Overview of Independent Component Analysis Technique with an Application to Synthetic Aperture Radar (SAR) Imagery Processing

Simone Fiori

Faculty of Engineering – Perugia University Loc. Pentima bassa, 21, I-05100 Terni (Italy)

E-mail: sfr@unipg.it

Keywords:

Blind signal processing by neural networks; Independent component analysis (ICA); Remote sensing; Synthetic aperture radar (SAR).

Pages: 38, Figures: 9, References: 100.

Accepted for publication on:

Neural Networks (Special issue on 'Neural Networks for Analysis of Complex Scientific Data: Astronomy, Geology and Geophysics')

October 15, 2002

Overview of Independent Component Analysis Technique with an Application to Synthetic Aperture Radar (SAR) Imagery Processing

Simone Fiori

Abstract

We present an overview of independent component analysis, an emerging signal processing technique based on neural networks, with the aim to provide an up-to-date survey of the theoretical streams in this discipline and of the current applications in the engineering area. We also focus on a particular application, dealing with a remote sensing technique based on synthetic aperture radar imagery processing: We briefly review the features and main applications of synthetic aperture radar and show how blind signal processing by neural networks may be advantageously employed to enhance the quality of remote sensing data.

Keywords. Blind signal processing by neural networks; Independent component analysis (ICA); Remote sensing; Synthetic aperture radar (SAR).

1 Introduction

Over the last few years there has been an enormous increase in the interest of applications of neural networks to astronomy, geology, geophysics and environment protection. These research efforts have addressed a variety of problems ranging from the management of scientific equipment to the representation and interpretation of measured data.

These applications comprise the use of neural networks in astronomy, for e.g. analysis of stellar light curves and solar data (Tagliaferri, Ciaramella, Milano, Barone & Longo, 1999), the analysis of photometric images as e.g. for star/galaxy separation, deblending of unresolved images and morphological classification of galaxies (Andreon, Capuano, Gargiulo, Longo & Tagliaferri, 2000; Odewahn, Stockwell, Pennington, Humphreys, & Zumach, 1992), the analysis of spectroscopic data (e.g. for spectral classification of stars and galaxies) the retrieval of information from very large databases of astronomical data

(Longo, Tagliaferri, Sessa, Ortiz, Capaccioli, Ciaramella, Donalek, Raiconi, Staiano, & A. Volpicelli, 2001; Longo, Donalek, Raiconi, Staiano, Tagliaferri, Sessa, Pasian, Smareglia & Volpicelli 2002), the automatic scheduling and the automatic evaluation of the performances of observing facilities. The known applications also comprise the use of artificial neural networks in geology and geophysics for e.g. stratigraphy, cycles detection and paleomagnetism (Tagliaferri, Pelosi, Ciaramella, Longo, Barone & Milano, 2001), the analysis of seismic data and geologic images (Acernese, Ciaramella, De Martino, De Rosa, Falanga & Tagliaferri, 2002; Acernese, Barone, De Rosa, Eleuteri, Milano, Tagliaferri, 2001), the analysis of synthetic aperture radar images for geological and geophysical prospecting (Zhong, Wicks, Power, Dzurisin Thatcher & Masterlark, 2002; Dowd, Vachon, Dobson & Olsen, 2001; Legarsky, Gogineni & Akins, 2001; Schulz-Stellenfleth, Horstmann, Lehner & Rosenthal, 2001; Lin, Alpers, Khoo, Lim, Lim & Kasilingam, 2001; Horstmann, Koch, Lehner & Tonboe, 2000), and environmental data analysis (Andretta, Eleuteri, Fortezza, Manco, Mingozzi, Serra & Tagliaferri, 2000).

The value and variety of these applications suggest the need of a cumulative investigation on the produced neural techniques, algorithms and theories for the analysis of complex scientific data in astronomy, geology and geophysics. In fact, in spite of the enormous amount of work done, no comprehensive summaries of the main techniques have appeared so far. These reviews would be worthwhile for all people active in the field.

In particular, in this paper we focus on neural independent component analysis with application to remote sensing by synthetic aperture radar imagery processing.

The Independent Component Analysis (ICA) is a well-established statistical signal/data processing technique that aims at decomposing a set of multivariate signals into a base of statistically independent data-vectors/streams with the minimal loss of information content. The main two recognized purposes of ICA are:

• Linear blind source separation: In this case the aim is to recover a number of statistically independent signals from their unknown linear mixtures, under simple consistency conditions. Namely, a linear mixture of independent source signals is supposed to be observed, and on the basis of these only available information, the original source signals are recovered from their mixtures. The classical hypotheses made on the unknown sources are: 1) each source signal is an independent identically distributed (IID) stationary random process; 2) the source signals are statistically independent at any time; 3) at most one among the source signals has Gaussian distribution. This minimal set of requirements ensures the existence and uniqueness of the solution to the blind separation problem (but for ordering, sign, and power scaling), the minimal loss of information about the sources and, as a by-product, the identification of the mixing-model features;

• Data representation and visualization: High-dimensionality data/signals are difficult to handle and to visualize, but often contain significant redundancies, which make their actual information-structure dimensionality considerably lower than their representation dimensionality. These concurrent facts suggest the possibility to design signal/data processing algorithms capable of finding a suitable lower-dimensionality representation of the signals at hand by reducing the statistical dependencies among them. In this context, the independent component analysis technique has proven to provide a suitable solution through the concept of independent latent variables: The ICA may discover a linear projection of the data into a low-dimensional basis of statistically independent signals, that carry on no mutual information, thus providing a parsimonious maximally-informative representation of the original data. The basis data-streams are termed latent variables, which do not necessarily possess a recognizable physical meaning.

The classical example used to informally explain the blind separation problem is the "cocktail party" scenario: Let us imagine a number of people stand in a room and speak together; if the room is equipped with a number of microphones, each sensor receives a different superposition of the speech signals uttered by each person in the room, so that the set of received signals may be described by a linear mixture model, that takes into account the emitters-sensors geometry and the pressure waves propagation phenomena. The aim of blind source separation algorithm is to recover, *from sensor observations only*, the single independent signals uttered by the people.

An example that helps clarifying the concept of latent variables extraction

is the analysis of vibrating machines: Le us suppose a rotating machine (e.g. a rotating pump under fault test) is equipped with a number of accelerometers that measure its vibration intensities versus time; we know that the recorded accelerometer signals are originated by a large number of small vibrating parts, that exceeds the reasonable number of different measurements, which in turn may exceed the number of reasonable signals we can process for fault testing. By seeking for a compact representation of the measured data by the ICA technique, the obtained basis of independent signals are the complex vibrations of a small number of virtual oscillators whose linear superposition generate the observed data.

A dual-pair of interesting examples that further clarifies the independent latent variables concept is related to the analysis of natural images and natural sounds: In a classical paper by Bell & Sejnowski (1997) it was shown that the latent variables for natural scenes are edge filters, that closely resemble the Gabor filters, while in a recent dual paper, Abdallah & Plumbley (2001) investigated the same question for natural sounds; it is also worth mentioning that recently Seifritz, Esposito, Hennel, Mustovic, Neuhoff, Bilecen, Tedeschi, Scheffler & Di Salle (2002) have investigated the temporal decomposition of sound-evoked, blood oxygen level-dependent signal response into independent transient and sustained components, which predominate in different portions (core and belt) of the a ditory cortex, confirming the suitability of ICA techniques for the analysis and explanation of some brain-related activities.

It is definitely apparent that much theoretical research work has been carried out by several researchers over recent years in order both to produce ever refined algorithms for performing independent component analysis, and to publish consistent results about the basic theory of ICA (concerning the signals models and the solvability of the analysis problem related to these models) and about the algorithmic-level theory (concerning the theoretical study of the relevant properties of the different classes of algorithms, such as convergence, reliability, computational burden, equivariancy, and implementation questions). The theoretical research work carried out so far has been summarized in two books (Lee, 1998; Hyvärinen, Karhunen & Oja, 2001) and in a series of journal papers (Cardoso, 1998; Cardoso 1999; Hyvärinen & Oja, 2000; Fiori, 2000a; Fiori 2001b).

