Computers and Mathematics with Applications 62 (2011) 3997–4006

Contents lists available at SciVerse ScienceDirect

Computers and Mathematics with Applications

journal homepage: www.elsevier.com/locate/camwa

Maximum contrast analysis for nonnegative blind source separation

Zuyuan Yang [a,*], Yong Xiang [b], Shengli Xie [c]

[a] School of Electronic and Information Engineering, South China University of Technology, Guangzhou, 510641, China

[b] School of Information Technology, Deakin University, Burwood Campus, Melbourne, VIC 3125, Australia

[c] Faculty of Automation, Guangdong University of Technology, Guangzhou, 510006, China

ARTICLE INFO

Article history:
Received 8 May 2011
Accepted 2 September 2011

Keywords:
Nonnegative blind source separation
Maximum contrast analysis
Iterative determinant maximization

ABSTRACT

In this paper, we propose a maximum contrast analysis (MCA) method for nonnegative blind source separation, where both the mixing matrix and the source signals are nonnegative. We first show that the contrast degree of the source signals is greater than that of the mixed signals. Motivated by this observation, we propose an MCA-based cost function. It is further shown that the separation matrix can be obtained by maximizing the proposed cost function. Then we derive an iterative determinant maximization algorithm for estimating the separation matrix. In the case of two sources, a closed-form solution exists and is derived. Unlike most existing blind source separation methods, the proposed MCA method needs neither the independence assumption, nor the sparseness requirement of the sources. The effectiveness of the new method is illustrated by experiments using X-ray images, remote sensing images, infrared spectral images, and real-world fluorescence microscopy images.

1. Introduction

Blind source separation (BSS) is an important problem in signal processing and computer application areas [1–4]. As BSS requires no or little prior knowledge of the sources, it can be used in a variety of applications such as speech identification [5], biomedical image processing [6], and remote sensing image interpretation [7]. So far, various methods have been developed for performing BSS, including independent component analysis (ICA) [2], sparse component analysis (SCA) [8], and nonnegative component analysis [9].

ICA is one of the traditional BSS approaches, aiming to recover the independent components from the observed mixtures. FastICA is a typical ICA method [10], and a detailed review of the ICA-based methods can be found in [11]. While ICA plays an important role in solving the problem of BSS and brings innovations to signal processing in both theory and applications [12,13], it also suffers from the independence condition imposed on the sources. To relax this condition, some second-order statistics methods are developed, such as slow feature analysis and temporal predictability analysis [14,15]. However, these methods are based on the assumption that the sources are uncorrelated.

To separate the correlative sources, the sparseness feature of the sources can be exploited. Recently, SCA for BSS has been intensively studied; it yields a number of BSS methods including the improved M-FOCUSS method [16], the morphological diversity-based method [17], and the spectral clustering method [18]. Besides, there is also considerable theoretical analysis of various aspects of SCA, such as identifiability analysis [8], performance analysis [19], and probability estimation for recoverability analysis [20]. The SCA-based methods do not require prior knowledge of the statistical features of the sources, but they often require the source signals to be sparse [21].

* Corresponding author. Tel.: +86 61 3 52271268; fax: +86 61 3 52272167.
E-mail addresses: zyuan.yang@mail.scut.edu.cn (Z. Yang), yxiang@deakin.edu.au (Y. Xiang), eeoshlxie@scut.edu.cn (S. Xie).

0898-1221/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.camwa.2011.09.003

Another approach to blind separation of correlative sources is the so-called nonnegative BSS (NBSS), where both the mixing matrix and the sources are nonnegative [9]. The nonnegative assumption holds in many practical applications, e.g., spectral unmixing in remote sensing image processing [22], and tumor feature analysis in biomedical image processing [6]. A typical method for processing nonnegative signals is the nonnegative matrix factorization (NMF) which aims to decompose a nonnegative matrix into two nonnegative factor matrices [23]. To improve the interpretability of the results, there were also developed some constrained NMF algorithms, such as non-smooth NMF (nsNMF) [24] and piecewise smooth NMF with a sparseness constraint (PSNMFSC) [25]. However, they need the prior knowledge of the constraints which is difficult to obtain. For instance, nsNMF requires the approximate smoothness degrees of the sources and the mixing matrix, and PSNMFSC needs the exact sparseness degrees of the sources.

More recently, the high contrast feature of the sources was exploited to achieve NBSS. This feature was originally found in processing biomedical images, and it brings new insight to the study of BSS [3,26]. In this paper, we further analyze the contrast feature of the sources and propose a maximum contrast analysis (MCA) method for NBSS. First, a joint contrast degree (JCD) measure is developed, based on the volume of the convex hull constructed by projected signal samples plus the origin of coordinates. Note that the constructed convex hull is based on the projections of column vectors of the signal matrix, instead of the traditional row vectors used in the algorithm nLCA-IVM in [6]. Exploiting this JCD measure, an MCA-based cost function related to determinant of the separation matrix is proposed. Our analysis shows that the separation matrix can be obtained by maximizing the proposed cost function. Then we develop an iterative determinant maximum (IDM) algorithm to optimize the cost function and thus find the separation matrix. Specifically, the columns of the separation matrix are iteratively estimated one by one, and the estimation of these columns is a linear programming (LP) problem at each iteration. Other than for the LP problems in the algorithms nLCA-IVM and CAMNS-LP in [3] where many inequality constraints are encountered, the inequality constraints used in IDM are much less. As a result, the proposed IDM algorithm is more efficient in computation. Compared with CAMNS-LP, IDM is more robust to the required pure-source sample condition, i.e., for each source, there exists at least one time instant at which the source is dominant. In addition, we derive a closed-form solution for the case of two sources. Further, in contrast to the traditional methods, our method does not impose independence or sparseness conditions on the sources. It should be noted that although the method in [27] can also deal with correlative sources, the source signals need to be precoded prior to transmission. So it can only be used in applications where the source signals are accessible, e.g., in communication systems.

The remainder of this paper is organized as follows. In Section 2, the MCA concept is introduced and the MCA-based cost function is proposed, together with the source identifiability analysis. The general IDM algorithm and the special closed-form solution are derived in Section 3. Section 4 presents extensive experimental results in order to illustrate the performance of the proposed method. Finally, Section 5 concludes the paper.

2. Maximum contrast analysis

We consider the following instantaneous mixing model:

$$\mathbf{X} = \mathbf{AS} \tag{1}$$

where $\mathbf{X} \in \Re^{m\times N}$ denotes the mixtures, $\mathbf{A} \in \Re^{m\times n}$ denotes the instantaneous mixing matrix, $\mathbf{S} \in \Re^{n\times N}$ represents the sources, and m, n, N denote the numbers of the mixed signals, the source signals, and the samples, respectively.

The corresponding unmixing model is

$$\mathbf{Y} = \mathbf{WX} \tag{2}$$

where $\mathbf{Y} \in \Re^{n\times N}$ denotes the recovered signals and $\mathbf{W} \in \Re^{n\times m}$ is the separation matrix.

