Medical Image Analysis 16 (2012) 239–251

Contents lists available at SciVerse ScienceDirect

Medical Image Analysis

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

A constrained independent component analysis technique for artery–vein separation of two-photon laser scanning microscopy images of the cerebral microvasculature

Hatef Mehrabian *, Liis Lindvere, Bojana Stefanovic, Anne L. Martel

Department of Medical Biophysics, University of Toronto, Sunnybrook Research Institute, Toronto, Ontario, Canada M4N 3M5

ARTICLE INFO

Article history:
Received 11 May 2010
Received in revised form 8 April 2011
Accepted 12 August 2011
Available online 25 August 2011

Keywords:
Artery–vein separation
Cerebral microvasculature
Constrained independent components analysis (CICA)
Non-negativity constraint
Gamma variate model

ABSTRACT

Understanding brain hemodynamics as well as the coupling between microvascular hemodynamics and neural activity is important in pathophysiology of cerebral microvasculature. When local increases in neuronal activity occur, the blood volume changes in the surrounding brain vasculature. Dynamic contrast enhanced imaging (DCE) is a powerful technique that quantifies these changes in the blood flow by repeatedly imaging the vasculature over time. Separating artery, vein and capillaries in the images and extracting their intensity-time curves from the DCE image sequence is an important first step in understanding vascular function. A constrained independent component analysis (ICA) technique is developed to analyze the two photon laser scanning microscopy (2PLSM) images of rat brain microvasculature, where a bolus of fluorescent dye is administered to the vascular system as the contrast agent. *A priori* information inferred from the gamma variate model of cerebral microvasculature is incorporated with the data driven technique in temporal and spatial domains using two constraints. The constraints are: no independent component (IC) is allowed to have negative contribution in forming the images (positivity constraint) and the component curves follow a gamma variate function (model fitting constraint). Experimental and simulation studies are conducted to demonstrate the improved performance of the proposed constrained ICA (CICA) technique over the most commonly used classical ICA algorithm (fast-ICA) in providing physiologically meaningful ICs and its ability to separate the model following factors from other factors are shown. The efficiency of CICA in handling noise is compared to model based techniques. Its capability in providing improved separation between artery, vein and capillaries compared to the other two techniques is also demonstrated.

1. Introduction

Increasing evidence has shown that dynamic information derived from dynamic contrast enhanced (DCE) imaging data can be used in understanding the pathophysiology of cerebral diseases (Calamante et al., 2000; Koenig et al., 2001). In DCE imaging of brain microvasculature, the contrast agent is injected intravenously and the vascular system is imaged repeatedly over time so as to track the passage of the contrast agent through the vascular bed (Wu and Liu, 2006). Passage of the tracer through the brain vasculature highlights the temporal changes of physical signals which results in intensity change in the captured image sequence. The tracer concentration change over time is used to assess the brain hemodynamic status (Wu and Liu, 2007).

The cerebral vascular system, like the neuronal network, exhibits both hierarchy and spatial specialization. Fenstermacher et al. (1991) showed that at rest, the blood flow varies up to 18-fold at different regions of a rat brain. This might result from regional differences in capillary density (Patlak et al., 1984). It can also be caused due to transit time variations in the vasculature (Rosen et al., 1991). A variety of central nervous system (CNS) conditions, such as stroke, dementia and epilepsy, involve a compromise in the resting blood flow (Mehrabian et al., 2010). The goal of this work is to characterize the spatial pattern of cerebral microvascular network reactivity under physiological conditions.

To date, there are limited data on the reactivity of deep cerebral microvessels and reported microvessel behavior has been highly variable. Unlike the prominent methodologies for non-invasive imaging of human brain function such as functional magnetic resonance imaging, positron emission tomography, and near-infrared spectroscopy, scanning confocal microscopy enables imaging of individual microvessels (Hutchinson et al., 2006). Two-photon laser scanning microscopy (2PLSM) provides additional important

* Corresponding author. Tel.: +1 4168263273.
E-mail addresses: hatef.mehrabian@sri.utoronto.ca (H. Mehrabian), liis.lindvere@sri.utoronto.ca (L. Lindvere), bojana@sri.utoronto.ca (B. Stefanovic), Anne.Martel@sri.utoronto.ca (A.L. Martel).