Concurrently, the scientific community has manifested an ever increasing in-

terest in the ICA technique and, generally speaking, in the blind signal processing research field, because it provides a powerful tool for signal/data processing, comparable in importance to the older principal component analysis. Known applications range from speech recognition to fault detection, from telecommunications to medical imaging, from financial data market analysis to biological data pre-processing, and from industrial plant identification to non-destructive evaluation (Clemente & Acha, 1997; Karhunen, Hyvärinen, Vigario, Hurri & Oja, 1997; Liu, 1996; Cichocki & Cao, 1998; Cichocki, Karhunen, Kasprzak & Vigario, 1999).

Synthetic aperture radar refers to a technique used to synthesize a very long antenna by combining signals received by the radar as it moves along its flight track, or, in other terms, a synthetic aperture is constructed by moving a real aperture or antenna through a series of positions along the flight track. As the radar moves, a pulse is transmitted at each position and the return echoes pass through the receiver and are stored in the memory of the acquiring equipment.

Synthetic aperture radar is a technique used to generate radar images in which fine details can be resolved. SARs provide unique capabilities as an imaging tool. Because they provide their own illumination (the radar pulses), they can image at any time of day or night, regardless of illumination, and because the radar wavelengths are longer than those of visible or infrared light, SARs imaging is independent of cloudy and dusty conditions.

The aim of the present paper is to provide a non-mathematical overview of the independent component analysis problem, its models and the state-of-theart algorithms, as well as of some interesting applications in the engineering area (section 2). After providing a brief overview of the remote sensing problem (section 3), an application of ICA technique to synthetic aperture radar image enhancement is discussed with experiments on real-world data (section 4).

2 Overview of Independent Component Analysis

The number of available contributions in the independent component analysis field is motivated by the wide variety of different observed-signals models considered in the applied areas. An informal but sufficiently representative list of models is given by:

- Instantaneous or convolutional mixture: When the relationship describing the formation of the observed signals/data from the source/latent signals does not take into account the temporal structure of the latter, i.e. the linear mixing operator is constant over time, the model is termed instantaneous. Otherwise, a convolutional model can be considered, which described the observed data/signals as the results of multiple linear filtering applied to the source/latent streams; in this case the linear mixing operator may be represented as a time-sequence of mixing matrices, a matrix of filters, or a matrix of functions in the frequency-domain;
- Linear or non-linear mixing: In the simplest (though much representative) case the mixing model is linear. Otherwise, some kinds of non-linear structures have been considered, such as the post-linear mixing model, in order to take into account e.g. the non-linear distortion introduced by the measurement systems;
- Real-valued or complex-valued models: In some applied fields, such as e.g. in telecommunications, it is useful to treat the involved signals as complex-valued data-streams, thus dedicated algorithms have been developed for complex-valued models handling;
- Square, over- and under-determined mixtures: The number of available measures may be equal to the number of source signals in blind separation, in which case the mixture is termed square; from a theoretical point of view this hypothesis allows carrying out some useful mathematics, while from a practical point of view it is not always realistic. The cases where the number of sources is larger or smaller than the number of observations are properly represented by the over- and under-determined models. Similar considerations may be carried out when latent-variable models are dealt with;
- Noiseless or noisy mixtures: The signals model may or may not take into account the possible presence of additive disturbance affecting the measured signals;
- Stationary or non-stationary model: The hypothesis of having IID source or latent signals is not always realistic; real-world random signals may exhibit time-fluctuations of their statistical features, and this phenomenon

may affect the performance of the independent component analysis if not properly taken into account. In opposition to stationary models, nonstationary ones try to capture the non-stationarity of the sources and to make the related ICA algorithm take advantage of this knowledge.

2.1 Overview of independent component analysis methods

The roots of basic ICA can be traced back to the work of Darmois in the 1950's and Rao in the 1960's, concerning the characterization of random variables in linear structure (Darmois, 1953; Kagan, Linnik & Rao, 1973).

The pioneering work on independent component analysis was carried out by Jutten & Hérault (1988; 1991), who introduced an adaptive algorithm in a feedback neuromimetic multiple filter. Their approach was further developed by Cichocki & Umbehauen (1996).

Later on, Comon (1994) formalized and developed the basic theory of independent component analysis, concerning the general conditions of existence, uniqueness and estimation indeterminacies; Comon also cast the ICA-computation problem as an optimization one and proposed a class of cost functions termed discriminant contrasts. The contrast-based approach to ICA has been further studied by Comon & Moreau (1997) who introduced simplified contrasts for certain classes of signals (i.e. for signals with positive or negative kurtosis). On the basis of simplified contrasts (i.e. on kurtosis optimization), Cardoso & Laheld (1996) proposed an adaptive algorithm relying on the relative gradient, which has been proven to enjoy the very desirable property of equivariancy, that is, the source extraction quality in blind source separation is independent of the mixing matrix and, in particular, from its conditioning number. Also, Delfosse & Loubaton (1995) proposed an algorithm based on a deflation procedure, whose idea has been recently developed by Thawonmas, Cichocki & Amari (1998) resulting in a cascade neural network. The deflation or sequential approach to ICA consists in extracting one component at a time from the original data, deflated from the previously-extracted components, while non-sequential or parallel approaches try to extract all the required components concurrently. Parallel independent component analysis algorithms may suffer from convergence difficulties, owing to the complexity of the search space, and from computational burden, especially for large-dimensionality data; on the other hand,

sequential algorithms may suffer from error accumulation from component to component. An experimental comparison presented in (Fiori, 2001b) suggests a hybrid sequential-parallel approach, i.e. sequential extraction of groups of components.

Concurrently to statistical signal processing techniques, unsupervised learning rules based on information theory were proposed. The aim was to maximize the mutual information between the inputs and the outputs of a neural network, so that each neuron matches features being as statistically independent as possible from the other neurons. Nadal, Brunel & Parga (1998) showed that in the low-noise case, the maximum of the mutual information between the input and the output of a network implies that the output joint probability density function (pdf) factorizes as a product of marginal probability density functions. Roth & Baram (1996) and Bell & Sejnowski (1996), independently devised stochastic gradient learning rules for ensuring mutual information maximization and applied them to forecasting and time-series analysis, blind separation of sources and blind deconvolution, respectively. Their approach was further developed by Yang and Amari (1997). Girolami & Fyfe (1997) employed neural exploratory projection pursuit algorithms for achieving separation. Generalized Hebbian learning algorithms for ICA have been developed by Karhunen, Oja, Wang, Vigario & Joutsensalo (1997), Hyvärinen & Oja (1998), Oja (1997) and Fiori (2000b; 2001b). The techniques based on maximum likelihood estimation (Pham, Garrat & Jutten, 1992; Pearlmutter & Parra, 1996; Belouchrani & Cardoso, 1995) and the recent proposal of Bayesian ICA (Knuth, 1998; Roberts, 1998) and maximum-a-posteriori (MAP) parameter estimation for blind source recovering (Parra, Mueller, Spence, Ziehe & Sajda, 2000) are of interest because they enable us to cast the ICA-computation problem in a way that allows to take into proper account the a-priori information on the sources or the signal models. As a meaningful example, in blind localization of sources emitting pressure or electromagnetic waves, the known features of the propagation model and physical constraints may be inserted into the separation algorithm. Also, Sagi, Nemat-Nasser, Kerr, Hayek, Downing & Hecht-Nielsen (2001) recently developed a new technique based on the 'cortronic' neural network, a biologicallyplausible neural model, while Welling and Weber (2001) proposed a EM-type learning procedure for ICA computation by neural networks.

The basic principle that the MMI-ICA technique is based on, is that the

mixing phenomena transform a set of statistically independent random signals into a set of statistically-dependent signals. Thus, in order to attain separation, the neural network should learn how to make the network's outputs become as independent as possible; under proper consistency conditions, this ensures the output coincide to the original independent sources, up to arbitrary order permutation and components scaling. In principle, the independence of the network's outputs may be measured in terms of the output signals joint probability density function and marginal densities: If these signals are statistically independent, the joint probability factorizes into the product of the marginal densities. A way to achieve separation is thus to define a measure of the mismatch between these two quantities, and a learning algorithm to learn the network connection-matrix in order to minimize such disagreement. A useful and widely employed measure of statistical dependency is the mutual information between network's output signals.

