Fasong WANG[1,3], Linrang ZHANG[1], Rui LI[2], Yu ZHOU[1], Nan LIU[1]

National Key Laboratory for Radar Signal Processing, Xidian University, China (1),
School of Sciences, Henan University of Technology, China (2),
China Electronics Technology Group Corporation, the 27th Research Institute, China (3).

Multidimensional independent subspace analysis by natural gradient

***Abstract**. Multidimensional Independent Subspace Analysis (MISA) as an extended Independent Component Analysis (ICA) method has been considered. The general and detailed definition, existence, uniqueness, separability of the MISA model are given and the relationships between ICA and MISA are also discussed. The natural gradient separation algorithm and corresponding simulation results for MISA are constructed based on the maximum likelihood theory and natural gradient method.*

***Streszczenie.** W artykule zaprezentowano metodę MISA – multidimensional independent subspace analysis. Przedstawiono też metode IOCA – independent component analysis. Opracowano algorytm separacji – natural gradient separation algorithm. (**Wielowymiarowa analiza podprzestrzeni MISA wykorzystująca metodę naturalnego gradient**)*

Keywords: independent component analysis (ICA), blind signal separation (BSS), independent subspace analysis (ISA), natural gradient.
Słowa kluczowe: analiza ICA – independent component analysis, ślepa separacja, naturalny gradient

Introduction

Standard Blind Signal Separation (BSS) model and methods have been successfully applied to many areas of science [1, 2]. The basic model assumes that the observed signals are linear superpositions of underlying hidden source signals. Most of the BSS algorithms are based on the independent assumption of the source signals, and are called Independent Component Analysis (ICA). However, the independence property of sources may not hold in some real-world situations, especially in biomedical signal processing and image processing, and therefore the standard ICA cannot give the expected results. Some techniques have been developed in recent years that relax the assumptions of basic ICA model and generalize the ICA problem. Among many extensions of the basic ICA model, several researchers have studied the case where the source signals are not statistically independent. Related models are generally recognized as dependent component analysis (DCA) model. Based on this basic extension of the ICA model, lots of DCA models and corresponding algorithms emerged [3-18].

Applications in which only certain groups of sources are independent may be highly relevant in practice. In this case, the independent sources can be multidimensional. The separation task requires an extension of ICA, which can be called independent subspace analysis (ISA) [3] or multi-dimensional independent component analysis (MICA) [4] or group ICA [5]. Throughout the paper, we will call it multi-dimensional ISA (MISA) model. MISA is a novel BSS model where ICA is incorporated with the idea of invariant feature subspaces [3]. In contrast to ordinary ICA, MISA does not assume that all sources are mutually independent. Instead, it assumes that the sources can be divided into couples, triplets, or in general *i*-tuples, such that the source signals inside a given *i*-tuple may be dependent on each other, but dependencies among different *i*-tuples are not allowed.

Various methods have been proposed to develop MISA algorithms [3-9,16]. MISA-related theoretical problems concern mostly the estimation of the entropy or of the mutual information. For this purpose, the *k*-nearest neighbors [6], FastISA methods [7], and the relative gradient methods [8] can be applied. Other recent approaches seeking independent subspaces via kernel methods [9] or joint block diagonalization [16] are also constructed.

In this paper, we present a new perspective of MISA method for BSS. After a general description of the MISA model, we discuss in detail the definition of MISA and the relationship between ICA and MISA from models to algorithms. As the solutions to MISA problem are not unique without extra constraints [15], we also discuss in detail the separateness and uniqueness of the MISA models and the corresponding separation theorem. Furthermore, based on the maximum likelihood theory and natural gradient method, the natural gradient separation algorithm for MISA model is constructed. Simulation result shows that the proposed algorithm is able to separate the MISA mixed source signals.