Like other works concerning NBSS, ours assumes that $\mathbf{A}$ and $\mathbf{S}$ are nonnegative. Since there exists scale indeterminacy in BSS, we also assume that each column of $\mathbf{A}$ is, without loss of generality, a sum-to-1 column. Further, we assume for the sake of simplicity that each column of the mixture matrix $\mathbf{X}$ is nonzero. This can be easily satisfied by removing zero columns from $\mathbf{X}$. In the rest of the paper, we focus on the case where the numbers of sources and mixtures are the same, i.e., $m = n$. For the case of $m > n$, this can be simplified to the case of $m = n$ by using the nonnegative principal component analysis method [6].

The concept of the contrast feature of the signals originates from biomedical image analysis [26]; it will be further developed in this paper. Denote as $\mathbf{x}_i$, $\mathbf{x}^t$, and x_{it} the ith row, the tth column, and the element in the ith row and the tth column of $\mathbf{X}$, respectively. Then we define the contrast degree (CD) measure of $\mathbf{x}_i$, $i \in \{1, \ldots, m\}$ as

$$CD(\mathbf{x}_i) = \max_{t\in\{1,\ldots,N\}} \left\{ \max \left\{ \frac{x_{it} - \bar{\mathbf{x}}^t}{m\bar{\mathbf{x}}^t - \bar{\mathbf{x}}^t}, 0 \right\} \right\} \tag{3}$$

where $\bar{\mathbf{x}}^t$ denotes the mean of $\mathbf{x}^t$.

Clearly, it holds that $1 \geq CD(\mathbf{x}_i) \geq 0$, where $CD(\mathbf{x}_i) = 1$ if and only if there exists a pure-source sample for $\mathbf{x}_i$, i.e., there exists one time instant at which $\mathbf{x}_i$ dominates, and $CD(\mathbf{x}_i) = 0$ if and only if $\forall t$, x_{it} is not greater than the mean of the tth column of $\mathbf{X}$.

Since the CD measure does not depend on the L1-norms of the columns of $\mathbf{X}$, we project $\mathbf{X}$ into $\tilde{\mathbf{X}}$ on the hyperplane $\sum_{i=1}^{m} \tilde{z}_i = 1$, i.e., $\tilde{x}_{it} = x_{it} / \sum_{j=1}^{m} x_{jt},\ \forall t$, for the convenience of further analysis. Let $\Omega_{\tilde{\mathbf{X}}}$ be the convex hull formed by the columns of $\tilde{\mathbf{X}}$ plus the origin of coordinates, and let $\mathrm{vol}(\Omega_{\tilde{\mathbf{X}}})$ be the volume of $\Omega_{\tilde{\mathbf{X}}}$. If each $\tilde{\mathbf{x}}_i$ has a high contrast, then $\mathrm{vol}(\Omega_{\tilde{\mathbf{X}}})$ can be determined by using only a few samples which largely determine the CD of $\tilde{\mathbf{x}}_i$. Alternatively, a larger volume tends to correspond to a signal having higher CD. Thus, we define the following JCD measure for $\tilde{\mathbf{X}}$:

$$J(\tilde{\mathbf{X}}) = \mathrm{vol}(\Omega_{\tilde{\mathbf{X}}}) \times c \tag{4}$$

where $c = m!$ is a scale factor for normalizing $J(\tilde{\mathbf{X}})$.

Since $\mathrm{vol}(\Omega_{\tilde{\mathbf{X}}})$ interpolates smoothly between 0 and $1/m!$, $J(\tilde{\mathbf{X}})$ interpolates smoothly between 0 and 1. In particular, $J(\tilde{\mathbf{X}}) = 0$ if $CD(\tilde{\mathbf{x}}_i) = 0, \forall i$, i.e., all elements of $\tilde{\mathbf{X}}$ are the same. And $J(\tilde{\mathbf{X}}) = 1$ if and only if $CD(\tilde{\mathbf{x}}_i) = 1, \forall i$, i.e., there exists a pure-source sample for each signal in $\tilde{\mathbf{X}}$.

Let $[\tilde{\mathbf{S}}]_{it} = \tilde{s}_{it} = s_{it} / \sum_{j=1}^{m} x_{jt},\ \forall t$. It follows from (1) that

$$\tilde{\mathbf{X}} = \mathbf{A}\tilde{\mathbf{S}}. \tag{5}$$

As $\tilde{\mathbf{X}}$ and $\mathbf{A}$ satisfy the column sum-to-1 rule, $\tilde{\mathbf{S}}$ also satisfies the column sum-to-1 rule. Denote the JCD measure of $\tilde{\mathbf{S}}$ by $J(\tilde{\mathbf{S}})$. We have the following theorem.

Theorem 1. *The JCD measures of $\tilde{\mathbf{X}}$ and $\tilde{\mathbf{S}}$ satisfy $J(\tilde{\mathbf{X}}) \le J(\tilde{\mathbf{S}})$, and the equality holds if and only if $\tilde{\mathbf{X}} = \mathbf{P}\tilde{\mathbf{S}}$, where $\mathbf{P}$ is a permutation matrix.*

Proof. See Appendix A. □

As shown in Theorem 1, $\tilde{\mathbf{S}}$ has higher JCD measure than $\tilde{\mathbf{X}}$. This property can be exploited to propose an MCA-based cost function for NBSS.

Let $\tilde{\mathbf{Y}} = \mathbf{W}\tilde{\mathbf{X}}$ denote the estimate of $\tilde{\mathbf{S}}$; then $\tilde{\mathbf{Y}}$ should be forced to satisfy nonnegativity and the column sum-to-1 rule. Moreover, since $\tilde{\mathbf{S}}$ is of full row rank and $\mathbf{A}$ is of full rank, it follows from (5) that $\tilde{\mathbf{X}}$ is of full rank. Thus, there exists an $n \times n$ full rank sub-matrix, denoted by $\tilde{\mathbf{X}}_S$, in $\tilde{\mathbf{X}}$. Let $\tilde{\mathbf{Y}}_S = \mathbf{W}\tilde{\mathbf{X}}_S$ be a sub-matrix of $\tilde{\mathbf{Y}}$ corresponding to $\tilde{\mathbf{X}}_S$. The matrix $\mathbf{W} = \tilde{\mathbf{Y}}_S \tilde{\mathbf{X}}_S^{-1}$ should be forced to satisfy the column sum-to-1 rule as both $\tilde{\mathbf{X}}_S$ and $\tilde{\mathbf{Y}}_S$ satisfy the column sum-to-1 rule. On the other hand, like in the proof of Theorem 1 in Appendix A, one can obtain $J(\tilde{\mathbf{Y}}) = |\det(\mathbf{W})| J(\tilde{\mathbf{X}})$. Clearly $J(\tilde{\mathbf{X}})$ is a constant. Thus, on the basis of Theorem 1 and the above discussion, the MCA-based cost function with respect to $\mathbf{W}$ for NBSS is proposed as follows, together with the corresponding constraints:

$$\begin{aligned} &\text{Maximize: } D(\mathbf{W}) = |\det(\mathbf{W})| \\ &\text{s.t. } \sum_{i=1}^{n} w_{ij} = 1, \qquad \tilde{y}_{it} = [\mathbf{W}\tilde{\mathbf{X}}]_{it} \ge 0. \end{aligned} \tag{6}$$

Note that since $J(\tilde{\mathbf{X}})$ is a positive constant, it is not included in (6).