Once that an appropriate criterion has been defined as a function of network's connection-matrix, the network learning phase may be formally conceived as an optimization procedure allowing to search for the connection pattern that minimizes or maximizes the criterion.

As improvements to original gradient-based MMI-ICA learning rules by Bell & Sejnowski (1996), several new optimization techniques, oriented to efficient extraction of independent components, have been recently developed, such as the fixed point technique (Hyvärinen & Oja, 1997), the natural-gradient one by Amari (Amari, 1998; Yang & Amari, 1997), and the Riemannian gradient on Stiefel manifold and orthogonal group (Fiori, 2001a; Fiori, 2002b; Fiori, 2002d).

A problem arising in MMI-ICA algorithms implementation by neural networks is source adaptivity: The MMI principle is based on non-linear transformations, implemented by the neurons' activation functions, which depend on components' probability density functions; these are of course unknown and need to be estimated from the networks' outputs. Some adaptivity is achieved by the use of adaptive activation function neurons, realized by the help of the mixture-of-kernels technique (Xu, Cheung, Yang & Amari, 1998; Xu, Cheung & Amari, 1998), standard multilayer perceptrons (Taleb & Jutten, 1997), quasi-Dirac-kernel functions (Gustaffson, 1998) and quasi-polynomial flexible functions as well as discrete-histograms (Fiori 1999; Fiori 2000a; Fiori 2002c; Fiori & Bucciarelli, 2001). Non-conventional neural optimization techniques have been recently applied to blind separation: As three examples, the present author employed mechanical-type learning algorithms to blind separation by the ICA (recently extended to complex-valued sources) (Fiori, 2000b; Fiori 2002a), Prieto & Puntonet (1998) and Puntonet & Prieto (1998) developed a purely geometrical approach to be applied in presence of unimodal-symmetrical sources density distributions, while Yoshioka and Omatu (1998) applied a genetic algorithm to minimize ICA cost functions.

Other algorithms have been proposed from different perspectives by several authors. It is in fact known that independent component analysis relies on some assumptions which limit its fields of application: Researchers have recently tried to overcome this problem by extending the classical algorithms to convolutional, under-determined, non-stationary, and non-linear mixtures.

The first goal that may be considered concerning non-stationarity in ICA consists in making ICA algorithms robust with respect to possible source nonstationarity. This has been achieved by extended versions of the classical Hérault-Jutten algorithm for linear instantaneous mixtures and of the corresponding Nguyen-Jutten algorithm for convolutional mixtures. The solution proposed consists in normalizing the adaptation terms of these algorithms by short-term estimates of the powers of the estimated sources. This may also be seen as a time-varying adaptation gain and makes the algorithms able to automatically track the time-varying levels (power values) of the sources. This is of special interest for such signals as speech, where fixed-gain algorithms cannot combine fast convergence and good component extraction accuracy. Other related features result from this normalized approach, especially the ability to track time-varying mixtures. More details about this approach may be found especially in (Deville, Damour & Charkani, 1999), which introduces this approach in the linear instantaneous case and discusses the above case of unknown but fixed source levels, and (Charkani & Deville, 1999a; 1999b), which present a much more complete approach for the convolutional case.

It has also been proposed to take advantage of the assumed non-stationarity of the considered sources. Two aspects may be distinguished in this framework: We may still restrict ourselves to the case that the number of observations is at least equal to the number of sources, or consider the underdetermined case, i.e. the situation arising when the number of sources is larger than the number of observations. In order to address the latter, more complex, case, the general concept of differential source separation was introduced (Deville & Savoldelli, 2001).

A somewhat related way to tackle ICA problems consists in using their time-frequency representations. One then takes advantage of assumed specific properties of the time-frequency signatures of the sources, which appear for non-stationary sources. For linear instantaneous mixtures, a preliminary description of these methods appeared in (Abrard, Deville & White, 2001a; 2001b).

2.2 Relationships with projection pursuit

It might be interesting to discuss the relationships among blind sources separation by the independent component analysis and the closely-related statistical theory of projection pursuit (PP) (Friedman & Tukey, 1974; Huber, 1985).

While it is true that the mathematics of ICA and PP are very closely related in the basic case, the following differences in the application, history, and also the theory of these methods deserve to be underlined:

- Independent component analysis was originally devised to solve a concrete separation problem, whereas projection pursuit was originally developed for visualization;
- In PP, the starting point is dimension reduction of a large-dimensional space. In ICA, the starting point is estimation of components or latent variables (basic signals), and this is meaningful even in two-dimension;
- If the time-structure of the signals and the mixing system or signal model are taken into account, ICA is very different from PP.

These points are clearly illustrated in one of the original motivations for ICA (Jutten & Héault, 1991): Two nervous fibers transmit mixtures of signals related to joint stretch and stretch speed, and these two signals need to be separated by the central nervous system; finding interesting projections is hardly the proper approach in this case.

Also, basically ICA searches simultaneously for many components with the criterion that they are mutually least dependent. On the other hand, PP extracts components sequentially with the criterion to be most interesting (non Gaussian). There is no penalty for extracting (nearly) the same component

again; instead, what is usually done is to subtract the extracted component from the data (through a regression) to avoid it being extracted again. However, if a component is not exactly extracted, it will contaminate all subsequent steps in PP; this problem is in fact encountered in sequential ICA as well.

In the blind source separation by the ICA, the model of the data can be taken into account: After separation is completed, as a by-product we could also obtain an estimate of the mixing model. However, the latter operation is not troublefree; apart from numerical/precision/size problems, there are the well-known indeterminacies: The separation algorithm causes an arbitrary re-ordering of the source signals, and, due to the multiplicative nature of the model, scale factors and phase rotations (sign switch, in the real-valued case) appear. This reflects in some indeterminacy in the elements of the estimated model. These problems are fundamental, in the sense that there is no way to suppress them. In fact, in many algorithms, it is supposed that the sources have unit powers, so that the 'true' powers result in mixing matrix column scaling. However, it is quite natural to suppose that if we know in advance that the mixing matrix has some special structure (i.e. induced by the physical properties of the problem at hand), it could be possible to remove some indeterminacy.

In blind source localization by ICA, for instance, it is possible to solve for the mixing matrix, however, at some point the mixing matrix coefficients become redundant as there will be fewer parameters describing the source positions than mixing matrix elements. Using a Bayesian methodology, one can derive more powerful algorithms that solve for these parameters rather than the elements of the mixing matrix. As mentioned, in this case, the algorithm provokes a cross-over from the source separation problem to the problem of source localization. The two are intimately related: Depending on the prior information possessed and the information sought for, we obtain either a source separation algorithm, a source localization algorithm, or simultaneous source separation and localization algorithm.

2.3 A pathway to geometric integration

In orthonormal independent component analysis (by pre-whitening) as well as in many unsupervised learning theories, the common way to design a learning algorithm arises from the following considerations. The lack of fundamental information in the engineering problems which require the use of unsupervised neural systems is overcome by introducing concepts referring to high-order statistics, information theory and optimization. The three mentioned mathematical theories result to be closely related and widely employed in the literature in order to cope with blind signal processing problems. By making use of the concept of the information theory, in fact, it is possible to define suitable objective functions which describe in informational terms the processing problem under analysis. The mathematical statistics gives then some formal instruments for elaborating theses objective functions, and the optimization theory constitutes the analytical basis for synthesizing global functions containing eventual physical constraints arisen from the problem, and for determining learning algorithms endowed with the necessary requisites about convergence speed, steady-state precision, and sensitivity to external disturbances.

The final result of the preceding design work is often a set of learning differential equations, written in matrix form, whose solution gives the learning trajectory where the mentioned physical constraints denote what mathematicians refer to as invariants associated to the differential equations. In order to respect the invariants, that is to preserve the qualitative characteristics of the solutions, the learning initial-value problems should be integrated numerically in a proper and efficient way. In particular:

- **Properness**: The differential equations should be integrated in a way that preserves the invariants both in order to ensure the quality of the signal processing solution provided by the neural system and to preserve some quantitative features of the learning theory such as intrinsic stability (for an expanded discussion see Fiori (2002d));
- Efficiency: An integration method that taks into account the structure of matrix-type expressions involved in the learning equations might possess contained computational complexity (Fiori & Celledoni, 2002).