Basic MISA Model

The idea of MISA is that we do not require full independence of transform

(1) $\mathbf{y}(t) = \mathbf{W}\mathbf{x}(t)$,

but mutual independence of certain tuples $(y_{i_1}^{(i)}, \cdots, y_{i_m}^{(i)})$, $i = 1, \cdots, d$, and

(2) $\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t)$,

where $\mathbf{s}(t) = (s_1(t), \cdots, s_N(t))^{\mathrm{T}}$ is an unknown source vector, and $s_i(t)$, $i = 1, 2, \cdots, N$ can be random variables or time series. The mixtures $\mathbf{x}(t) = (x_1(t), \cdots, x_M(t))^{\mathrm{T}}$ are called sensor outputs. $\mathbf{n}(t) = (n_1(t), \cdots, n_M(t))^{\mathrm{T}}$ is a vector of additive noise. Matrix $\mathbf{A} = [a_{ij}] \in \mathbf{R}^{M \times N}$ is an unknown non-singular mixing matrix; no particular assumptions on the mixing coefficients are required. If the size of all tuples is restricted to one, the model becomes a general ICA model. In general, the tuples could have different sizes, but for the sake of simplicity, we assume that all tuples have the same size m (we define it as regular MISA) and that dm sources are to be extracted from equally numbered mixtures. If the model has one tuple only and components in the tuple are statistically dependent, it becomes the general DCA model.

Hyvarinen and Hoyer [3] first presented the regular MISA model by combining the principle of invariant feature subspace analysis, where the dependence within a *k*-tuple is explicitly modeled enabling the authors to propose better algorithms without having to resort to the problematic multidimensional density estimation.

In the case of regular MISA, it is assumed that the source random vector $\mathbf{s}$ is only *m*-independent (i.e.,

$(s_1^{(1)},\cdots,s_m^{(1)})^{\mathrm{T}}$, $\cdots$, $(s_{md-m+1}^{(d)},\cdots,s_{md}^{(d)})^{\mathrm{T}}$ are mutually independent). In the following, we give the MISA model definition and its indeterminacy. Cardoso [4], Theis [5], and Blanchard [19] generalized ICA to DCA from different views. In this paper, we define MISA consulting the definition of DCA of them.

Definition 1. A random vector $\mathbf{y}$ is called an independent component of the random vector $\mathbf{x}$, if there exists an invertible matrix $\mathbf{A}$ and a decomposition $\mathbf{x}=\mathbf{A}(\mathbf{y}_1,\cdots,\mathbf{y}_d)$, where $\mathbf{y}_i=(y_1^{(i)},\cdots,y_m^{(i)})^{\mathrm{T}}$, such that $\mathbf{y}_i$ and $\mathbf{y}_j$ are stochastically independent, $i,j=1,\cdots,d$.

Definition 2. Let $E_1,\cdots,E_d$ be d linear subspaces of $\mathbf{R}^{dm}$. They are recognized as linearly independent if any vector $\mathbf{x}$ of $E_1\oplus\cdots\oplus E_d$ admits of a unique decomposition as

(3) $$\mathbf{x}=\sum_{k=1}^{d}\mathbf{x}_k,$$

with $\mathbf{x}_k\in E_k$ for $1\le k\le d$. In such a case, the vectors $\mathbf{x}_1,\cdots,\mathbf{x}_d$ are called the linear components of $\mathbf{x}$ on the set $E_1,\cdots,E_d$.

Definition 3. A random vector $\mathbf{x}$ is defined as irreducible if it contains no lower-dimensional independent component.

Definition 4. An invertible matrix $\mathbf{W}$ is called a general MISA of random vector $\mathbf{x}$ if $\mathbf{Wx}=(\mathbf{s}_1,\cdots,\mathbf{s}_d)$ with pairwise independent, irreducible random vectors $\mathbf{s}_i$, $i=1,\cdots,d$, the sizes m_i of each $\mathbf{s}_i$ are arbitrary and satisfy $m_1+\cdots+m_d=N$. If all m_i are equals and $dm_i=N$, the separation is defined as regular MISA.

Theorem 1. (**Existence and Uniqueness of MISA**)[15-16]**.** Given a random vector $\mathbf{x}$ with existing covariance, an MISA of $\mathbf{x}$ exists and is unique except for permutation of components of the same dimension and invertible transformations within each independent component and within the Gaussian part.

Definition 5. The canonical MISA decomposition (if it exist) of a vector $\mathbf{x}$ is the unique MISA decomposition of $\mathbf{x}$ into $\mathbf{x}=\sum_{p=1}^{d}\mathbf{x}_d$ such that: 1) there is at most one Gaussian component; 2) no non-Gaussian component is irreducible.

