{W} \propto [\mathbf{I} - \tanh(\mathbf{u})\mathbf{u}^T - \mathbf{u}\mathbf{u}^T]\mathbf{W} \tag{10}$$

Therefore, using these two equations, we can obtain a generalized learning rule, using the switching criterion in order to distinguish between the sub- and super-Gaussian sources by the sign before the hyperbolic tangent function as,

$$\Delta\mathbf{W} \propto [\mathbf{I} - \mathbf{K}\tanh(\mathbf{u})\mathbf{u}^T - \mathbf{u}\mathbf{u}^T]\mathbf{W} \tag{11}$$

where $\mathbf{K}$ is an N-dimensional diagonal matrix composed of k_i's, defined as,

$$k_i = \mathrm{sign}(kurt(u_i)) \tag{12}$$

4.4 Fast Fixed-Point (FastICA) Algorithms

One iteration of the generalised fixed-point algorithm for finding a row vector $\mathbf{w}_i^T$ of $\mathbf{W}$ is [9]:

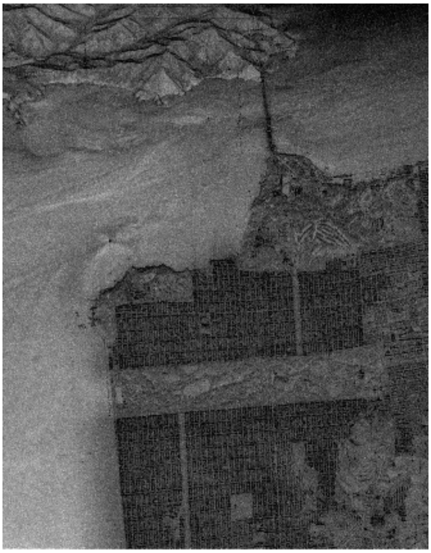
$$\mathbf{w}_i^* = \mathbf{E}\{\mathbf{v}g(\mathbf{w}_i^T\mathbf{v})\} - \mathbf{E}\{g'(\mathbf{w}_i^T\mathbf{v})\}\mathbf{w}_i \tag{13}$$

$$\mathbf{w}_i = \mathbf{w}_i^* / \| \mathbf{w}_i^* \| \tag{14}$$

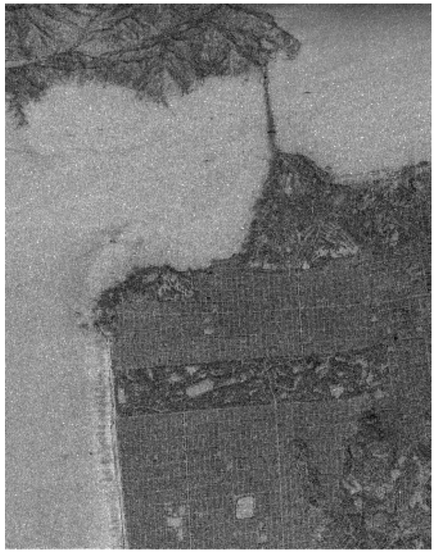
Here $g(t)$ is a suitable nonlinearity, typically $g(t) = t^3$ or $g(t) = \tanh(t)$, and $g'(t)$ is its derivative. The expectations are in practice replaced by their sample means.Hence the fixed-point algorithm is not a truly neural adaptive algorithm. The algorithm requires prewhitening of the data. The vectors $\mathbf{w}_i$ must be orthogonalised against each other; this can be done either sequentially or symmetrically. Usually the algorithm (13) converges after 5-20 iterations.

5 Simulations and Results

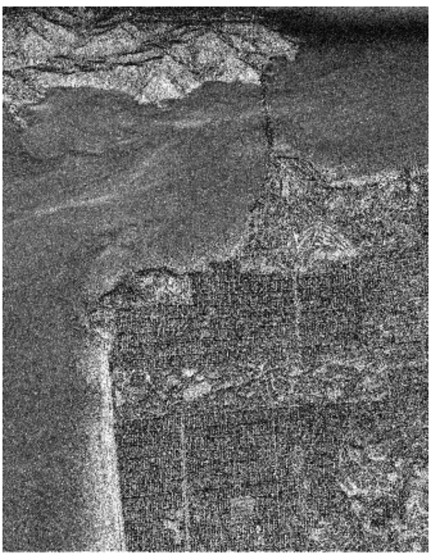
JPL AIRSAR L-band data from San Francisco are used for illustration. This San Francisco scene contains a rich variety of scatterers: specular scattering from the ocean at the top of the scene, double bounce scattering from the city blocks, volume scattering from trees, and surface scattering from grass. The original images are shown in Fig.1, including three polarimetric modes, they are HH (Horizontal- Horizontal),



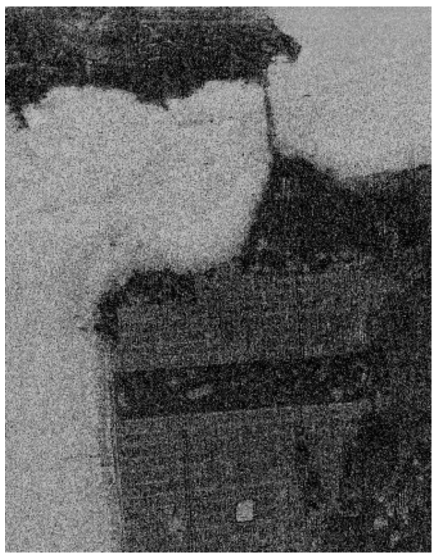
(a) HH image



(b) HV image



(c) VV image



(d) HH/VV radio image

Fig. 1. L band PSAR images

HV (Horizontal-Vertical) and VV (Vertical-Vertical). In polarimetric datas, the amplitude ratio has a number of important uses, both as a means of inferring physical propertied of a medium and as a way of removing terrain effects. In order to get better speckle reduction image, we also add HH/VV radio image to be the input data.