The proper mathematical theoretical setting for discussing and formalizing the mentioned topics is the Geometric Integration (GI). The classical efforts in numerical analysis have been to model physical phenomena into algorithms that produce sufficiently accurate and affordable numerical approximations of their behavior. Geometric integration is concerned also with producing numerical approximations preserving the qualitative attributes of the solutions: Some examples of GI algorithms for differential equations include Lie group integrators, volume and energy preserving integrators, integrators preserving first integrals, Lyapunov functions and Casimirs, as well as Lagrangean and variational integrators (Hairer, Lubich & Wanner, 2002).

Preliminary and encouraging results in the field of independent component analysis in presence of linear mixtures of real-world signals have been obtained recently as reported in (Fiori & Celledoni, 2002; Fiori & Rossi, 2002).

3 Remote Sensing and Synthetic Aperture Radar (SAR)

Remote sensing is the science of acquiring information about material objects, areas, or phenomena, without coming into physical contact with the objects, areas, phenomena under investigation.

In absence of direct contact, some mean of transferring information through space is necessary: In remote sensing, information transfer is accomplished by the use of electromagnetic radiation. On the basis of the physical type of used energy sources, we distinguish between *passive remote sensing*, which makes use of sensors that detect the reflected or emitted electromagnetic radiation from natural sources, and *active remote sensing*, which exploits sensors that detect reflected responses from objects that are irradiated from artificially-generated energy sources, such as radar. A schematic of common remote sensing platforms with on-board radar sensors is depicted in the Figure 1. Remote sensing measures are recorded in digital form and then digitally processed in order to produce images for interpretation purposes. Variations in the scene's characteristics are represented as variations in brightness on recorded images, as usually a particular part of a scene which reflects more energy appears bright, while a part of the same scene that reflects less energy appears darker.

Digital images consist of discrete pixels, whose intensity value represents the average radiance of a relatively small area within a scene, where the size of the elemental area affects the reproduction of details within the scene; then, digital image processing techniques are necessary to retrieve the desired information from the raw images.

Digital image processing is a collection of techniques for the manipulation of digital images by computers, which encompass the operations of image data manipulation and management, geometric and radiometric corrections, enhance-



Figure 1: Schematic of common platforms with radar sensors on board. (Reprinted, with permission, from www.gisdevelopment.net).

ment of images quality, and information extraction. Image processing methods may be grouped into four functional categories: Image management, image restoration, image enhancement, and information extraction. These categories may be briefly described as follows:

- Image management: Image acquisition and storage is the first step in image analysis. In particular, pictorial data storing requires special attention due to the large size of this kind of data, that make it useful to perform some kind of compression and size-reduction in order to save memory supports. Usually natural images are highly redundant (correlated), i.e. parts of them can be reconstructed from the knowledge of few features; principal component analysis (or Karhunen-Loeve transform) is a widely known technique that allows representing real-world data by a small amount of uncorrelated features (for a recent review see e.g. Costa and Fiori (2001)).
- *Image restoration*: The process of image acquisition may degrade the observed scenes because of measurement errors, as for example the blur effect caused by a non-focused optical system, and of external causes, such

as the opacity of the mean that the images are acquired through. The images restoration procedures tend to ameliorate the quality of recorded scenes by removing noise and distortion effects introduced by the imaging system(for a recent review see e.g. Kundur and Hatzinakos (1996)).

- Image enhancement: Even if perfectly reproduced, a digital image may not yet be suitable for elaboration purpose, because its intrinsic quality does not allow to reveal the interesting features of the phenomena it corresponds to nor to emphasize the information content it brings on. To this purpose, image enhancement techniques have been developed; they introduce usually quite subjective warping of the original data in order to emphasize certain attributes, as brightness, to calibrate other attributes, as saturation, and equalize other features, such as hue (for a recent review see e.g. Fiori, Grimani and Burrascano (2002)).
- Information extraction: The most "intelligent" operation to be performed on the basis of restored/enhanced available data is the extraction of relevant information, that facilitates scene interpretation. The subject of feature extraction, ranking and selection is covered by pattern analysis research field.

In the present paper we focus our attention to radar-based remote sensing and, in particular, to *synthetic-aperture* radar.

3.1 Synthetic aperture radar

Synthetic Aperture Radar (SAR) is an active microwave device, producing highresolution imagery of the Earth's surface (for a recent review see e.g. Hogda, Guneriussen & Lauknes (2002), Lou (2002), Brown & Bennett (2001), and Jao (2001)). Known past and present Earth observation satellites are the ERS-1/ERS-2, JERS-1, the Shuttle imaging radar SIR-C/X-SAR, and RADARSAT. An example of their use is the ESA/Eurimage "Earthwatch" program, producing imagery of natural and artificial disasters when weather conditions prevent other forms of surveillance from being profitable.

Two important properties distinguish SAR from classical optical imagery:

• The SAR is an active device: It generates its own illumination of the scene to be viewed. The illumination is coherent: All the light in any flash is

exactly in phase, as in a laser, so it does not disperse over the distance between the satellite and the Earth's surface. A SAR device can measure both intensity and phase of the reflected light, resulting highly sensitive to textures. Experiments with the technique of interferometry have shown that SAR can accurately model relieves, and appears also able to reveal small changes over time. Some consequences are that it works day and night as well, it can be used to gain additional information with respect to optical imagery, especially when different polarization are available on the same platform, it needs more power than passive sensors to be operated (and can therefore only operate intermittently), and that it suffers from speckle, an artifact of interference patterns in coherent light.

• The SAR is a radar: It uses microwave frequency radiation, which penetrates cloud and haze, so it views the Earth's land and sea surface in all weather, that is the major advantage of SAR for general-purpose remote sensing.

The first implementation of radar interferometry came in Earth-based observations of Venus (Rogers & Ingalls, 1969). The first reported experiments to determine terrain elevation of the Earth were by Graham (1974). Ten years later, interferometric radar experiments on the airborne system Convair-990 and on the spaceborne systems Seasat and SIR-B took place. Since 1990 the interest in SAR interferometry has grown due to the impressive amount of data suitable for interferometry from ERS-1 and the many airborne systems available such as the AIRSAR.

SAR images are suitable for vegetation studies, as well as ocean waves, winds, currents, seismic activity and moisture content. In practice, by properly processing the complex SAR images, it is possible to obtain high-resolution topographic maps (5 m or less height resolution), measure very small (1 cm or less) Earth surface motion over large swaths, measure water surface currents (with an accuracy around 5 cm/s) and classify land surfaces. In particular, the following main applications are worth citing:

• Ground topography: Studies on vulcanology and Earth surface motion related to differential SAR interferometry are well known. ERS-1 helps the evaluation of digital elevation models with a grid spacing of about 50 m and a height accuracy of about 5 m. AIRSAR and Do-SAR allow the evaluation of the digital elevation models with a grid spacing smaller than 10 m and a height accuracy around 1 m. The SAR-derived digital elevation models (DEM) have a big impact in the field of the topography: They are replacing the stereo DEMs derived from optical systems (Hogda, Guneriussen & Lauknes, 2002).

- Ocean surface current measurements: Ocean surface currents having speeds of less than 4 cm/s have been observed by Goldstein & Zebker (1987).
- Earth surface motion detection: Massonet (1993) shows the possibility of measuring the residual displacement caused by the earthquakes. The differential interferometry is clearly validated for long term survey of slow faults (typically 10 mm/year measured with ERS-1). This is a remarkable result, because, due to the motion errors of the aircraft, it is very difficult to implement the differential interferometry in an operational way.
- Land surface classification: By carrying out repeat-pass interferometry, coherence maps and change detection of SAR images can be used to provide properties of land surfaces. Results using ERS-1 data show the capability of the classification of forest, open fields, urban areas and open water (Lin, Alpers, Khoo, Lim, Lim & Kasilingam, 2001).

An exemplary commercial SAR system is depicted in the Figure 2: It is a PREDATOR Lynx synthetic aperture radar system.

3.2 Some details on imaging radar and SAR

An imaging radar uses an antenna and a digital computer to store the acquired images. A radar image is generated only by the light that gets reflected back towards the antenna.