Let $\mathbf{e}_i$ denote the unit vector whose ith element is 1 where the other elements are all 0. For a nonnegative matrix $\tilde{\mathbf{X}}$ satisfying the column sum-to-1 rule, the points corresponding to the columns scatter in the bounded hyperplane $\sum_{i=1}^{m} \tilde{z}_i = 1, \tilde{z}_i \ge 0$. Intuitively, if the points of the set $\{\mathbf{e}_1, \ldots, \mathbf{e}_m\}$ belong to $\tilde{\mathbf{X}}$, the JCD measure of $\tilde{\mathbf{X}}$ defined in (4) will be the global maximum. Then, we have the following identifiability theorem.

Theorem 2. *Given that $\hat{\mathbf{W}}$ is the optimal solution of (6), if the sources satisfy the pure-source sample assumption, i.e., for each source, there exists at least one time instant at which the source is dominant, then*

$$\hat{\mathbf{W}}\mathbf{A} = \mathbf{P} \tag{7}$$

where $\mathbf{P}$ is a permutation matrix.

Proof. See Appendix B. □

Theorem 2 clearly shows that the separation matrix $\mathbf{W}$ can be obtained by optimizing the cost function (6).

3. Algorithm development

In this section, the IDM algorithm will be developed for optimizing the cost function in (6) and thus estimating the separation matrix $\mathbf{W}$. Moreover, a closed-form solution will be derived for the case of two sources. After the separation matrix $\mathbf{W}$ is estimated, the estimations of the sources can be easily obtained by $\mathbf{Y} = \mathbf{W}\mathbf{X}$.

3.1. The IDM algorithm

We first simplify the constraint in (6). Denoting the kth row of $\tilde{\mathbf{X}}$ by $\tilde{\mathbf{x}}_k$ and the kth column of $\mathbf{W}$ by $\mathbf{w}^k = [w_{1k}, \ldots, w_{nk}]^{\mathrm{T}}$, it follows that

$$\mathbf{W}\tilde{\mathbf{X}} = \sum_{k=1}^{n} \mathbf{w}^k \tilde{\mathbf{x}}_k = \sum_{k=1,k\neq j}^{n} \mathbf{w}^k \tilde{\mathbf{x}}_k + \mathbf{w}^j \tilde{\mathbf{x}}_j. \tag{8}$$

Let $\mathbf{C} = \sum_{k=1,k\neq j}^{n} \mathbf{w}^k \tilde{\mathbf{x}}_k$. Then, (8) can be rewritten in the following componentwise style:

$$[\mathbf{W}\tilde{\mathbf{X}}]_{it} = \left[\sum_{k=1,k\neq j}^{n} \mathbf{w}^k \tilde{\mathbf{x}}_k + \mathbf{w}^j \tilde{\mathbf{x}}_j\right]_{it} = c_{it} + [\mathbf{w}^j \tilde{\mathbf{x}}_j]_{it}. \tag{9}$$

Thus, $\tilde{y}_{it} = [\mathbf{W}\tilde{\mathbf{X}}]_{it} \geq 0$ is equivalent to $[\mathbf{w}^j \tilde{\mathbf{x}}_j]_{it} \geq -c_{it}$, i.e.,

$$w_{ij} \geq \max_t \left\{\frac{-c_{it}}{\tilde{x}_{jt}}\right\}. \tag{10}$$

On the other hand, it is well-known that

$$\det(\mathbf{W}) = \sum_{i=1}^{n} (-1)^{i+j} w_{ij} \det(\mathbf{W}_{ij}) \tag{11}$$

where $\mathbf{W}_{ij}$ is the sub-matrix of $\mathbf{W}$ with the ith row and jth column removed.

As $\mathbf{W}_{ij}$ is independent of $\mathbf{w}^j$, $\det(\mathbf{W})$ is a linear function with respect to $\mathbf{w}^j$. To utilize this feature, we iteratively optimize the columns of $\mathbf{W}$. Substituting (10) and (11) into (6), we obtain

$$\begin{aligned} &\text{Maximize: } g(\mathbf{w}^j) = \left|\sum_{i=1}^{n} (-1)^{i+j} w_{ij} \det(\mathbf{W}_{ij})\right| \\ &\text{s.t. } \sum_{i=1}^{n} w_{ij} = 1, \quad w_{ij} \geq \max_t \left\{\frac{-c_{it}}{\tilde{x}_{jt}}\right\}. \end{aligned} \tag{12}$$

To solve the cost function in (12), we decompose it into the following two LP problems:

$$\begin{aligned} &\text{Maximize: } g_1(\mathbf{w}^j) = \sum_{i=1}^{n} (-1)^{i+j} w_{ij} \det(\mathbf{W}_{ij}) \\ &\text{s.t. } \sum_{i=1}^{n} w_{ij} = 1, \quad w_{ij} \geq \max_t \left\{\frac{-c_{it}}{\tilde{x}_{jt}}\right\} \end{aligned} \tag{13}$$

and

$$\begin{aligned} &\text{Maximize: } g_2(\mathbf{w}^j) = -\sum_{i=1}^{n} (-1)^{i+j} w_{ij} \det(\mathbf{W}_{ij}) \\ &\text{s.t. } \sum_{i=1}^{n} w_{ij} = 1, \quad w_{ij} \geq \max_t \left\{\frac{-c_{it}}{\tilde{x}_{jt}}\right\}. \end{aligned} \tag{14}$$

Let $p = \max_{\mathbf{w}^j}\left\{g_1(\mathbf{w}^j)\right\}$, $q = \max_{\mathbf{w}^j}\left\{g_2(\mathbf{w}^j)\right\}$. Then the optimal solution $\hat{\mathbf{w}}^j$ of (12) is equivalent to the optimal solution of (13) if $|p| \geq |q|$ or the optimal solution of (14) if $|p| < |q|$. Furthermore, (13) and (14) can be solved directly by using the interior-point methods [28]. In this paper, the MATLAB LP toolbox is invoked. Here is the construction of the LP problem posed for (13) for reference: $\mathbf{x} = \mathbf{w}^j$, $f = -[(-1)^{1+j}\det(\mathbf{W}_{1j}), \ldots, (-1)^{n+j}\det(\mathbf{W}_{nj})]^{\mathrm{T}}$, $\mathbf{A}\mathrm{eq} = 1_n^{\mathrm{T}}$, $\mathrm{beq} = 1$, $\mathbf{A} = -\mathbf{I}_{n\times n}$, $\mathrm{b} = -\left[\max_t\left\{\frac{-c_{1t}}{\tilde{x}_{jt}}\right\}, \ldots, \max_t\left\{\frac{-c_{nt}}{\tilde{x}_{jt}}\right\}\right]^{\mathrm{T}}$.