In the ICA model, given the signals, sources $s_i(i=1,\cdots,N)$ can be recovered only up to sign, arbitrary scaling factors, and an arbitrary permutation. The MISA task has more freedom: signals $\mathbf{s}_i$ can be recovered up to an arbitrary permutation and an *m*-dimensional linear, invertible transformation. It is easy to see this by considering matrix $\mathbf{C}\in\mathbf{R}^{dm\times dm}$ made of a permutation matrix of size $d\times d$, where each element is made of an $m\times m$ block-matrix having invertible $\mathbf{C}_i$ blocks replacing the non-zero elements of the permutation matrix. Then,

(4) $$\mathbf{x}=\mathbf{As}=\mathbf{AC}^{-1}\mathbf{Cs},$$

and since $\mathbf{s}_i$ is independent of $\mathbf{s}_j$, $\mathbf{C}_i\mathbf{s}_i$ is independent of $\mathbf{C}_j\mathbf{s}_j$, $\forall i\neq j$. That is, in the MISA model, matrices $\mathbf{A}$ and $\mathbf{AC}^{-1}$, sources $\mathbf{s}_i$ and $\mathbf{C}_i\mathbf{s}_i$ are indistinguishable. This ambiguity of the MISA task can be lowered by assuming $E\{\mathbf{s}\}=0$, and $E\{\mathbf{ss}^{\mathrm{T}}\}=\mathbf{I}_{md}$, where $E\{\bullet\}$ is the expected value operator, $\mathbf{I}_{md}$ is the *md*-dimension identity matrix. Similarly, by scaling observed signals $\mathbf{x}$, one can assure that $E\{\mathbf{x}\}=0$ and $E\{\mathbf{xx}^{\mathrm{T}}\}=\mathbf{I}_{md}$, which is called the whitening of the inputs. Then,

(5) $$E\{\mathbf{xx}^{\mathrm{T}}\}=\mathbf{A}E\{\mathbf{ss}^{\mathrm{T}}\}\mathbf{A}^{\mathrm{T}}=\mathbf{AA}^{\mathrm{T}}=\mathbf{I}_{md}.$$

It follows that under our assumptions, signals $\mathbf{s}_i$ can be recovered up to permutation and *m*-dimensional orthogonal transformation in the MISA problem. In other words, if $\mathbf{C}_i\in\mathbf{R}^{m\times m}$ is an arbitrary orthogonal matrix, then signals $\mathbf{x}$ will not provide information whether the original sources correspond to $\mathbf{s}_i$ or, instead, to $\mathbf{C}_i\mathbf{s}_i$. For the 1-D case this is equivalent to the uncertainty that $\mathbf{C}_i=1$ or $\mathbf{C}_i=-1$. That is, in 1-D, the sign of $\mathbf{s}_i$ cannot be determined. Thus, without any loss of generality, it is satisfactory to restrict the search for mixing matrix $\mathbf{A}$ (or, for its inverse, i.e., for separation matrix $\mathbf{W}$) to the set of orthogonal matrices.

Definition 6. We call matrix $\mathbf{A}$ ***k*-admissible** if for each $r,s=1,\cdots,d$, the (r,s) sub-*k*-matrix of $\mathbf{A}$ is either invertible or zero. (Note that this is not a strong restriction.)

**Theorem 2. (Separability of MISA)[15]. ** Let $\mathbf{A}\in\mathbf{R}^{N\times N}$ and $\mathbf{s}$ be a *m*-independent *N*-dimensional random vector having no Gaussian *k*-tuple $(s_{rm},\cdots,s_{rm+m-1})^{\mathrm{T}}$. Assume that $\mathbf{A}$ is *m*-admissible. If $\mathbf{As}$ is again *m*-independent, then $\mathbf{A}$ is *m*-equivalent to the identity.

Note that for $m=1$, this is linear ICA separability because every matrix is 1-admissible. As a result, we give the MISA model by taking the multiplicative model partition of the entries of matrix and vectors as:

$$\mathbf{x}(t)=\mathbf{As}(t)\quad\text{or}$$

(6) $$\begin{bmatrix}\mathbf{x}_1(t)\\ \vdots\\ \mathbf{x}_d(t)\end{bmatrix}=\begin{bmatrix}\mathbf{A}_{11}&\cdots&\mathbf{A}_{1d}\\ \vdots&\ddots&\vdots\\ \mathbf{A}_{d1}&\cdots&\mathbf{A}_{dd}\end{bmatrix}\begin{bmatrix}\mathbf{s}_1(t)\\ \vdots\\ \mathbf{s}_d(t)\end{bmatrix},$$