1361-8415/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.media.2011.08.002

advantages particularly for in vivo biological applications (Denk et al., 1990). Recently, the applicability of two-photo microscopy to the investigation of cerebral microcirculation has been shown in rat dorsal olfactory bulb during odor stimulation (Chaigneau et al., 2003), rat somatosensory cortex (Hutchinson et al., 2006; Kleinfeld et al., 1998), freely moving awake rat (Helmchen et al., 2001), etc. The current work investigates the use of independent component analysis (ICA) to separate the artery, vein and capillaries from a dynamic sequence of high resolution 2PLSM images that show the transit of a fluorescent bolus through the 3D microvascular tree.

In DCE imaging of microvasculature, the signal enhancement after injection of contrast agent reflects the passage of the contrast agent through the vascular system. This change in signal intensity (contrast concentration) is used to differentiate arteries and veins leading characterization of the vascular response based on the difference in the onset time and time to peak of contrast concentration (Hutchinson et al., 2006). Two main approaches are available to analyze these dynamic datasets, the model based techniques and the data driven techniques. The model based techniques involve fitting the data into a physiological model that characterizes passage of contrast agent through vascular system. The data-driven techniques deal with the dataset as a number of observations and try to extract their main latent or hidden features. Various data-driven techniques are available such as PCA (principal component analysis) (Martel et al., 2001), NMF (non-negative matrix factorization) (Lee and Seung, 1999; Chen et al., 2008) and ICA (independent component analysis) (Lu and Rajapakse, 2005). Among these methods the latter is a very powerful technique which also suits our problem as the spatial distribution of different vessels (arteries, veins and capillaries) in the imaging field are independent of each other. It should be noted that these vessels are not independent in the temporal domain as the signal intensity variations of the vessels are correlated.

1.1. Modeling vascular system (gamma variate model)

Modeling the passage of the contrast agent through the vascular system allows us to quantitatively assess its behavior. We have used a gamma variate function (Thompson et al., 1964) to model the signal intensity variations in each pixel of the DCE image dataset. The parameters of the model fit are used to characterize the contrast concentration, for example we can obtain parametric maps of onset time and time to peak using this approach (Hutchinson et al., 2006). The main advantage of model based techniques is that the parameter maps generated from this model are physiologically meaningful and therefore are minimally dependant on the acquisition protocol. However, due to the requirement of curve fitting for every individual pixel, this model based approach is very sensitive to noise and is very time consuming particularly in case of large 3D datasets.

1.2. Independent component analysis of dynamic structures

Independent component analysis (ICA) is a statistical method to describe variability among a large number of observed variables in terms of a few unobserved variables called the Independent Components (ICs). The observed variables are assumed to be linear combinations of the ICs (Nijran and Barber, 1986; Kim and Mueller, 1978). ICA is a very powerful technique that tries to map the DCE data space into a space constructed by a number of statistically independent components (Comon, 1994). It is a promising method that uses higher order statistical information and is increasingly used in various image analysis areas such as functional MRI, structural MRI, electroencephalography (EEG), etc. (McKeown and Sejnowski, 1998; Biswal and Ulmer, 1999; Calhoun and Adali, 2006). ICA has also been applied to DCE-CT images (Koh et al., 2006) and assessment of DCE images of cerebral blood perfusion (Wu and Liu, 2007).

This data driven method assumes having no prior information about the variations of the contrast agent and tries to extract the main underlying features (ICs) of the dataset. It is capable of handling noise efficiently and is less time consuming compared to model based techniques. On the other hand, the ICs extracted with this method do not necessarily carry physiologically meaningful information about the contrast agent behavior in the vasculature. These results are difficult to interpret and may vary with acquisition set up and conditions.