In summary, the proposed IDM algorithm is formulated as follows:

1. *Projection: Project* $\mathbf{X}$ *into* $\tilde{\mathbf{X}}$ *on the hyperplane* $\sum_{i=1}^{m} \tilde{z}_i = 1$, *i.e.*, $\tilde{x}_{it} = x_{it} / \sum_{j=1}^{m} x_{jt}$, $\forall t$.
2. *Optimization: Optimize the LP problems* (13) *and* (14), *and let* $p = \max_{\mathbf{w}^j}\left\{g_1(\mathbf{w}^j)\right\}$, $q = \max_{\mathbf{w}^j}\left\{g_2(\mathbf{w}^j)\right\}$.
3. *Updating: Update the* jth *column* $\mathbf{w}^j$ *of* $\mathbf{W}$ *with the optimal solution of* (13) *if* $|p| \geq |q|$, *or the optimal solution of* (14) *if* $|p| < |q|$.
4. *Stop: If a stop criterion is satisfied, stop the algorithm; otherwise, set* $j = \begin{cases} j+1 \text{ if } j < n \\ 1 \text{ if } j = n \end{cases}$ *and go to the third step.*

In the last step of this algorithm, either the maximum iteration number or the relative error can be used as the stop criterion [6,10].

3.2. Closed-form solution for the case of $n = 2$

In this case, (6) can be rewritten as

$$\begin{aligned} &\text{Maximize: } D(\mathbf{W}) = |\det(\mathbf{W})| = |w_{11}w_{22} - w_{12}w_{21}| \\ &\text{s.t. } \begin{cases} w_{11} + w_{21} = 1 \\ w_{12} + w_{22} = 1 \\ \tilde{y}_{1t} = w_{11}\tilde{x}_{1t} + w_{12}\tilde{x}_{2t} \geq 0, \quad \forall t \\ \tilde{y}_{2t} = w_{21}\tilde{x}_{1t} + w_{22}\tilde{x}_{2t} \geq 0, \quad \forall t. \end{cases} \end{aligned} \tag{15}$$

Substituting the equality constraints into the cost function and the inequality constraints, (15) is simplified to

$$\begin{aligned} &\text{Maximize: } D(\mathbf{W}) = |w_{11} - w_{12}| \\ &\text{s.t. } \begin{cases} w_{11}\tilde{x}_{1t} + w_{12}\tilde{x}_{2t} \geq 0, \quad \forall t \\ w_{11}\tilde{x}_{1t} + w_{12}\tilde{x}_{2t} \leq 1, \quad \forall t. \end{cases} \end{aligned} \tag{16}$$

Since $\forall t$, $\tilde{x}_{1t} + \tilde{x}_{2t} = 1$, we can rewrite (16) as

$$\begin{aligned} &\text{Maximize: } D(\mathbf{W}) = |w_{11} - w_{12}| \\ &\text{s.t. } \begin{cases} (w_{11} - w_{12})\tilde{x}_{1t} + w_{12} \geq 0, \quad \forall t \\ (w_{11} - w_{12})\tilde{x}_{1t} + w_{12} \leq 1, \quad \forall t. \end{cases} \end{aligned} \tag{17}$$

The optimization of (17) can be considered in two scenarios.

Firstly, if $w_{11} \geq w_{12}$, then $D(\mathbf{W}) = w_{11} - w_{12}$, and the constraints in (17) can be written as

$$\begin{cases} \min\limits_t \left\{\tilde{x}_{1t}\right\}(w_{11} - w_{12}) + w_{12} \geq 0 \\ \max\limits_t \left\{\tilde{x}_{1t}\right\}(w_{11} - w_{12}) + w_{12} \leq 1. \end{cases} \tag{18}$$

Or equivalently

$$\begin{cases} \min\limits_t \left\{\tilde{x}_{1t}\right\} D(\mathbf{W}) + w_{12} \geq 0 \\ \max\limits_t \left\{\tilde{x}_{1t}\right\} D(\mathbf{W}) + w_{12} \leq 1. \end{cases} \tag{19}$$

From (19), it follows that

$$\frac{1 - w_{12}}{\max\limits_t \left\{\tilde{x}_{1t}\right\}} \geq D(\mathbf{W}) \geq \frac{-w_{12}}{\min\limits_t \left\{\tilde{x}_{1t}\right\}}.$$

Since both $\frac{1-w_{12}}{\max_t\{\tilde{x}_{1t}\}}$ and $\frac{-w_{12}}{\min_t\{\tilde{x}_{1t}\}}$ are monotone decreasing with respect to w_{12}, $D(\mathbf{W})$ reaches the maximum if and only if $\frac{1-w_{12}}{\max_t\{\tilde{x}_{1t}\}} = \frac{-w_{12}}{\min_t\{\tilde{x}_{1t}\}}$. And in this case, the maximum of $D(\mathbf{W})$ is $\frac{1}{\max_t\{\tilde{x}_{1t}\}-\min_t\{\tilde{x}_{1t}\}}$. Thus, for $w_{11} \geq w_{12}$, the optimal solution of (17) is

$$\begin{cases} \hat{w}_{11} = \dfrac{1 - \min\limits_t \left\{\tilde{x}_{1t}\right\}}{\max\limits_t \left\{\tilde{x}_{1t}\right\} - \min\limits_t \left\{\tilde{x}_{1t}\right\}} \\ \hat{w}_{12} = \dfrac{\min\limits_t \left\{\tilde{x}_{1t}\right\}}{\min\limits_t \left\{\tilde{x}_{1t}\right\} - \max\limits_t \left\{\tilde{x}_{1t}\right\}}. \end{cases} \tag{20}$$

Furthermore, since $\hat{w}_{11} + \hat{w}_{21} = 1$, $\hat{w}_{12} + \hat{w}_{22} = 1$, the optimal solution $\hat{\mathbf{W}}^*$ of (15) in the scenario of $w_{11} \geq w_{12}$ is

$$\hat{\mathbf{W}}^* = \begin{bmatrix} \dfrac{1 - \min\limits_t \left\{\tilde{x}_{1t}\right\}}{\max\limits_t \left\{\tilde{x}_{1t}\right\} - \min\limits_t \left\{\tilde{x}_{1t}\right\}} & \dfrac{\min\limits_t \left\{\tilde{x}_{1t}\right\}}{\min\limits_t \left\{\tilde{x}_{1t}\right\} - \max\limits_t \left\{\tilde{x}_{1t}\right\}} \\ \dfrac{\max\limits_t \left\{\tilde{x}_{1t}\right\} - 1}{\max\limits_t \left\{\tilde{x}_{1t}\right\} - \min\limits_t \left\{\tilde{x}_{1t}\right\}} & \dfrac{-\max\limits_t \left\{\tilde{x}_{1t}\right\}}{\min\limits_t \left\{\tilde{x}_{1t}\right\} - \max\limits_t \left\{\tilde{x}_{1t}\right\}} \end{bmatrix}. \tag{21}$$

Secondly, if $w_{11} < w_{12}$, then $D(\mathbf{W}) = w_{12} - w_{11}$ and the constraints in (17) can be written as

$$\begin{cases} \max\limits_t \left\{\tilde{x}_{1t}\right\}(w_{12} - w_{11}) - w_{12} \leq 0 \\ \min\limits_t \left\{\tilde{x}_{1t}\right\}(w_{12} - w_{11}) - w_{12} \geq -1. \end{cases} \tag{22}$$