where $\mathbf{x}_i$ and $\mathbf{s}_i$ of $\mathbf{x}$ and $\mathbf{s}$ are vectors of dimension *m* respectively and can be defined as $\mathbf{x}_i=(x_1^{(i)},\cdots,x_m^{(i)})$ and $\mathbf{s}_i=(s_1^{(i)},\cdots,s_m^{(i)})$ for $i=1,\cdots,d$. Partitioned matrix $\mathbf{A}$ is of size $dm\times dm$ since its entries $\mathbf{A}_{ij}$ are matrices of size $m\times m$. The following assumptions are needed: 1) Components $\mathbf{s}_i$ are vector-valued, non-Gaussian, mutually independent and of identity covariance; 2) Entries of each $\mathbf{s}_i$ are not independent and all are of equal dimension *m*; 3) Sample data is centered and whitened. The whole mixed-separation MISA process is described in Fig.1.

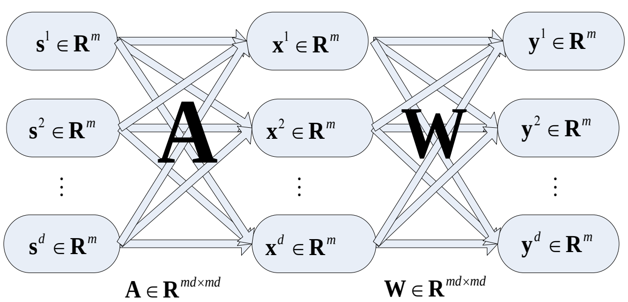


Fig.1. The whole mixed-separation MISA process.

In Fig.1, $\mathbf{s}\in\mathbf{R}^{dm}$ is the hidden independent subspaces; $\mathbf{x}\in\mathbf{R}^{dm}$ is the observation and $\mathbf{x}(t)=\mathbf{As}(t)$; $\mathbf{y}\in\mathbf{R}^{dm}$ is the estimated sources and $\mathbf{y}(t)=\mathbf{Wx}(t)$.

Proposed Algorithm for MISA

Based on the estimation of the entropy or of the mutual information, a number of methods have been proposed by various authors to develop MISA algorithms. In this section, we first give the performance index of the MISA algorithm, and subsequently the natural gradient based MISA algorithm is derived.

Performance Index of MISA Algorithm

If the MISA algorithm works properly, the product of the estimated separation matrix $\mathbf{W}$ and the original mixing matrix $\mathbf{A}$ produces a permutation matrix made of $m \times m$ blocks. The distance of $\mathbf{WA}$ and the block permutation matrix are measured by using a generalization of the Amari-distance [6, 16]. Let $\mathbf{P} = \mathbf{WA}$ and b_{ij} denote the sum of the absolute values of elements at the intersection of the $i(m-1)+1, \cdots, im$ rows and the $j(m-1)+1, \cdots, jm$ columns of matrix $\mathbf{P}$. Then the generalized Amari-distance *PI* is defined as follows:

(7)
$$pI(P) \triangleq \frac{1}{2d}\sum_{i=1}^{d}\left(\frac{\sum_{j=1}^{d}|b_{ij}|}{\max_j |b_{ij}|}-1\right) + \frac{1}{2d}\sum_{j=1}^{d}\left(\frac{\sum_{i=1}^{d}|b_{ij}|}{\max_i |b_{ij}|}-1\right) \geq 0$$

where $b_{ij} = \sum_{p=i(m-1)+1}^{im}\sum_{q=j(m-1)+1}^{jm}|\mathbf{P}_{pq}|$ for $1 \leq i, j \leq d$. Clearly, $PI(\mathbf{P}) \geq 0$ and it is zero if and only if matrix $\mathbf{P}$ is a permutation matrix permuting $m \times m$ block matrices.

Natural Gradient Learning Algorithm

For simplicity purpose, we consider here only the case where $M = N$ and $\mathbf{s}(t) \in \mathbf{R}^N$ are divided into *d* number of m-tuple (where m represents the dimension of subspace, i.e., $N = dm$). In this case, the data matrix given by $\mathbf{X} \in \mathbf{R}^{N\times T}$ and a linear transform that MISA seeks is given by $\mathbf{W} \in \mathbf{R}^{N\times N}$. We also assume an identical dimension, m, for every feature subspace.