Radar measures the strength and round-trip time of the microwave signals that are emitted by an antenna and reflected from a distant surface or object. Its antenna alternately transmits and receives pulses at particular microwave wavelengths (in the range of 1 cm to 1 m, which corresponds to a frequency range of about 300 MHz to 30 GHz) and polarizations (waves polarized in a single vertical or horizontal plane). About 1500 high-power pulses per second are transmitted towards the target, with each pulse having a pulse duration of



Figure 2: A PREDATOR Lynx synthetic aperture radar system. (Reprinted, with permission, from www.lynxsar.com.)

typically 10-50 microseconds. The pulse normally span a small band of frequencies, centered on the frequency selected for the radar; typical bandwidths are in the range 10 to 200 MHz. At the Earth's surface, the energy content of the incoming radar pulse is scattered in all directions, while only a fraction of it is reflected back towards the antenna. Such energy returns to the radar as a weaker radar echo and is received by the antenna in a specific polarization (not necessarily the same as the transmitted pulse). Since the radar pulse travels at the speed of light, it is relatively straightforward to use the measured time corresponding to the roundtrip of a particular pulse to calculate the distance or range to the reflecting object.

In the case of imaging radar, the radar moves along a flight path and the area illuminated by the radar, or footprint, is moved along the surface, building the image (see Figure 3). The chosen pulse bandwidth determines the resolution in the range direction (higher bandwidth means finer resolution in this dimension), while the length of the radar antenna determines the resolution in the azimuth direction of the image (the longer the antenna, the finer the resolution in this dimension). In an imaging radar, the term "aperture" means the opening used to collect the reflected energy. Because the radar is moving with respect to the ground, the returned echoes are Doppler-shifted (negatively when the radar approaches a target, positively when it moves away). Comparing the



Figure 3: Exemplary representation of SAR footprint.

Doppler-shifted frequencies to a reference frequency allows returned signals to be "focused" on a single point, effectively increasing the length of the antenna. SAR data focusing consists in correctly matching the variation in Doppler frequency for each point in the image; this operation requires a precise knowledge of the relative motion between the platform and the imaged objects.

Some SARs can transmit pulses in either horizontal (H) or vertical (V) polarization and receive in either H or V modes, with the resultant combinations of HH (Horizontal transmit, Horizontal receive), VV, HV, or VH. Additionally, some SARs can measure the phase of the incoming pulse and therefore measure the phase difference (in degrees) in the return of the HH and VV signals. This difference is frequently retained as an indicator of structural characteristics of the areas or objects under observation. These SARs can also measure the correlation coefficient for the HH and VV returns, which can be considered as a measure of how alike the portions of the areas or objects are.

Radar images are composed of many dots, or picture elements. Each pixel in the radar image represents the radar backscatter for an area on the ground: Bright areas represent high backscatter (bright features mean that a large fraction of the radar energy was reflected back to the radar), while darker areas in the image represent low backscatter (dark features imply that very little energy was reflected back to the antenna). Backscatter for a target area at a particular wavelength varies because of several conditions, as the size of the scatterers in the target area, the moisture content of the target area, the polarization of the pulses, the values of emitted wavelengths, and the observation angles.

A rule that helps interpreting the radar images is that the brighter the backscatter on the image, the rougher the surface being imaged. Flat surfaces that reflect little microwave energy always appear dark in radar images. Vegetation is usually moderately rough on the scale of most radar wavelengths and appears as gray in a radar image. Some areas not illuminated by the radar, like the back slope of mountains, are in shadow, and appear dark. Roads and freeways are flat surfaces so they appear dark. Backscatter is also sensitive to the target's electrical properties, such as water content: Wetter objects appear bright and drier targets appear dark (with the exception of smooth bodies of water, which behave as flat surfaces and reflect incoming pulses away, thus they appear dark). Backscatter also varies depending on the use of different polarization and observations angles: Low incidence angles (perpendicular to the surface) will result in high backscatter, while it decrease with increasing incidence angles.

4 Application of ICA to SAR imagery processing and Experimental Results

Over the next section it is shown how independent component analysis by neural networks may be advantageously employed to enhance remote sensing data on real-world SAR imagery.

4.1 Independent component analysis applications to remote sensing

Very recently there has been interest in the use of independent component analysis methods for remote sensing. The use of principal component analysis for SAR imagery and remote sensing in general has been examined and well understood (Fiori & Piazza, 2000; Costa & Fiori, 2001; Cheng & Miller, 2002), while there has been much less work carried out about the use of ICA in image analysis. The existing contributions in this fields concern mainly the three following topics:

- Speckle reduction: A role played by the independent component analysis in synthetic aperture radar data processing is image speckle reduction: The SAR data can be considered as a mixture of target signals and some complex noises. ICA methods may be used to separate the noise from the signals as much as possible.
- Feature extraction: Perhaps the most important role of independent component analysis in SAR imagery processing is feature extraction for classification. Selected independent components may be used to form feature vectors for classification. ICA has been found to perform better than PCA, as comparable classification percentage is achieved with less independent components than principal components (Chen & Zhan, 2000).
- Data fusion: Another important role of independent component analysis is in SAR data fusion that allows to advantageously employ the data redundancy provided by radar measurements. When more sensors are available to observe the same area or objects on the best extracted component images may be merged using appropriate criteria. It has been shown that the image obtained can be more informative than the original images in any spectral band (Chen & Zhan, 2000).

Other interesting notes on the theory and applications of neural ICA to remote sensing may be found on the recent book by Chicocki and Amari (2002).

In the present paper, we consider the problem of noise removal by independent component extraction. The algorithms were implemented in MATLAB language and ran on a 500MHz, 64MB machine.

4.2 SAR data and results of noise removal by ICA

The considered image data is from an agricultural area near the village of Feltwell, United Kingdom and consists of 9 channels of SAR images: The data consist of three frequency bands (denoted, respectively, with c, p, and l); in each band there are three different polarizations (HH, HV, and VV); the available SAR images are depicted in the Figure 4. It is interesting to observe the differences in the recorded images for the different frequency-bands: It clearly appears



Figure 4: Nine channels of SAR images of an agricultural area in the United Kingdom.

that SAR is sensitive to different features of the observed areas depending on the frequency of the emitted electromagnetic waves.

The practical application of ICA algorithms to these 2-dimensional data requires to perform the following pre-processing steps:

- 1. Each image is 250×250 pixels bitmap, that should first be vectorized, thus the set of available images gives rise to a 9×62500 data-matrix;
- 2. The data-matrix may be linearly normalized in order to transform the range of pixel intensity from the natural domain into a conventional datadomain, for instance [-1, +1]; in the present case, this operation is not

necessary and is therefore omitted.

- 3. The normalized data should now be subjected to pre-whitening, which consists in removing first-order and second-order statistics from the data. Some ICA algorithms explicitly require this operation to be performed, while other known algorithms do not need pre-whitening; however, it has been experimentally observed that whitening facilitates independent components extraction, because it facilitates convergence (Giannakopoulos, Karhunen & Oja, 1999); also, the analysis of the covariance matrix of data, and in particular of its spectrum, may reveal some interesting property about the data's information-structure dimension.
- 4. ICA applied to the whitened data allows extracting the desired number of independent components. In the present context the model of the data is linear, real-valued, and instantaneous.
- 5. Re-normalization of extracted components and restoration of the original 2-dimensional size of data allows then results visualization. In the present context we perform blind separation of a SAR image from noises, thus it is supposed that the obtained results have a recognizable meaningful.

On the basis of the above-recalled algorithm, we proceeded first to datapreparation and pre-processing. The whitening operation, in particular, is quite meaningful: The 9×9 data-covariance matrix has well-ordered eigenvalues, as clearly shown by the Figure 5; the fact that 4 eigenvalues out of 9 are significantly larger than the remaining eigenvalues leads to conclude that by projecting the data over the basis formed by the four principal eigenvectors a significant data-size reduction may be achieved with a negligible loss of information. The resulting transformed data-matrix has then the size of 4×62500 . the smallest remaining (non-zero) eigenvalue is 375.06 and the largest remaining (nonzero) eigenvalue is 3536.62. The pre-whitening makes the covariance matrix of the whitened data an identity: In this case the actual covariance matrix of the reduced-size data differs from an identity by a factor of 2.2×10^{-13} . The whitened data have now been processed by three different independent component analysis algorithms:

• The DICA+ software (Frulloni & Fiori, 2002) is based on a cascade neural network by Thawonmas *et. al* (1999); this method extracts the com-



Figure 5: Eigenvalues of the 9×9 SAR-data covariance matrix.

ponents sequentially and iterates the whole data-stream several times in order for the active neuron in the cascade to properly extract each source signal.