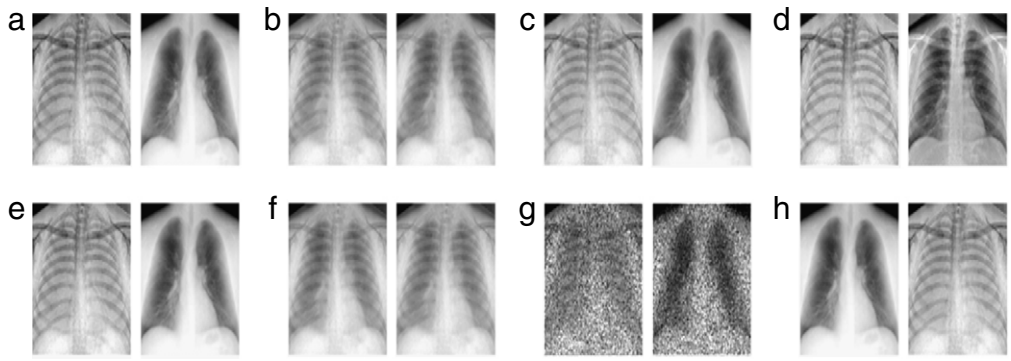

Fig. 1. X-ray source images, mixtures and recovered images obtained using different algorithms. (a) Two sources of bone and soft tissue. (b) Two mixtures. (c) Images recovered by the proposed method. (d) Images recovered by FastICA. (e) Images recovered by CAMNS-LP. (f) Images recovered by PSNMFSC. (g) Images recovered by nsNMF. (h) Images recovered by nLCA-IVM.

Similarly, one can find that the optimal solution $\hat{\mathbf{W}}^{**}$ of (15) in the scenario of $w_{11} < w_{12}$ is

$$\hat{\mathbf{W}}^{**} = \begin{bmatrix} \dfrac{\max\limits_t \{\tilde{x}_{1t}\} - 1}{\max\limits_t \{\tilde{x}_{1t}\} - \min\limits_t \{\tilde{x}_{1t}\}} & \dfrac{-\max\limits_t \{\tilde{x}_{1t}\}}{\min\limits_t \{\tilde{x}_{1t}\} - \max\limits_t \{\tilde{x}_{1t}\}} \\ \dfrac{1 - \min\limits_t \{\tilde{x}_{1t}\}}{\max\limits_t \{\tilde{x}_{1t}\} - \min\limits_t \{\tilde{x}_{1t}\}} & \dfrac{\min\limits_t \{\tilde{x}_{1t}\}}{\min\limits_t \{\tilde{x}_{1t}\} - \max\limits_t \{\tilde{x}_{1t}\}} \end{bmatrix}. \tag{23}$$

It can be seen from (21) and (23) that $\hat{\mathbf{W}}^{**}$ is the same as $\hat{\mathbf{W}}^{*}$ up to a row permutation. Since permutation is an intrinsic feature of BSS, either $\hat{\mathbf{W}}^{*}$ or $\hat{\mathbf{W}}^{**}$ can be chosen as the optimal closed-form solution of (15) (or equivalently (6)).

4. Experiments

In this section, four sets of experiments are carried out to evaluate the validity of the proposed MCA method. If there are more than two sources, the proposed IDM algorithm is used to estimate the separation matrix. In the case of two sources, the separation matrix is obtained from (21), which is a closed-form solution. The performance of our method is compared with those of CAMNS-LP, FastICA, nLCA-IVM, nsNMF, and PSNMFSC.

The accuracy of source separation is measured by the cross-correlation coefficient ρ between the source matrix $\mathbf{S}$ and its estimate $\hat{\mathbf{S}}$, which is defined as follows:

$$\rho = \frac{1}{n} \max_{\pi_i \in \Pi_n, c_i \in \{1,-1\}} \sum_{i=1}^{n} \frac{(\mathbf{s}_i - \mathbf{q}(\mathbf{s}_i))(c_i \hat{\mathbf{s}}_{\pi_i} - c_i \mathbf{q}(\hat{\mathbf{s}}_{\pi_i}))^{\mathrm{T}}}{\|\mathbf{s}_i - \mathbf{q}(\mathbf{s}_i)\| \cdot \left\| c_i \hat{\mathbf{s}}_{\pi_i} - c_i \mathbf{q}(\hat{\mathbf{s}}_{\pi_i}) \right\|}. \tag{24}$$

Here $\mathbf{s}_i$ denotes the ith source, $\mathbf{q}(\mathbf{s}_i)$ is an N-dimensional vector composed of the means of the $\mathbf{s}_i$, $\Pi_n = \{\pi = (\pi_1, \ldots, \pi_n) | \pi_i \in \{1, \ldots, n\}, \pi_i \neq \pi_j, \ \forall i \neq j\}$ is the set of all the permutations of $\{1, \ldots, n\}$, and c_i is the sign (or the polarity) ambiguity between the source $\mathbf{s}_i$ and its estimate $\hat{\mathbf{s}}_{\pi_i}$. Note that $\rho \in [0, 1]$ and the larger the value, the higher the source separation accuracy.

Furthermore, the running CPU time T is utilized to give an intuitive comparison of the computational costs of the algorithms. Each algorithm is implemented with MATLAB R2009a installed on a personal computer which is equipped with Intel(R) Celeron(R) CPU 2.4 GHz, 2 GB memory and Microsoft Windows 7 system.

4.1. Unmixing X-ray images

X-ray imaging is often used for the detection of lung cancers in medical diagnostics. However, since the ribs often overlap with soft issue, the observed images are mixtures of bone and soft tissue. In this experiment, the proposed MCA method is used to recover two source images shown in Fig. 1(a), which are taken from [29], from their mixtures. It can be found that the CD measures of these two sources are both 1, meaning that there are pure-source samples for each source. The averaged cross-correlation coefficient ρ_{ave} computed from fifty Monte Carlo runs with different random mixing matrices is used to show the source separation performance and the averaged CPU time T_{ave} is used to illustrate computational cost.

As shown in Table 1, the proposed method, CAMNS-LP and nLCA-IVM achieve perfect source separation. This is not surprising because the pure-source sample condition is satisfied in this experiment. On the other hand, since the two image sources are neither independent nor sufficiently sparse, the other three methods do not achieve perfect source separation.

Table 1
Indices ρ_{ave} and T_{ave} (s) of different algorithms in unmixing X-ray images.

	MCA	FastICA	CAMNS-LP	PSNMFSC	nsNMF	nLCA-IVM
ρ_{ave}	1.0000	0.9334	1.0000	0.8953	0.5007	1.0000
T_{ave}	0.0122	0.0834	1.5861	8.1709	2.2056	0.0166

Table 2
Indices ρ_{ave} and T_{ave} (s) of different algorithms in recovering remote sensing images.

	MCA	FastICA	CAMNS-LP	PSNMFSC	nsNMF	nLCA-IVM
ρ_{ave}	0.9950	0.7426	–	0.6180	0.6042	0.9951
T_{ave}	23.1666	4.6528	–	263.5825	115.4175	117.3956

Table 3
Indices T_{ave} (s) of different algorithms in decomposing infrared spectra under different SNR (dB) levels.