MISA finds a linear transform $\mathbf{W}$ that maximizes the independence of the norms of the projection in linear subspaces. The cost function $\mathscr{J}(\mathbf{W},\mathbf{X})$ is taken as the negative normalized log-likelihood, which has the form

(8)
$$\mathscr{J}(\mathbf{W},\mathbf{X}) = -\frac{1}{T}\sum_{t=1}^{T}\sum_{j=1}^{m}\log p\left(\sum_{i\in\mathscr{F}_j} y_i^2(t)\right) - \log|\det(\mathbf{W})|.$$

We define

(9) $$\psi\Big(\sum_{i\in\mathscr{F}_j} y_i^2(t)\Big) = -\log p\Big(\sum_{i\in\mathscr{F}_j} y_i^2(t)\Big),$$

where the probability distribution $p(\bullet)$ is assumed to be super-Gaussian as in [3]. We also define $\boldsymbol{\xi}(t) = [\xi_1(t), \cdots, \xi_N(t)]^{\mathrm{T}}$, where $\xi_i(t) = \sum_{l\in\mathscr{F}_{j(i)}}(\mathbf{w}_l^{\mathrm{T}}\mathbf{x}(t))^2$ and $\mathscr{F}_{j(i)}$ is the feature subspace to which $\mathbf{w}_i$ belongs. Then, applying the gradient descent method, leads to the following updating rule for $\mathbf{W}$ [8]:

(10)
$$\Delta\mathbf{W} = \eta\frac{\partial\mathscr{J}(\mathbf{W},\mathbf{X})}{\partial\mathbf{W}} = \eta\left\{\mathbf{W}^{-\mathrm{T}} - \sum_{t=1}^{T}\{[\psi'(\boldsymbol{\xi}(t))\odot\mathbf{y}(t)]\mathbf{x}^{\mathrm{T}}(t)\}\right\}$$

where $\eta > 0$ is a learning rate, $\psi' = -p'/p$ (negative score function), i.e., $\psi'(\xi_i(t)) = \frac{1}{2}\alpha\xi_i^{-1/2}(t)$, and □ is the Hadamard product (which is the element-wise product). Define a matrix

(11) $$\mathbf{\Phi} = [\varphi(1)\cdots\varphi(T)] \in \mathbf{R}^{N\times T}$$

where $\boldsymbol{\varphi}(t) = \boldsymbol{\psi}'(\boldsymbol{\xi}(t))\odot\mathbf{y}(t)$. Then, (10) can be written in a compact form

(12) $$\Delta\mathbf{W} = \eta\left\{\mathbf{W}^{-\mathrm{T}} - \mathbf{\Phi}\mathbf{X}^{\mathrm{T}}\right\}.$$

In the case where the data matrix $\mathbf{X}$ is already whitened, the linear transform $\mathbf{W}$ is constrained to be an orthogonal matrix. The cost function is hence simplified as

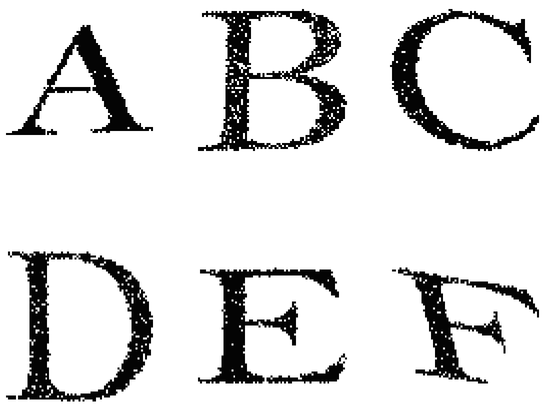
(13) $$\mathscr{J}(\mathbf{W},\mathbf{X}) = -\frac{1}{T}\sum_{t=1}^{T}\sum_{j=1}^{J}\log p\left(\sum_{i\in\mathscr{F}_j} y_i^2(t)\right).$$

The associated gradient descent algorithm is also in a simpler form $\Delta\mathbf{W} = -\eta\mathbf{\Phi}\mathbf{X}^{\mathrm{T}}$, which was originally proposed by [4]. Based on the theory of Riemann manifold and the natural gradient method of ICA, we can derive the natural gradient based MISA algorithm:

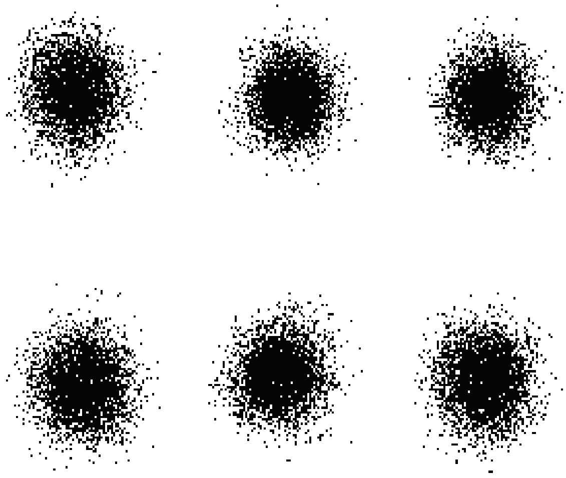
(14) $$\Delta\mathbf{W} = \eta(k)\left\{\mathbf{I} - \mathbf{\Phi}^{(k)}[\mathbf{Y}^{(k)}]^{\mathrm{T}}\right\}\mathbf{W}^{(k)}.$$

Simulations

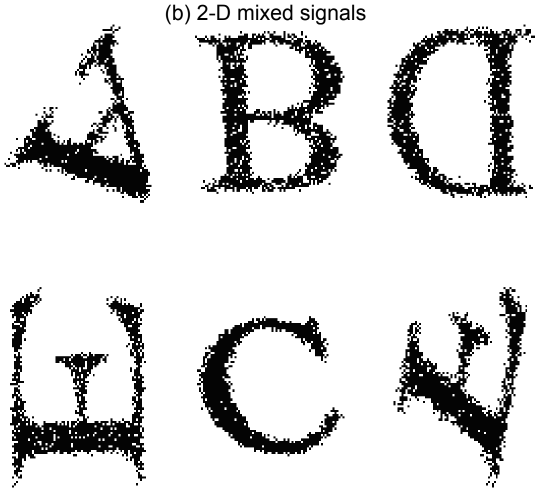
In this section, we will give some simulations to verify the efficiency of the proposed natural gradient based MISA algorithm. 6 pieces of 2-dimensional independent sources are chosen, none of them are linearly separable in 2-dimensional spaces. For the sake of visualization, sources formed simple 2-D patterns. 2-D samples are generated form letters, alike in Fig.2 (a). Random matrix of dimension 12×12 is used to mix the sources. 6 pieces of 2-D projections of the mixed sources are shown in Fig.2 (b). The proposed algorithm is applied to the mixed signals. Results of the separation are shown in Fig.2 (c). The MISA algorithm could recover the sources up to permutation and the directions within the subspaces. This feature is illustrated in Fig.2 (d) by the product of the true mixing matrix and the estimated separation matrix. This matrix is close to a permutation matrix made of 2×2-sized blocks as expected. After convergence, the Amari index equals to 0.0154.



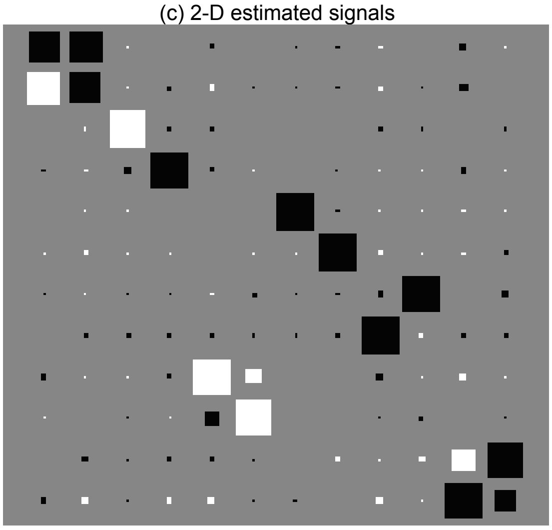
(a) 2-D source signals



(b) 2-D mixed signals



(c) 2-D estimated signals



(d) 2-dimensional performance matrix

Fig.2. MISA results for 2-D data.(a)2-D source signals; (b) 2-D mixed signals; (c) 2-D estimated signals;(d) 2-D performance matrix

Conclusions

We discussed the MISA definition and the relationships between ICA and MISA from models to algorithms. Moreover, the separateness and uniqueness of the MISA models have been discussed in detail and the corresponding separation theorems are also derived. Then, the natural gradient based MISA algorithms were derived. Simulation results showed that the proposed algorithm can work well.