- The FASTICA software is based on a generalized-Hebbian learning theory and fixed-point optimization (Hyvärinen & Oja, 1997); it has been used in parallel mode.
- The MEC learning rule is based on a rigid-body learning theory applied to kurtosis extremization (Fiori, 2002b); it extracts the independent components in a parallel way.

The final result of elaboration is shown in the Figure 6 for the DICA+, Figure 7 for the FASTICA, and Figure 8 for the MEC-ICA algorithm: Four independent components where extracted, and one of them clearly appears as the true SAR image. In the three experiments, the extracted meaningful component exhibits a better quality than the original SAR images.

In particular, for the DICA+ algorithm, the component IC₁ looks the best (least noisy) one, while the others appear as complex noises. The algorithm autonomously ran for 83, 59, 32, 10 iterations of the data set for the first, second, third and fourth neuron in the cascade, respectively, and a total computation time of about 2 minutes and 6 seconds.

For the FASTICA algorithm, the component IC_2 looks the best one, the others being complex noises. The algorithm autonomously ran for 21 iterations



Figure 6: Final result of ICA elaboration by the DICA+ software.



Figure 7: Final result of ICA elaboration by the FASTICA software.



Figure 8: Final result of ICA elaboration by the MEC-ICA algorithm.

of the whole data set in the parallel mode and a total computation time of about 1 minute and 36 seconds.

For the MEC-ICA algorithm, the component IC_4 looks the least noisy one. The shown results pertain to the use of just the first 20,000 learning data out of the 62,500 available and the algorithm ran for about 16 seconds. The number of learning steps was chosen by observing the two curves characterizing the MEC-theory: The network's kinetic and potential energy (Fiori, 2002b); they are shown in the Figure 9.

5 Conclusion

The aim of this work was to present an overview of independent component analysis technique based on neural networks, with the aim to provide a state-ofthe-art survey of the theoretical streams in this research area and a brief review



Figure 9: MEC network's kinetic and potential energies during learning over the first 20,000 data-points out of the 62,500 available.

of some applications of engineering interest. We also focused on a particular application dealing with synthetic aperture radar imagery processing, and briefly reviewed the features and main applications of synthetic aperture radar and showed how independent component analysis by neural networks may be of use in order to remove noise from these remote sensing data.

References

S.A. Abdallah and M.D. Plumbley (2001). If the independent components of natural images are edges, what are the independent components of natural sounds ?, Proc. of the International Workshop on Independent Component Analysis and Blind Separation (ICA'01), pp. 534 - 539

F. Abrard, Y. Deville, P. White (2001a). A new source separation approach based on time-frequency analysis for instantaneous mixtures, Proceedings of the 5th International Workshop on Electronics, Control, Modeling, Measurement and Signals (ECM2S'2001), pp. 259-267, Toulouse, France

F. Abrard, Y. Deville, P. White (2001b). From blind source separation to blind source cancellation in the underdetermined case: a new approach based on time-frequency analysis, Proc. of the International Workshop on Independent Component Analysis and Blind Signal Separation (ICA'2001), pp. 734 - 739

F. Acernese, A. Ciaramella, S. De Martino, R. De Rosa, M. Falanga, and R. Tagliaferri (2002). Neural Networks for Blind Source Separation of Stromboli explosion quakes, IEEE Transactions on Neural Networks. Accepted for publication

F. Acernese, F. Barone, R. De Rosa, A. Eleuteri, L. Milano, and R. Tagliaferri (2001). A neural network based identification system for VIRGO Seismic noise, Proceeding of IJCNN'2001, IEEE Press, Vol. 1, pp. 252 - 257

S.-i. Amari (1998). Natural Gradient Works Efficiently in Learning, Neural Computation, Vol. 10, pp. 251 - 276

M. Andretta, A. Eleuteri, F. Fortezza, D. Manco, L. Mingozzi, R. Serra, and R. Tagliaferri (2000). Neural networks for sulfur dioxide ground level concentrations forecasting, Neural Computing and Applications, Vol. 9, pp. 93 -100

S. Andreon, N. Capuano, G. Gargiulo, G. Longo and R. Tagliaferri (2000). Wide Field Imaging. I. Applications of Neural Networks to Object Detection and Star/Galaxy Classification, Monthly Notices of the Royal Astronomical Society, Vol. 319, pp. 700 - 716

A.J. Bell and T.J. Sejnowski (1997). The 'independent components' of natural scenes are edge filters, Vision Research, Vol. 37, No. 23, pp. 3327 - 3338

A.J. Bell and T.J. Sejnowski (1996). An Information Maximisation Approach to Blind Separation and Blind Deconvolution, Neural Computation, Vol. 7, No. 6, pp. 1129 - 1159

A. Belouchrani and J.-F. Cardoso (1995). Maximum Likelihood Source Separation by the Expectation-Maximization Technique: Deterministic and Stochastic Implementation, Proc. of International Symposium on Non-Linear Theory and Applications (NOLTA), pp. 49 - 53

S.C.M. Brown and J.C. Bennett (2001). Reduced bandwidth algorithm for near field imaging of tree targets using synthetic aperture radar, IEE Proceedings - Radar, Sonar and Navigation, Vol. 148, No. 5, pp. 259 - 266

J.-F. Cardoso (1998). Blind signal separation: statistical principles, Proceedings of the IEEE (Special issue on "Blind identification and Estimation", R.-W. Liu and L. Tong, Ed.s), Vol. 90, No. 8, pp. 2009 - 2026

J.-F. Cardoso (1999). High-order contrasts for independent component analysis, Neural Computation, Vol. 11, No. 1, pp. 157 - 192

J.-F. Cardoso and B. Laheld (1996). Equivariant Adaptive Source Separation, IEEE Trans. Signal Processing, Vol. 44, No. 12, pp. 3017 - 3030

N. Charkani, Y. Deville (1999a). Self-adaptive separation of convolutively mixed signals with a recursive structure. Part I: stability analysis and optimization of asymptotic behaviour, Signal Processing, vol. 73, no. 3, pp. 225 - 254

N. Charkani, Y. Deville (1999b). Self-adaptive separation of convolutively mixed signals with a recursive structure. Part II: Theoretical extensions and application to synthetic and real signals, Signal Processing, vol. 75, no. 2, pp. 117-140

C.H. Chen and X. Zhan (2000). On the roles of independent component analysis in remote sensing, Proc. of Progress in Electromagnetics Research Symposium (PIERS'2000), Cambridge (MA, USA)

J. Cheng and E. Miller (2002). Model-based principal component techniques for detection of buried landmines in multiframe synthetic aperture radar images, 2002 IEEE International Geoscience and Remote Sensing Symposium (IGARSS '02), Vol. 1, pp. 334 - 336

A. Cichocki and J. Cao (1998). A self-adaptive neural network for on-line blind separation of convolved sources, Topics on Non-destructive Evaluation Series, Eds. B.B. Djordjevic and H.D. Reis, Vol. 3, The American Society for Non-destructive Evaluation Testing, Inc., pp. 207 - 212

A. Cichocki and R. Unbehauen (1996). Robust neural networks with on-line learning for blind identification and blind separation of sources, IEEE Trans. Circuit and Systems-I, Vol. CAS-43, pp. 894 - 906

A. Chicocki, J. Karhunen, W. Kasprzak, R. Vigario (1999). Neural Networks for Blind Separation with Unknown Number of Sources, Neurocomputing, Vol. 24, pp. 55 - 93, 1999

A. Chicocki and S.-i. Amari (2002), Adaptive Blind Signal and Image Processing, J. Wiley & Sons

R.M. Clemente and J.I. Acha (1997). Blind separation of sources using a new polynomial equation, Electronics Letters, Vol. 33, No. 3, pp 176 - 177