SNR	MCA	FastICA	CAMNS-LP	PSNMFSC	nsNMF	nLCA-IVM
15	0.3050	0.0362	2.7941	2.0094	0.7853	2.1075
20	0.2831	0.0225	2.0116	1.9259	0.7863	1.9997
25	0.2863	0.0278	2.9188	1.8928	0.7881	1.9700
30	0.2931	0.0294	3.6513	1.8916	0.7837	1.9244
35	0.2919	0.0259	6.5159	1.8875	0.7931	1.8763
40	0.2938	0.0278	4.4310	1.8875	0.7925	1.8300
45	0.2909	0.0266	5.2230	1.8919	0.7878	1.6925

Actually, the performance of nsNMF is poor in this experiment. It can also be seen from Table 1 that our method uses the least time to achieve BSS. Furthermore, we provide a visual comparison of these methods in Fig. 1. One can see that the images recovered by our method, CAMNS-LP and nLCA-IVM are the same as the original ones. In contrast, the images recovered by the other methods still contain certain levels of mixtures.

4.2. Interpreting remote sensing images

Remote sensing images are widely used in both industry and agriculture, especially for natural resource exploration in a large region. However, due to the restriction of the space resolution of the sensors, the collected images are often mixtures of the abundance maps of the end-members. This affects the correct interpretation of the collected images. In the experiment, seven mineral images from [22] are used as sources, which are the maps of sphene, kaolinite, montmorillonite, chalcedony, desert, alunite, and buddingtonite. The CD measures of these sources are 0.7376, 0.8704, 0.7517, 0.4335, 0.7526, 0.7682, and 0.5508, respectively. This implies that there is no pure-source sample for any source. We carry out fifty Monte Carlo runs with different random mixing matrices for all algorithms.

From Table 2, we can see that both the proposed MCA method and nLCA-IVM achieve satisfactory separation performance, though the pure-source sample condition is violated. Moreover, while the performance of our method is almost identical to that of nLCA-IVM, it only uses one fifth of the time that the latter utilizes to recover the sources. Compared with the performances of these two methods, those of FastICA, PSNMFSC and nsNMF are much poorer. In particular, due to the small CD measures of some sources, CAMNS-LP converged very slowly and could not yield an outcome after 24 h continuous running.

4.3. Decomposing infrared spectra with noise

Infrared spectra are often used to identify and quantify a solvent in the liquid or gas state. If there exist multiple solvents, the measured spectra are mixtures of their spectra. Also, the mixture spectra may be polluted by noise in practice. Decomposing these polluted mixture spectra is challenging, especially if the noise is strong. The noise level is measured by the signal-to-noise ratio (SNR) index defined as $\mathrm{SNR} = 10\log(E(\mathbf{x}\mathbf{x}^{\mathrm{T}})/E(\mathbf{y}\mathbf{y}^{\mathrm{T}}))\mathrm{dB}$, where $\mathbf{x}$ denotes the real signal, $\mathbf{y}$ denotes the corresponding noise, and $E()$ is the expectation operator. Three pure spectra (acetone, ethanol, and ethylene) with 1.929 cm^{-1} resolution and 4000 samples, which are obtained from a public database [30], are used to generate five mixtures contaminated by additive Gaussian noise. We consider seven SNR levels: 15, 20, 25, 30, 35, 40, and 45 dB. At a given SNR level, fifty Monte Carlo runs using different random mixing matrices are conducted for each BSS algorithm. At each run, we force the negative noisy samples to be zero to maintain the nonnegativity of the mixtures.

The indices ρ_{ave} and T_{ave} under different SNR levels are shown in Fig. 2 and Table 3, respectively. We can see from Fig. 2 that the proposed MCA method performs the best at high SNR. In the case of strong noise, it is comparable to nsNMF, but much better than CAMNS-LP, nLCA-IVM, and PSNMFSC. As regards computation time, MCA uses much less CPU time T_{ave} than the other algorithms except for FastICA.

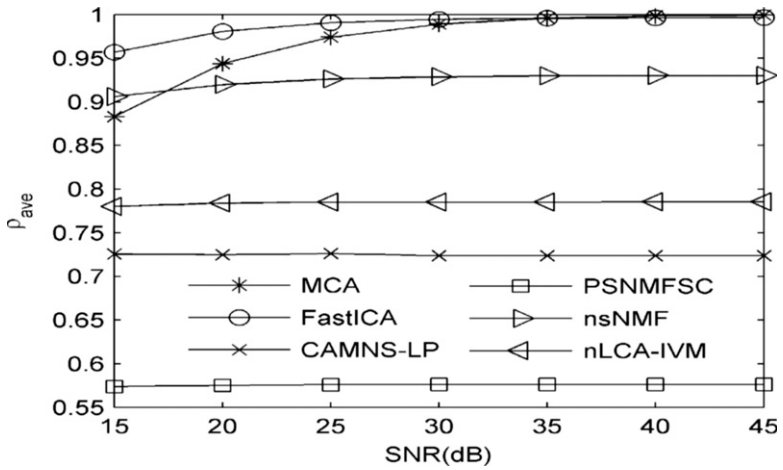

Fig. 2. Indices ρ_{ave} for the proposed method, FastICA, CAMNS-LP, PSNMFSC, nsNMF, and nLCA-IVM versus SNR.

4.4. Analyzing fluorescence microscopy images

Fluorescence microscopy images are often produced by multispectral optical sensor arrays and widely used in biomedical signal processing. Due to the problem of spectral overlap among the array probes, the observed images are often mixtures of spectral biomarkers. The decomposition of the mixtures can be treated as a BSS problem. In this experiment, we decompose the real-world newt lung cell images (see Fig. 3(a)) taken from [31] by using the proposed MCA method, FastICA, CAMNS-LP, PSNMFSC, nsNMF, and nLCA-IVM. The decomposed images are shown in Fig. 3(b)–(g). Comparing with the mixture cases in Fig. 3(a), we can see that the images recovered by the MCA method show clearer contours of the objects (see Fig. 3(b)).

5. Conclusion

It is shown in this paper that the contrast degree of source signals is larger than that of mixed signals. Motivated by this observation, we propose a novel determinant-based MCA cost function for nonnegative blind source separation, and show that the optimization of the cost function yields the separation matrix. On the basis of the proposed cost function, we then derive a general IDM algorithm for estimating the separation matrix. A closed-form solution is also derived for the case of two sources. Unlike most existing BSS methods, the proposed MCA method does not require the sources to be independent or sparse. Compared with the CAMNS-LP method and the nLCA-IVM method, our method is more efficient in computation. Experiments also show that the proposed MCA method is more robust to the pure-source sample condition than CAMNS-LP.

Acknowledgments

This work was supported by the Australian Research Council under grant DP0773446, the National Natural Science Foundation of China (No. 61104053, 61170193, 61103122), China Postdoctoral Science Foundation (No. 201104336, 20100480758), and the National Basic Research Program of China (937 Program, No. 2010CB731800).