1.3. Constrained DCE image analysis technique

The goal of this work is to develop a DCE image analysis technique that combines the two analysis techniques (model based and data driven) to minimize the shortcomings of each method while taking advantage of their powerful aspects. A data driven independent component analysis technique is developed which incorporates *a priori* information about the gamma variate model of the vascular system into the temporal domain. It is known that the constrained ICA approach that incorporates additional requirements (Erdogmus et al., 2004; Lu and Rajapakse, 2005; Hesse and James, 2006) and prior information (Lu and Rajapakse, 2006) in the form of constraints into the ICA cost function helps to improve the performance of data decomposition. We have developed a Constrained ICA (CICA) technique that extracts ICs subject to two constraints: the non-negativity constraint (ICs are not allowed to have negative impact on the resultant images) in the spatial domain (Buvat et al., 1993; Barber, 1980) as well as the model fitting constraint (component curves follow a gamma variate function) in the temporal domain. These constraints impose the conditions of gamma variate model into the ICA routine and result in ICs that carry maximum information from both data driven and model based techniques. There are many methods of implementing spatial ICA; we have opted to use fast-ICA with maximizing negentropy (Hyvärinen and Oja, 2000) as this is the most commonly used ICA technique.

In the Section 2, classical ICA is introduced and the fast-ICA method is described. Constrained ICA (CICA) is then derived from ICA by adding temporal model fitting and spatial non-negativity constraints. Synthetic DCE data is generated to model the function of cerebral microvasculature and experimental DCE dataset that tracks passage of a bolus of dye through rat brain microvasculature is acquired. In Section 3, the proposed method has been validated by simulation as well as experimental study. The simulation study compares the Constrained ICA (CICA) with fast-ICA (Hyvärinen and Oja, 2000) and parametric mapping (model fitting) to highlight its superior performance. In the experimental study the results of fast-ICA, CICA and parametric mapping are compared.

2. Theory and methods

2.1. Independent component analysis

Independent component analysis (ICA) is a statistical signal processing approach that aims to extract underlying features of the data set (unobserved components or source signals) from observed mixtures such that the extracted features are mutually independent, without assuming any knowledge of the mixing coefficients (Comon, 1994). In this article we will use uppercase bold letter for 2-D matrices, lowercase bold letters for column vectors, lowercase letters (not bold) for scalars and bold, italic letters for functions.

The core idea of ICA is motivated from blind source separation problem for data model of the form

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

where $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N]^T$ is a matrix of the N observed mixtures (frames or images), $\mathbf{S} = [\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_M]^T$ is a matrix containing the M source signals, independent components or ICs (usually $M \leqslant N$) and $\mathbf{A} \in \mathbb{R}^{N\times M}$ is the mixing matrix. We call the subspace that the columns of **A** span, the mixing space and the columns of **A,** the component curves. We also call the space that the rows of **S** span as the source space and the space that the rows of **X** span as the observed data space. The aim of ICA is to estimate the independent components **S** and the mixing matrix **A** having the observed mixture signals **X**.

Classical ICA algorithms try to find an unmixing matrix $\mathbf{W} \in \mathbb{R}^{\mathbb{M}\times\mathbb{N}}$ (we call the space that columns of the unmixing matrix span, the unmixing space) and estimate the IC matrix $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_M]^T$ such that:

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

where rows of **Y** are statistically independent. The ICs can be recovered up to a scaling and permutation.

Several estimation methods have been proposed for ICA such as ICA by maximizing non-Gaussianity (Novey and Adali, 2005), ICA by maximum likelihood estimation (Jang and Lee, 2004) and ICA by maximizing mutual information (Almeida, 2004). In this study an ICA approach is developed based on maximizing non-Gaussianity of the estimated ICs.