P. Comon (1994). Independent Component Analysis, A new concept ?, Signal Processing, Vol. 36, pp. 287 - 314

P. Comon and E. Moreau (1997). Improved Contrast Dedicated to Blind Separation in Communications, Proc. International Conference on Acoustics, Speech and Signal Processing, pp. 3453 - 3456

S. Costa and S. Fiori (2001). Image Compression Using Principal Component Neural Networks, Image and Vision Computing Journal (special issue on "Artificial Neural Network for Image Analysis and Computer Vision"), Vol. 19, No. 9-10, pp. 649 - 668

G. Darmois (1953). Analyse générale des liaisons stochastiques. Rev. Inst. Internat. Stat., Vol. 21, pp. 2 - 8

N. Delfosse and P. Loubaton (1995). Adaptive Blind Separation of Independent Sources: A Deflation Approach, Signal Processing, Vol. 45, pp. 59-83

Y. Deville, J. Damour, N. Charkani (1999). Improved multi-tag radiofrequency identification systems based on new source separation neural networks, Proceedings of the International Workshop on Independent Component Analysis and Blind Signal Separation (ICA'99), pp. 449 - 454, Aussois, France

Y. Deville, S. Savoldelli (2001). A second-order differential approach for underdetermined convolutive source separation, Proceedings of the 2001 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2001), Salt Lake City, USA, May 7-11

M. Dowd, P.W. Vachon, F.W. Dobson, and R.B. Olsen (2001). Ocean wave extraction from RADARSAT synthetic aperture radar inter-look image crossspectra IEEE Transactions on Geoscience and Remote Sensing, Vol. 39, No. 1, pp. 21 - 37

S. Fiori (1999). Entropy Optimization by the PFANN Network: Application to Independent Component Analysis, Network: Computation in Neural Systems, Vol. 10, No. 2, pp. 171 - 186

S. Fiori and F. Piazza (2000). A General Class of ψ -APEX PCA Neural Algorithms, IEEE Transactions on Circuits and Systems - Part I, Vol. 47, No. 9, pp. 1394 - 1398

S. Fiori (2000a). Blind Signal Processing by the Adaptive Activation Function Neurons, Neural Networks, Vol. 13, No. 6, pp. 597 - 611

S. Fiori (2000b). Blind Separation of Circularly Distributed Source Signals by the Neural Extended APEX Algorithm, Neurocomputing, Vol. 34, No. 1-4, pp. 239 - 252

S. Fiori and P. Bucciarelli (2001). Probability Density Estimation Using Adaptive Activation Function Neurons, Neural Processing Letters, Vol. 13 No. 1, pp. 31 - 42

S. Fiori (2001a). A Theory for Learning by Weight Flow on Stiefel-Grassman Manifold, Neural Computation, Vol. 13, No. 7, pp. 1625 - 1647

S. Fiori (2001b). On Blind Separation of Complex-Valued Sources by Extended Hebbian Learning, IEEE Signal Processing Letters, Vol. 8, No. 8, pp. 217 - 220

S. Fiori (2002a), Complex-Weighted One-Unit 'Rigid-Bodies' Learning Rule for Independent Component Analysis, Neural Processing Letters, Vol. 15, No. 3, pp. 275 - 282

S. Fiori (2002b). A Theory for Learning Based on Rigid Bodies Dynamics, IEEE Trans. on Neural Networks, Vol. 13, No. 3, pp. 521 - 531

S. Fiori (2002c). Hybrid Independent Component Analysis by Adaptive LUT Neurons, Neural Networks, Vol. 15, No. 1, pp. 85 - 94

S. Fiori (2002d). Unsupervised Neural Learning on Lie Group, International Journal of Neural Systems. Accepted for publication

S. Fiori, F. Grimani, P. Burrascano (2002). Novel Neural Network Feature

Selection Procedure by Generalization Maximization with Application to Automatic Robot Guidance, International Journal of Smart Engineering System Design, 4(2), pp. 91 - 106

S. Fiori and E. Celledoni (2002). Neural Learning by Geometric Integration of Reduced 'Rigid-Body' Equations. Unpublished Technical Report of the Center for Advanced Study, Oslo, Norway

S. Fiori and R. Rossi (2002). Rigid-Body Learning Theory in the 3×3 -Network Case: Improvements by the Rodriguez Formula for the Exponential Map. Unpublished Technical Report of the University of Perugia, Italy

J. Friedman and J.W. Tukey (1974). A projection pursuit algorithm for exploratory data analysis, IEEE Trans. on Computers, Vol. C-23, No. 9, pp. 881 - 889

E. Frulloni and S. Fiori (2002). Application of Neural Independent Component Analysis to ECT/NDE Data Processing. Submitted to a Journal

X. Giannakopoulos, J. Karhunen, and E. Oja (1999). An Experimental Comparison of Neural Algorithms for Independent Component Analysis and Blind Separation, International Journal of Neural Systems, Vol. 9, No. 2, pp. 99 - 114

M. Girolami and C. Fyfe (1997). Extraction of independent signal sources using a deflationary exploratory projection pursuit network with lateral inhibition, IEE Proceedings - Vision, Image and Signal Processing, Vol. 14, No. 5, pp. 299 - 306

R.M. Goldstein and H.A. Zebker (1987). Interferometric radar measurement of ocean surface currents, Nature, Vol. 328, pp. 707 - 709

L.C. Graham (1974). Synthetic interferometer radar for topographic mapping, Proceedings of the IEEE, Vol. 62, No. 6, pp.763 - 768

M. Gustaffson (1998). Gaussian Mixture and Kernel Based Approach to Blind Source Separation Using Neural Networks, Proc. of International Conference on Artificial Neural Networks, Vol. 2, pp. 869 - 874

E. Hairer, C. Lubich, and G. Wanner (2002). Geometric Numerical Integration, Springer Series in Computationa Mathematics, Springer

K.A. Hogda, T. Guneriussen, and I. Lauknes (2002). Synthetic aperture radar for DEM generation in snow-covered mountain terrain, 2002 IEEE International Geoscience and Remote Sensing Symposium (IGARSS '02), Vol. 4, pp. 2193 - 2195 J. Horstmann, W. Koch, S. Lehner, and R. Tonboe (2000). Wind retrieval over the ocean using synthetic aperture radar with C-band HH polarization, IEEE Transactions on Geoscience and Remote Sensing, Vol. 38, No. 5, pp. 2122 - 2131

P.J. Huber (1985). Projection pursuit, Annals of Statistics, Vol. 13, No. 2, pp. 435 - 475

A. Hyvärinen, J. Karhunen, and E. Oja (2001). Independent Component Analysis, John Wiley & Sons

A. Hyvärinen and E. Oja (1997). A fast fixed-point algorithm for Independent Component Analysis, Neural Computation, Vol. 9, No. 7, pp. 1483 -1492

A. Hyvärinen and E. Oja (1998). Independent Component Analysis by General Non-Linear Hebbian-Like Rules, Signal Processing, Vol. 64, No. 3, pp. 301
- 313

A. Hyvärinen and E. Oja (2000). Independent component analysis: algorithms and applications, Neural Networks, 13, 411 - 430

J.K. Jao (2001). Theory of synthetic aperture radar imaging of a moving target, IEEE Transactions on Geoscience and Remote Sensing, Vol. 39, No. 9, pp. 1984 - 1992

C. Jutten and J. Hérault (1988). Independent Component Analysis Versus Principal Component Analysis, Proc. of European Symposium on Signal Processing (EUSIPCO), Vol. 2, pp. 643 - 646

C. Jutten and J. Hérault (1991). Blind Separation of Sources, Part I: An Adaptive Algorithm Based on Neuromimetic Architecture, Signal Processing, Vol. 24, pp. 1 - 10

A.M. Kagan, Y.V. Linnik, and C.R. Rao (1973). Characterization problems in mathematical statistics. J. Wiley.