We would like to thank the anonymous reviewers for their constructive comments which improved the original manuscript. This research is financially supported by the National Natural Science Foundation of China (No. 61001213), the National Defense Basic Scientific Research program of China (No. B1420110193).

REFERENCES

[1] Comon P., Jutten C., *Handbook of Blind Source Separation: Independent Component Analysis and Applications*. Elsevier, Oxford 2010.

[2] Cichocki A., Amari S., *Adaptive Blind Signal and Image Processing: Learning Algorithms and Applications*. John Wiley & Sons, New York 2002.

[3] Hyvarinen A., Hoyer P., Emergence of phase and shift invariant features by decomposition of natural images into independent feature subspaces, *Neural Computation*, 12 (2000), No.7, 1705-1720.

[4] Cardoso J.F., Multidimensional independent component analysis, *Proc. of the IEEE 1998 Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP'98)*, 1998, 1941-1944, Seattle, WA, USA.

[5] Theis F.J., Blind signal separation into groups of dependent signals using joint block diagonalization, *Proc. ISCAS 2005*, 2005, Kobe, Japan.

[6] Poczos B., Lorincz A., Independent Subspace Analysis Using k-Nearest Neighborhood Distances, *Lecture Notes in Computer Science*, 3697 (2005), 163-168.

[7] Hyvarinen A., Koster U., FastISA: A fast fixed-point algorithm for independent subspace analysis, *Proc. European Symposium on Artificial Neural Networks*, 2006, Bruges, Belgium.

[8] Choi H., Choi S., Relative Gradient Learning for Independent Subspace Analysis, *Proc. Int. Joint Conf. Neural Networks (IJCNN)*, 2006, Vancouver, BC, Canada, 3919-3924.

[9] Bach F.R., Jordan M.I., Kernel Independent Component Analysis. *Journal of Machine Learning Research*, 3 (2002), 1-48.

[10] Caiafa C.F., Proto A.N., Separation of statistically dependent sources using an L^2-distance non-Gaussianity measure. *Signal Processing*, 86 (2006), No.11, 3404-3420.

[11] Wang F.S., Li H.W., Li R., Novel NonGaussianity Measure Based BSS Algorithm for Dependent Signals, *Lecture Notes in Computer Science*, 4505 (2007), 837-844.

[12] Caiafa C.F., Salerno E., Protoa A.N., Fiumi L., Blind spectral unmixing by local maximization of non-Gaussianity. *Signal Processing*, 88 (2008), No.1, 50-68.

[13] Zhang K., Chan L.W., An Adaptive Method for Subband Decomposition ICA. *Neural Computation*, 18 (2006), No.1, 191-223.

[14] Abrard F., Deville Y., A time-frequency blind signal separation method applicable to underdetermined mixtures of dependent sources, *Signal Processing*, 85 (2005), No.7, 1389-1403.

[15] Theis F.J., Uniqueness of complex and multidimensional independent component analysis, *Signal Processing*, 84 (2004), No.5, 951-956.

[16] Theis F.J., Towards a general independent subspace analysis, *Proc. NIPS*, 2007.

[17] Aghabozorgi M.R., Doost-Hoseini A.M., Blind separation of jointly stationary correlated sources. *Signal Processing*, 84 (2004), No.2, 317-325.

[18] Li H., Shen Y.-H., Wang J.-G., Online Blind Separation of Dependent Sources Using Nonnegative Matrix Factorization Based on KL Divergence. *Przeglad Elektrotchniczny*, 88 (2012), No.1b, 17-20.

[19] Blanchard G., Kawanabe M., Sugiyama M., Spokoiny V., Muller K.R., In search of non-gaussian components of a high-dimensional distribution. *Journal of Machine Learning Research*, 7 (2006), 247-282.

***Authors**: Dr. Fasong Wang, Prof. Lirang Zhang, Dr. Yu Zhou, Dr. Nan Liu, National Key Laboratory for Radar Signal Processing, Xidian University, Xi'an 710071, China, E-mail: fasongwang@126.com; Dr. Rui Li, School of Sciences, Henan University of Technology, Zhengzhou 450001, China, E-mail: liruilyric@126.com.*

The correspondence address is:

e-mail:*fasongwang@126.com*