Based on the central limit theorem, the distribution of a sum of independent random variables tends towards a Gaussian distribution (Araujo and Giné, 1980). Thus, to find the independent components of an observed dataset, one has to maximize the non-Gaussianity of the extracted components. A robust measure of non-Gaussianity is Negentropy which is based on the information theoretic quantity of differential entropy. Entropy quantifies the randomness of an observed variable. The differential entropy $\boldsymbol{H}$ of a random variable with density $p_y(v)$ is defined as:

$$\boldsymbol{H}(\mathbf{y}) = -\int p_y(v)\log p_y(v)\,\mathrm{d}v \tag{2}$$

A fundamental result of information theory is that among all random variables of equal variance, a Gaussian variable has the largest entropy. Thus, the Negentropy as a measure of non-Gaussianity of a signal **y** is defined as:

$$\boldsymbol{J}(\mathbf{y}) = \boldsymbol{H}(\mathbf{y}_{\mathbf{Gauss}}) - \boldsymbol{H}(\mathbf{y}) \tag{3}$$

where $\mathbf{y}_{\mathbf{Gauss}}$ is a Gaussian random variable having zero mean and the same variance as the signal **y** (Hyvärinen and Oja, 2000). Based on this definition, Negentropy is always non-negative. It is also invariant for invertible linear transformations (Comon, 1994; Hyvärinen, 1999a). Estimating Negentropy given in (3) is complex and difficult, therefore; an estimate to Negentropy, given in (4), is usually used for ICA purposes:

$$\boldsymbol{J}(\mathbf{y}) = \rho(\mathrm{E}\{\boldsymbol{G}(\mathbf{y})\} - \mathrm{E}\{\boldsymbol{G}(\mathbf{y}_{\mathbf{Gauss}})\})^2 \tag{4}$$

where ρ is a positive constant, $\boldsymbol{G}(\cdot)$ is a non-quadratic function and $\mathrm{E}\{\cdot\}$ is the expectation operator. Various non-quadratic functions can be chosen for $\boldsymbol{G}(\cdot)$ but to have a robust estimator it is recommended to use functions that do not grow fast (Hyvärinen and Oja, 2000). The following function is used for $\boldsymbol{G}(\cdot)$ in our algorithm.

$$\mathbf{G}(\mathbf{y}) = \frac{1}{a_1}\log\cosh a_1\mathbf{y} \tag{5}$$

where $1 \leqslant a_1 \leqslant 2$ and in this study $a_1 = 1$ is used in $\boldsymbol{G}(\cdot)$. In classical ICA, the independent component as well as its mixing vector is estimated by maximizing the Negentropy ($\boldsymbol{J}(\cdot)$) in an iterative optimization process. Here, we have used the optimization algorithm introduced in the fast-ICA technique (Hyvärinen, 1999b).

So far estimation of only one IC is described. In order to estimate more ICs we use the deflation method which involves deflationary orthogonalization of **W** using the Gram-Schmidt method. This means that we estimate the independent components one by one. Once p ICs or p vectors $\mathbf{w}_1, \ldots, \mathbf{w}_p$ are estimated, the IC estimation algorithm is run for $\mathbf{w}_{p+1}$ and after every iteration the projections $(\mathbf{w}_{p+1}^T\mathbf{w}_j)\mathbf{w}_j$, $j = 1, \ldots, p$ of the previously estimated p vectors are subtracted from $\mathbf{w}_{p+1}$ and it is renormalized.

2.2. Constrained ICA (CICA)

We have developed a constrained ICA technique similar to the one unit ICA-R (ICA with reference) algorithm developed by Lu and Rajapakse (2005). ICA-R tries to find an $N \times 1$ unmixing vector $\mathbf{w}^*$ such that the output signal $\mathbf{y} = \mathbf{w}^{*T}\mathbf{X}$ is equal to the desired IC signal $\mathbf{s}^*$ subject to *a priori* information about $\mathbf{s}^*$. The ICA-R method applied the constraint to the spatial domain (ICs) or the rows of $\mathbf{S} = [\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_M]^T$ matrix.