J. Karhunen, A. Hyvärinen, R. Vigário, J. Hurri, and E. Oja (1997). Applications of Neural Blind Separation to Signal and Image processing. Proc. International Conference on Acoustics, Speech and Signal Processing (IEEE-ICASSP), Vol. 1, pp. 131 - 134

J. Karhunen, E. Oja, L. Wang, R. Vigário, and J. Joutsensalo (1997). A class of neural networks for independent component analysis, IEEE Trans. on Neural Networks, Vol. 8, No. 3, pp. 486 - 504

K.H. Knuth (1998). Bayesian source separation and localization, In A.

Mohammad-Djafari (Ed.), SPIE'98 Proceedings: Bayesian Inference for Inverse Problems, San Diego, pp. 147 - 158

D. Kundur and D. Hatzinakos (1996). Blind image deconvolution. IEEE Signal Processing Magazine, pp. 43 - 64.

T.-W. Lee (1998). Independent Component Analysis - Theory and Practice, Kluwer Academic Publisher

J.J. Legarsky, S.P. Gogineni, and T.L. Akins (2001). Focused synthetic aperture radar processing of ice-sounder data collected over the Greenland ice sheet, IEEE Transactions on Geoscience and Remote Sensing, Vol. 39, No. 10, pp. 2109 - 2117

I.-I. Lin, W. Alpers, V. Khoo, H. Lim, T.K. Lim, and D. Kasilingam (2001). An ERS-1 synthetic aperture radar image of a tropical squall line compared with weather radar data, IEEE Transactions on Geoscience and Remote Sensing, Vol. 39, No. 5, pp. 937 - 945

R.-w. Liu (1996). Blind Signal Processing: An Introduction, Proc. of International Symposium on Circuits and Systems (IEEE-ISCAS), Vol. 2, pp. 81 - 84

G. Longo, C. Donalek, G. Raiconi, A. Staiano, R. Tagliaferri, S. Sessa, F. Pasian, R. Smareglia, and A. Volpicelli (2002). Data Mining of Large Astronomical Databases with Neural Tools, Proceeding of SPIE Meeting Astronomical Telescopes and Instrumentation. In press

G. Longo, R. Tagliaferri, S. Sessa, P. Ortiz, M. Capaccioli, A. Ciaramella, C. Donalek, G. Raiconi, A. Staiano, and A. Volpicelli (2001). Advanced Data Mining Tools for Exploring Large Astronomical Databases, Proceeding of SPIE Meeting N. 4477 on Astronomical Data Mining, SPIE, pp. 61 - 75

Y. Lou (2002). Review of the NASA/JPL airborne synthetic aperture radar system, 2002 IEEE International Geoscience and Remote Sensing Symposium (IGARSS '02), Vol. 3, pp. 1702 - 1704

D. Massonnet (1993). Validation of ERS1 Interferometry at CNES, Proceedings of the second ERS-1 symposium, Vol. 2, pp. 703 - 709, October 11-14, Hamburg, Germany

J.P. Nadal, N. Brunel, and N. Parga (1998). Nonlinear Feedforward Networks with Stochastic Inputs: Infomax Implies Redundancy Reduction, Network: Computation in Neural Systems, Vol. 9, No. 2

S. Odewahn, E. Stockwell, R. Pennington, L.Q. Humphreys, and W. Zu-

mach. Automated star/galaxy discrimination with neural networks, Astronomical Journal, Vol. 103(1), pp. 318 - 331

E. Oja (1997). The nonlinear PCA learning rule in independent component analysis, Neurocomputing, Vol. 17, pp. 25 - 45

L. Parra, K.-R. Mueller, C. Spence, A. Ziehe, and P. Sajda (2000). Unmixing Hyperspectral Data, Advances in Neural Information Processing Systems (NIPS*12), pp. 942 - 948

B.A. Pearlmutter and L.C. Parra (1996). Maximum Likelihood Blind Source Separation: A Context-Sensitive Generalization of ICA, Proc. of Neural Information Processing System (NIPS*9), M.M. Mozer, M.I. Jordan, and T. Petsche Eds., pp. 613 - 619

D. Pham, P. Garrat, and C. Jutten (1992). Separation of a Mixture of Independent Sources Through a Maximum Likelihood Approach, Proc. of European Signal Processing Conference, pp. 771 - 774

A. Prieto and B. Puntonet (1998). A neural learning algorithm for blind separation of sources based on geometric properties, Signal Processing, Vol. 64, No. 3, pp. 315 - 331

C.G. Puntonet and B. Prieto (1998). Neural net approach for blind separation of sources based on geometric properties, Neurocomputing, Vol. 18, No. 1-3, pp. 141 - 164

S.J. Roberts (1998). Independent Component Analysis: Source Assessment & Separation, a Bayesian Approach, IEE Proceedings - Vision, Image and Signal Processing, Vol. 145, No. 3, pp. 149 - 154

A.E.E. Rogers, and R.P. Ingalls (1969). Venus: Mapping the surface reflectivity by radar interferometry, Science, Vol. 165, pp. 797 - 799

Z. Roth and Y. Baram (1996). Multidimensional Density Shaping by Sigmoids, IEEE Trans. on Neural Networks, Vol. 7, No. 5, pp. 1291 - 1298

B. Sagi, S.C. Nemat-Nasser, R. Kerr, R. Hayek, C. Downing, and R. Hecht-Nielsen (2001). A Biologically Motivated Solution to the Cocktail Party Problem, Neural Computation, Vol. 13, No. 7, 1575 – 1602, July 2001

J. Schulz-Stellenfleth, J. Horstmann, S. Lehner, and W. Rosenthal (2001). Sea surface imaging with an across-track interferometric synthetic aperture radar: The SINEWAVE experiment, IEEE Transactions on Geoscience and Remote Sensing, Vol. 39, No. 9, pp. 2017 - 2028

E. Seifritz, F. Esposito, F. Hennel, H. Mustovic, J.G. Neuhoff, D. Bilecen,

G. Tedeschi, K. Scheffler, and F. Di Salle (2002). Spatiotemporal Pattern of Neural Processing in the Human Auditory Cortex, Science, Vol. 297, pp. 1706
- 1708

R. Tagliaferri, A. Ciaramella, L. Milano, F. Barone, G. Longo (1999). Spectral Analysis of Stellar Light Curves by Means of Neural Networks, Astronomy and Astrophysics Supplement Series, vol. 137, pp. 391 - 405

R. Tagliaferri, N. Pelosi, A. Ciaramella, G. Longo, F. Barone, and M. Milano (2001). Soft Computing Methodologies for Spectral Analysis in Cyclostratigraphy, Computers and Geosciences, Vol. 27, pp. 535 - 548

A. Taleb and C. Jutten (1997). Entropy Optimization - Application to Source Separation, Proc. of International Conference on Artificial Neural Networks, pp. 529 - 534

R. Thawonmas, A. Cichocki, and S.-i. Amari (1998). A Cascade Neural Network for Blind Signal Extraction without Spurious Equilibria, IEICE Trans. Fundamentals, Vol. E81-A, No. 9, pp. 1833 - 1846

M. Welling and M. Weber (2001). A Constrained E.M. Algorithm for Independent Component Analysis, Neural Computation, Vol. 13, No. 3, pp. 677 -689

L. Xu, C.C. Cheung, H.H. Yang, and S.-i. Amari (1998). Independent Component Analysis by the Information-Theoretic approach with Mixture of Densities, Proc. International Joint Conference on Neural Networks, (IEEE-IJCNN), pp. 1821 - 1826

L. Xu, C.C. Cheung, and S.-i. Amari (1998). Learned Parametric Mixture Based ICA Algorithm, Neurocomputing (Special issue on Independence and Artificial Neural Networks), Vol. 22, No. 1 - 3, pp. 69 - 80

H.H. Yang and S.-i. Amari (1997). Adaptive Online Learning Algorithms for Blind Separation: Maximum Entropy and Minimum Mutual Information, Neural Computation, Vol. 9, pp. 1457 - 1482

M. Yoshioka and S. Omatu, Signal Separation Method Using Genetic Algorithm, Proc. International Joint Conference on Neural Networks (IJCNN'98), pp. 909 - 912

L. Zhong, C. Wicks, J. Power, D. Dzurisin, W. Thatcher, and T. Masterlark (2002). Interferometric synthetic aperture radar studies of alaska volcanoes, 2002 IEEE International Geoscience and Remote Sensing Symposium (IGARSS '02), Vol. 1, pp. 191 - 194