During the ICA application, pre-processing should be taken: Every image of 700×900 pixels should be transformed to vector, and a 4×630000 matrix was produced from the four images; The data matrix should be normalized in order to transform the pixel intensity from the nature data field to traditional data field.

In order to analyze the ability of speckle reduction by quantity, we define equivalent number of look (ENL), a good approach of estimating the speckle noise level in a SAR image, to measure the performance of speckle intensity over a uniform image region [1]. That is:

$$\mathrm{ENL} = \frac{(\mathrm{mean})^2}{\mathrm{variance}} \tag{14}$$

The ENL is equivalent to the number of independent intensity values averaged per pixel. The larger the ENL, the less the speckle effect and the stronger the ability of speckle reduction.

The output matrix **u** of ICA has become 4×630000. Comparing the four independent components (ICs), the IC 4 is complex noise. Therefore we only compute ENL of IC 1, IC 2 and IC 3. The ability of speckle reduction with different algorithms is showed in Table. 1, the ENL of original images, PCA, NG, EICA Infomax and FastICA were listed. Table 2 shows the runtime of different ICA methods. Compared with the three original images, the ENL of three PCs were increased obviously. But the ENL of PCs is lower than that of ICs. The results shown in Table 1 indicate that the FastICA and EICA algorithms performed best, with NG and Infomax having close values, while PCA is the most remote. In addition, Table 2 shows that the FastICA's speed is fastest, while Infomax has slowest speed.

Table 1. Comparison ENL of different ICA algorithms

Origin PSAR image	HH mode	4.89	EICA	IC 1	24.82
	VV mode	3.77		IC 2	17.87
	HV mode	7.79		IC 3	7.15
PCA	PC 1	16.11	Infomax	IC 1	29.41
	PC 2	7.71		IC 2	9.13
	PC 3	4.58		IC 3	6.61
NG	IC 1	20.98	FastICA	IC 1	28.76
	IC 2	10.49		IC 2	16.24
	IC 3	9.79		IC 3	6.38

Table 2. Comparison runtime of different ICA algorithms

Algorithm	NG	EICA	Infomax	FastICA
Runtime(s)	687	92	1267	23

In conclusion, the original images were improved after PCA processing, and four ICA are efficient optimizing algorithm. After ICA processing, IC 1 is the best component, the speckle index is decreased more, and the speckle is farthest separated from the original images. Comparing with other algorithms, FastICA is a fast and efficient method.

6 Conclusions

Based on rigorous statistical formulation of PSAR image, a new speckle reduction approach using ICA is proposed. In addition, we apply four ICA algorithms on real PSAR images and compare their performances. The experiment shows that ICA has effectively reduced the speckle noise of SAR image, has improved the image quality and manifested its strong ability in image separation. ICA has been widely used in blind source separation, but it is not widely used in image processing and is rarely used in remote sensing. We expect that ICA will be widely applied in remote sensing and will accelerate the development of it.

Acknowledgment. This work is supported by the National Natural Science Foundation of China (60375003) , Aeronautics and Astronautics Basal Science Foundation of China (03I53059).

References

[1] Oliver, C. and Quegan, S.: Understanding Synthetic Aperture Radar Images. Artech-House, London, 1998.
[2] Chitroub, S. Houacine, A. and Sansal, B.: Statistical characterisation and modelling of SAR images. Signal Processing, Vol. 82, No. 1, (2002) 69-92.
[3] Pi, Y. et al.: Polarimetric speckle reduction using multi-texture maximum likelihood method. IEE Electronic Letter, 39(2003) 18, 1348-1349.
[4] Common, P.: Independent component analysis, a new concept? Signal processing. 1994,36:287-314.
[5] Cichocki, A. and Unbehauen, R.: Robust neural networks with on-line learning for blind identification and blind separation of sources. IEEE Trans. on Circuits and Systems, 43(11): (1996) 894-906.
[6] Yang, H. and Amari, S. -I.: Adaptive online learning algorithms for blind separation: Maximum entropy and minimum mutual information. Neural Computation. 9(7): 1457-1482, October 1997.
[7] Cruces, S. Castedo, L. Cichocki. A.: Robust blind source separation algorithms using cumulants. Neurocomputing. vol. 49, (2002) 87-118.
[8] Lee T-W, Girolami M, Sejnowski T J.: Independent Component Analysis Using an Extended Infomax Algorithm for Mixed Subgaussian and Supergaussian Sources [J]. Neural Computation, 1999, 11 (2): 417-441.
[9] Hyvärinen, A.: Fast and Robust Fixed-Point Algorithms for Independent Component Analysis. IEEE Transactions on Neural Networks 10(3): 626-634, 1999.
[10] Fiori, S.: Overview of independent component analysis technique with an application to Synthetic Aperture Radar (SAR) imagery processing, Neural Networks, 16(2003) Special Issue, 453-467.
[11] Giannakopoulos, X. Karhunen, J. and Oja, E.: An experimental comparison of neural ICA algorithms. In Proc. Int. Conf. on Artificial Neural Networks, Skovde, Sweden, 1998.M. Girolami and C. Fyfe. Generalised independent