The *a priori* information used in this study is based on the assumption that the component curves (columns of **A**) representing the cerebral vasculature hemodynamics should take the form of a gamma variate function. Therefore our methods differs significantly from that described by Lu and Rajapakse (2005) in two important respects; firstly the constraint in our algorithm is set on the temporal domain rather than the spatial domain and, secondly we do not need to know the exact form of the reference curve *a priori*, we only assume that it follows a gamma variate form. The general form of the gamma variate function that was presented by Thompson et al., 1964 is given in the following equation:

$$\mathbf{C}(t) = \begin{cases} a & t \leqslant AT \\ a + K(t-AT)^{\alpha}\mathrm{e}^{-(t-AT)/\beta} & AT < t \leqslant t_b \\ b & t_b < t \end{cases} \tag{6}$$

where t is time after injection, $\boldsymbol{C}(t)$ is concentration at time t, K is a constant scale factor, AT is the appearance time, a is the bias (the signal before arrival of the bolus), b is the residue (the concentration that remains in the blood flow after passage of the bolus), t_b is the time at which the concentration drops to b and α, β are arbitrary parameters. The vessel specific dynamics are inferred from the parameters extracted from fitting this gamma curve to the data. The onset time of the contrast agent (bolus of dye) is equal to the appearance time of its gamma curve fit; the relationship between time-to-peak and gamma function parameters is $t_p = AT + \alpha\beta$.

As our constraint is on the mixing matrix **A** whereas the ICA technique operates on the unmixing matrix **W**, a transformation from mixing space to unmixing space is required. At each iteration the estimated unmixing vector $\mathbf{w}_k$ is transformed into the mixing space $\mathbf{a}_k$, then the gamma variate curve introduced in (5) is fitted into the vector $\mathbf{a}_k$ using the Levenberg–Marquardt Curve-Fitting Algorithm (Ahearn et al., 2005) to generate the reference concentration curve $\mathbf{c}_k$.

For an ideal separation $\mathbf{AW} = \mathbf{I}$ and assuming that $\mathbf{c}'_k$ is the transformation of $\mathbf{c}_k$ into the unmixing space, the *a priori* constraint added to the ICA learning algorithm $(\boldsymbol{\epsilon}(\mathbf{w}, \mathbf{c}'_k) = \|\mathbf{w} - \mathbf{c}'_k\|^2)$ is a measure of how well the estimated component curve fits the gamma variate function. Setting a threshold ζ, the desired closeness of the extracted curve to its gamma variate fit is defined as:

$$\boldsymbol{g}(\mathrm{w}) = \varepsilon(\mathbf{w}, \mathbf{c}'_k) - \zeta$$

This equation is incorporated to the ICA optimization algorithm as an inequality constraint of the form $\boldsymbol{g}(\mathbf{w} \leqslant 0)$. The other constraint that is also inferred from the gamma variate model of the system,

known as the non-negativity constraint, is that no component image is allowed to have negative contribution to the resultant mixed image (Barber, 1980; Buvat et al., 1993).

Two pre-processing steps on the observed data are necessary before application of ICA, centering and whitening. Centering is subtracting the mean of each observation to make rows of $\mathbf{X}$ zero-mean variables. Whitening involves a linear transformation of the centered data such that the components of the resultant data matrix $\hat{\mathbf{X}}$ are uncorrelated and their variance equals unity.

Incorporating $\boldsymbol{g}(\mathbf{w})$ and positivity as the two constraints to the negentropy function of (4), the cost function of the constrained ICA algorithm can be formulated as (7):

$$\begin{array}{ll} \text{maximize} & \boldsymbol{J}(\mathbf{y}) \approx \rho[\mathrm{E}\{\boldsymbol{G}(\mathbf{y})\} - \mathrm{E}\{\boldsymbol{G}(\mathbf{y}_{\mathbf{Gauss}})\}]^2 \\ \text{Subject to} & \boldsymbol{g}\mathbf{w} \leqslant 0 \quad \text{(model fitting)} \\ \& & \mathbf{y}'(\mathrm{i}) \geqslant 0 \quad \text{(non-negativity)} \end{array} \tag{7}$$