Appendix A

Proof of Theorem 1. Decompose the convex hull $\Omega_{\tilde{\mathbf{X}}}$ formed by $\{\tilde{\mathbf{x}}_1, \ldots, \tilde{\mathbf{x}}_N, O_m\}$, i.e., the columns of $\tilde{\mathbf{X}}$ plus the origin of coordinates, into a series of simplexes $\Delta^1_{\tilde{\mathbf{X}}}, \ldots, \Delta^K_{\tilde{\mathbf{X}}}$ whose vertices belong to $\{\tilde{\mathbf{x}}_1, \ldots, \tilde{\mathbf{x}}_N, O_m\}$. Then, the volume of $\Omega_{\tilde{\mathbf{X}}}$ can be calculated as

$$\mathrm{vol}(\Omega_{\tilde{\mathbf{X}}}) = \sum_{i=1}^{K} \mathbf{vol}(\Delta^i_{\tilde{\mathbf{X}}}). \tag{25}$$

On the basis of the volume feature of the simplex [32] and the relation $\tilde{\mathbf{X}} = \mathbf{A}\tilde{\mathbf{S}}$, it holds that for $\forall i$,

$$\mathrm{vol}(\Delta^i_{\tilde{\mathbf{X}}}) = |\det(\mathbf{A})|\, \mathbf{vol}(\Delta^i_{\tilde{\mathbf{S}}}) \tag{26}$$

and $\sum_{i=1}^{K} \mathrm{vol}(\Delta^i_{\tilde{\mathbf{S}}}) = \mathbf{vol}(\Omega_{\tilde{\mathbf{S}}})$. Here $\Delta^i_{\tilde{\mathbf{S}}}$ is a simplex decomposed from the convex hull $\Omega_{\tilde{\mathbf{S}}}$ which is formed by $\tilde{\mathbf{S}}$ plus the origin of coordinates.

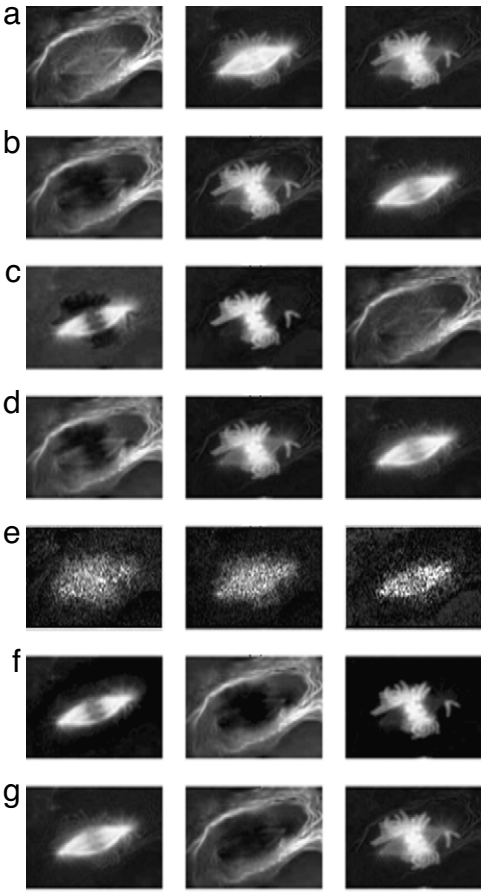

Fig. 3. Real-world newt lung cell images (mixtures), and images recovered using different BSS algorithms. (a) Three mixtures. (b) Images recovered by the proposed MCA method. (c) Images recovered by FastICA. (d) Images recovered by CAMNS-LP. (e) Images recovered by PSNMFSC. (f) Images recovered by nsNMF. (g) Images recovered by nLCA-IVM.

Moreover, since $\mathbf{A}$ is nonnegative and satisfies the column sum-to-1 rule, $|\det(\mathbf{A})| \leq 1$ and the equation holds if and only if $\mathbf{A}$ is a permutation matrix [6]. From (26), we obtain $\mathrm{vol}(\Delta^i_{\tilde{\mathbf{X}}}) \leq \mathbf{vol}(\Delta^i_{\tilde{\mathbf{S}}})$, $\forall i$. This inequality leads to $\mathrm{vol}(\Omega_{\tilde{\mathbf{X}}}) \leq \mathbf{vol}(\Omega_{\tilde{\mathbf{S}}})$ or $J(\tilde{\mathbf{X}}) \leq J(\tilde{\mathbf{S}})$, and the equality holds if and only if $\mathbf{A}$ is a permutation matrix, i.e., $\tilde{\mathbf{X}} = \mathbf{P}\tilde{\mathbf{S}}$. This completes the proof. □

Appendix B

Proof of Theorem 2. Since the sources satisfy the pure-source sample condition, there exists one permutation sub-matrix $\tilde{\mathbf{S}}_{\{i_1,i_2,\ldots,i_n\}}$ in $\tilde{\mathbf{S}}$, where $\{i_1, i_2, \ldots, i_n\}$ denotes its column index set in $\tilde{\mathbf{S}}$. On the other hand, $\hat{\mathbf{W}}$ is the optimal solution of (6) and $\hat{\mathbf{Y}} = \hat{\mathbf{W}}\tilde{\mathbf{X}}$. Then according to the interpretation of the joint contrast degree in Section 2, there exists at least one permutation sub-matrix $\hat{\mathbf{Y}}_{\{j_1,j_2,\ldots,j_n\}}$ in $\hat{\mathbf{Y}}$, where $\{j_1, j_2, \ldots, j_n\}$ denotes its column index set in $\hat{\mathbf{Y}}$. For example, if $\hat{\mathbf{W}} = \mathbf{A}^{-1}$, $\hat{\mathbf{Y}}_{\{j_1,j_2,\ldots,j_n\}} = \tilde{\mathbf{S}}_{\{i_1,i_2,\ldots,i_n\}}$ is a permutation matrix. Considering $\hat{\mathbf{Y}} = \hat{\mathbf{W}}\tilde{\mathbf{X}} = \hat{\mathbf{W}}\mathbf{A}\tilde{\mathbf{S}}$, we have

$$\begin{cases} \hat{\mathbf{Y}}_{\{i_1,i_2,\ldots,i_n\}} = \hat{\mathbf{W}}\mathbf{A}\tilde{\mathbf{S}}_{\{i_1,i_2,\ldots,i_n\}} \\ \hat{\mathbf{Y}}_{\{j_1,j_2,\ldots,j_n\}} = \hat{\mathbf{W}}\mathbf{A}\tilde{\mathbf{S}}_{\{j_1,j_2,\ldots,j_n\}}. \end{cases} \tag{27}$$

This yields

$$\hat{\mathbf{W}}\mathbf{A} = \hat{\mathbf{Y}}_{\{i_1,i_2,\ldots,i_n\}}\tilde{\mathbf{S}}^{-1}_{\{i_1,i_2,\ldots,i_n\}} = \hat{\mathbf{Y}}_{\{j_1,j_2,\ldots,j_n\}}\tilde{\mathbf{S}}^{-1}_{\{j_1,j_2,\ldots,j_n\}} \tag{28}$$

which leads to

$$\hat{\mathbf{Y}}_{\{i_1,i_2,\ldots,i_n\}} = \hat{\mathbf{Y}}_{\{j_1,j_2,\ldots,j_n\}}\tilde{\mathbf{S}}^{-1}_{\{j_1,j_2,\ldots,j_n\}}\tilde{\mathbf{S}}_{\{i_1,i_2,\ldots,i_n\}}. \tag{29}$$

Since $\hat{\mathbf{Y}}_{\{i_1,i_2,\ldots,i_n\}}$ and $\tilde{\mathbf{S}}_{\{j_1,j_2,\ldots,j_n\}}$ are nonnegative with the column sum-to-1 rule and $\hat{\mathbf{Y}}_{\{j_1,j_2,\ldots,j_n\}}$ and $\tilde{\mathbf{S}}_{\{i_1,i_2,\ldots,i_n\}}$ are permutation matrices, $\tilde{\mathbf{S}}_{\{j_1,j_2,\ldots,j_n\}}$ must be a permutation matrix. Thus, $\hat{\mathbf{Y}}_{\{i_1,i_2,\ldots,i_n\}}$ is a permutation matrix. Therefore, $\hat{\mathbf{W}}\mathbf{A} = \mathbf{P}$ is a permutation matrix, where $\mathbf{P} = \hat{\mathbf{Y}}_{\{i_1,i_2,\ldots,i_n\}}\tilde{\mathbf{S}}^{-1}_{\{i_1,i_2,\ldots,i_n\}}$. This completes the proof. □

References

[1] D. Peng, Y. Xiang, H. Trinh, A new blind method for separating $M + 1$ sources from M mixtures, Computers & Mathematics with Applications 60 (7) (2010) 1829–1839.
[2] T. Lee, M. Girolami, A.J. Bell, T.J. Sejnowski, A unifying information-theoretic framework for independent component analysis, Computers & Mathematics with Applications 39 (11) (2000) 1–21.
[3] T.H. Chan, W.K. Ma, C.Y. Chi, Y. Wang, A convex analysis framework for blind separation of non-negative sources, IEEE Transactions on Signal Processing 56 (10) (2008) 5120–5134.
[4] Y. Xiang, W. Yu, J. Zhang, S. An, Blind source separation based on phase and frequency redundancy of cyclostationary signals, IEICE—Transactions on Fundamentals E87-A (12) (2004) 3343–3349.
[5] M.Z. Ikram, D.R. Morgan, Permutation inconsistency in blind speech separation: investigation and solutions, IEEE Transactions on Speech and Audio Processing 13 (1) (2005) 1–13.
[6] F.Y. Wang, C.Y. Chi, T.H. Chan, Y. Wang, Nonnegative least-correlated component analysis for separation of dependent sources by volume maximization, IEEE Transactions on Pattern Analysis and Machine Intelligence 32 (5) (2010) 875–888.
[7] C.F. Caiafa, E. Salerno, A.N. Proto, L. Fiumi, Blind spectral unmixing by local maximization of non-gaussianity, Signal Processing 88 (1) (2008) 50–68.
[8] P. Georgiev, F. Theis, A. Cichocki, Sparse component analysis and blind source separation of underdetermined mixtures, IEEE Transactions on Neural Networks 16 (4) (2005) 992–996.
[9] A. Cichocki, R. Zdunek, S.I. Amari, New algorithms for non-negative matrix factorization in applications to blind source separation, in: IEEE International Conference on Acoustics, Speech and Signal Processing, 2006, pp. 5479–5482.
[10] A. Hyvrinen, Fast and robust fixed-point algorithms for independent component analysis, IEEE Transactions on Neural Networks 10 (3) (1999) 626–634.
[11] S. Choi, A. Cichocki, H.-M. Park, S.-Y. Lee, Blind source separation and independent component analysis: a review, Neural Information Processing—Letters and Reviews 6 (1) (2005) 1–57.
[12] J. Eriksson, V. Koivunen, Identifiability, separability, and uniqueness of linear ICA models, IEEE Signal Processing Letters 11 (7) (2004) 601–604.
[13] G.Q. Wang, Q.Z. Ding, Z.Y. Hou, Independent component analysis and its applications in signal processing for analytical chemistry, Trac—Trends in Analytical Chemistry 27 (4) (2008) 368–376.
[14] L. Wiskott, T.J. Sejnowski, Slow feature analysis: unsupervised learning of invariances, Neural Computation 14 (4) (2002) 715–770.
[15] J.V. Stone, Blind source separation using temporal predictability, Neural Computation 13 (7) (2001) 1559–1574.
[16] R. Zdunek, A. Cichocki, Improved M-FOCUSS algorithm with overlapping blocks for locally smooth sparse signals, IEEE Transactions on Signal Processing 56 (10) (2008) 4752–4761.
[17] J. Bobin, J.L. Starck, J. Fadili, Y. Moudden, Sparsity and morphological diversity in blind source separation, IEEE Transactions on Image Processing 16 (11) (2007) 2662–2674.
[18] S. Van Vaerenbergh, I. Santamaria, A spectral clustering approach to underdetermined postnonlinear blind source separation of sparse sources, IEEE Transactions on Neural Networks 17 (3) (2006) 811–814.
[19] I. Takigawa, M. Kudo, J. Toyama, Performance analysis of minimum l(1)-norm solutions for underdetermined source separation, IEEE Transactions on Signal Processing 52 (3) (2004) 582–591.
[20] Y.Q. Li, S.I. Amari, A. Cichocki, C.T. Guan, Probability estimation for recoverability analysis of blind source separation based on sparse representation, IEEE Transactions on Information Theory 52 (7) (2006) 3139–3152.
[21] J. Bobin, J.L. Starck, Y. Moudden, M.J. Fadili, Blind source separation: the sparsity revolution, Advances in Imaging and Electron Physics 152 (1) (2008) 221–302.
[22] L. Miao, H. Qi, Endmember extraction from highly mixed data using minimum volume constrained nonnegative matrix factorization, IEEE Transactions on Geoscience and Remote Sensing 45 (3) (2007) 765–777.
[23] D.D. Lee, H.S. Seung, Learning the parts of objects by non-negative matrix factorization, Nature 401 (6755) (1999) 788–791.
[24] A. Pascual-Montano, J.M. Carazo, K. Kochi, R.D. Pascual-Marqui, Nonsmooth nonnegative matrix factorization (nsNMF), IEEE Transactions on Pattern Analysis and Machine Intelligence 28 (3) (2006) 403–415.
[25] S. Jia, Y. Qian, Constrained nonnegative matrix factorization for hyperspectral unmixing, IEEE Transactions on Geoscience and Remote Sensing 47 (1) (2009) 161–173.
[26] F.Y. Wang, Y. Wang, T.H. Chan, C.Y. Chi, Blind separation of multichannel biomedical image patterns by non-negative least-correlated component analysis, in: Proceedings of Pattern Recognition in Bioinformatics, 2006, pp. 151–162.
[27] Y. Xiang, S.K. Ng, V.K. Nguyen, Blind separation of mutually correlated sources using precoders, IEEE Transaction on Neural Networks 21 (1) (2010) 82–90.
[28] I.J. Lustig, R.E. Marsten, D.F. Shanno, Interior point methods for linear programming: computational state of the art, ORSA Journal of Computing 6 (1) (1994) 1–14.
[29] K. Suzuki, R. Engelmann, H. MacMahon, K. Doi, Virtual dual-energy radiography: improved chest radiographs by means of rib suppression based on a massive training artificial neural network, Radiology 238 (2006) 932.
[30] http://webbook.nist.gov/chemistry/quant-ir/.
[31] http://publications.nigms.nih.gov/insidethecell/chapter1.html.
[32] S. Lang, Undergraduate Analysis, second ed., Springer-Verlag, 1983.