where $\mathbf{y}'$ is transformation of $\mathbf{y}$ into the space of un-whitened and un-centered data and $\mathbf{y}'(i) \geqslant 0$ means every element in vector $\mathbf{y}'$ has to be non-negative. The method of Lagrange multipliers (Bertsekas, 1982) is adapted to search for the optimal solution. The corresponding augmented Lagrangian function $\mathbf{L}$ is given by (Lin et al., 2007):

$$\boldsymbol{L}(\mathbf{w},\mu) = \boldsymbol{J}(\mathbf{y}) - \left(\mu\hat{\boldsymbol{g}}(\mathbf{w}) + /\frac{1}{2}\gamma\|\hat{\boldsymbol{g}}(\mathbf{w})\|^2\right) \tag{8}$$

where $\hat{\boldsymbol{g}}(\mathbf{w}) = \boldsymbol{g}(\mathbf{w}) + z^2$ transforms the original inequality constraints into equality constraints with a vector of slack variables z, μ is the positive Lagrange multiplier corresponding to inequality constraint. γ is the penalty parameter and $\|\cdot\|$ denotes the Euclidean norm. The quadratic penalty term $\frac{1}{2\gamma}\|\cdot\|^2$ ensures that the maximization problem holds the condition of local convexity: the Hessian matrix of $\boldsymbol{L}$ is positive–definite (Lu and Rajapakse, 2005). The inequality constraint is further translated to eliminate slack variable z by using the procedure explained in (Bertsekas, 1982). The CICA objective function is given by (Lin et al., 2007):

$$\boldsymbol{L}(\mathbf{w},\mu) = \boldsymbol{J}(\mathbf{y}) = -\frac{1}{2\gamma}[\max^2\{\mu + \gamma\boldsymbol{g}(\mathbf{w}), 0\} - \mu^2] \tag{9}$$

The non-negativity constraint ($\mathbf{y}'(i) \geqslant 0$) is not shown in the augmented cost functions as it is applied by truncating the negative elements of the estimated IC and re-mapping it to $\mathbf{X}$ to calculate the new unmixing vector $\mathbf{w}$ that satisfies the non-negativity constraint at each iteration. A Newton-like learning algorithm, similar to the one presented in Lu and Rajapakse (2005) and Lin et al. (2007) is used to solve the iterative optimization problem of (9) to update the unmixing vector $\mathbf{w}$:

$$\mathbf{w}_{\mathbf{k}+1} = \mathbf{w}_{\mathbf{k}} - \eta\boldsymbol{L}'_{\mathbf{w}_{\mathbf{k}}}/\delta(\mathbf{w}_{\mathbf{k}})$$

where η is the learning rate and

$$\boldsymbol{L}'_{\mathbf{w}_{\mathbf{k}}} = \bar{\rho}\mathrm{E}\left\{\hat{\mathbf{X}}\boldsymbol{G}'_{\mathbf{y}}(\mathbf{y})\right\} - 0.5\mu\mathrm{E}\left\{\boldsymbol{g}'_{\mathbf{w}_{\mathbf{k}}}(\mathbf{w}_{\mathbf{k}})\right\}$$

$$\delta(\mathbf{w}_{\mathbf{k}}) = \bar{\rho}\mathrm{E}\left\{\hat{\mathbf{X}}\boldsymbol{G}''_{\mathbf{y}^2}(y)\right\} - 0.5\mu\mathrm{E}\left\{\boldsymbol{g}''_{w_k^2}(\mathbf{w}_{\mathbf{k}})\right\}$$

where $\bar{\rho} = \pm\rho$, whose sign is determined by the sign of $\mathbf{E}\{\boldsymbol{G}(\mathbf{y})\} - \mathbf{E}\{\boldsymbol{G}(\mathbf{y}_{\mathbf{Gauss}})\}$, $\boldsymbol{G}'_y(\mathbf{y})$ and $\boldsymbol{G}''_{y^2}(\mathbf{y})$ are the first and second derivatives of $\boldsymbol{G}(\cdot)$ with respect to $\mathbf{y}$ and $\boldsymbol{g}'_{\mathbf{w}_{\mathbf{k}}}(\mathbf{w}_{\mathbf{k}})$ and $\boldsymbol{g}''_{\mathbf{w}_{\mathbf{k}}^2}(\mathbf{w}_{\mathbf{k}})$ are the first and second derivatives of $\boldsymbol{g}(\cdot)$ with respect to $\mathbf{w}$. The updating is followed by normalization of the vector $\mathbf{w}_{\mathbf{k}}$:

$$\mathbf{w}_{\mathbf{k}+1} = \mathbf{w}_{\mathbf{k}+1}/\|\mathbf{w}_{\mathbf{k}+1}\|$$

The non-negativity constraint (details in Appendix) is then applied by mapping the observed data space via the estimated unmixing vector $\mathbf{w}_{\mathbf{k}}$ into the source space, truncating the negative elements of the estimated IC and then re-mapping the IC into the observed data space to get the updated unmixing vector $\mathbf{w}_{\mathbf{k}}$. This is followed by normalization and orthogonalization of the estimated unmixing vector. The Lagrange multiplier is learned by the following gradient ascent method (Lu and Rajapakse, 2005).

$$\mu_{\mathrm{k}+1} = \max\{0, \mu_k + \gamma\boldsymbol{g}(\mathbf{w}_{\mathbf{k}})\}$$

The value of learning rate η is set to 1 ($\eta = 1$) in both fast-ICA and constrained ICA. A small value (with respect to the norm of the estimated curve) is chosen for threshold ζ as it shows how close the estimated curve is to its gamma variate fit. The initial value of Lagrange multiplier μ is chosen intuitively such that the first and second terms in the right hand side of (8) have the same order of magnitude. The penalty parameter γ is chosen such that the Lagrange multiplier μ does not increase or decrease quickly. The constrained ICA algorithm can be described as follows:

1. Center and whiten the observed signals $\mathbf{X}$ to make $\hat{\mathbf{X}}$.
2. Choose a proper scalar penalty parameter γ.
3. Take a random initial vector $\mathbf{w}_0$ of norm 1, or just let $\mathbf{w}_0 = 0$.
4. Choose an initial value for the Lagrange multiplier μ.
5. Compute the constant $\bar{\rho} = \mathbf{E}\{\boldsymbol{G}(\mathbf{y})\} - \mathbf{E}\{\boldsymbol{G}(\mathbf{y}_{\mathbf{Gauss}})\}$.
6. Update the Lagrange multiplier μ by
   $\mu_{K+1} = \max\{0, \mu_k + \gamma\boldsymbol{g}(\mathbf{w}_{\mathbf{k}})\}$.
7. Update the unmixing vector $\mathbf{w}$ by
   $\mathbf{w}_{\mathbf{k}+1} = \mathbf{w}_{\mathbf{k}} - \eta\boldsymbol{L}'_{\mathbf{w}_{\mathbf{k}}}/\delta(\mathbf{w}_{\mathbf{k}})$.
8. Normalize the unmixing vector $\mathbf{w}$ by
   $\mathbf{w}_{\mathbf{k}+1} = \mathbf{w}_{\mathbf{k}+1}/\|\mathbf{w}_{\mathbf{k}+1}\|$.
9. Apply non-negativity by truncating the estimated IC and re-mapping it to the data space to update $\mathbf{w}$.
10. Normalize the unmixing vector $\mathbf{w}$ after non-negativity.
11. Orthogonalize the unmixing vector $\mathbf{w}$ by
    $\mathbf{w}_{\mathbf{k}+1} = \mathbf{w}_{\mathbf{k}+1} - \sum_{j=0}^{k}(\mathbf{w}_{\mathbf{k}+1}^{\mathbf{T}}\mathbf{w}_{\mathbf{j}})\mathbf{w}_{\mathbf{j}}$.
12. Until $|\mathbf{w}_{\mathbf{k}}^{\mathbf{T}}\mathbf{w}_{\mathbf{k}+1}|$ approaches 1 (up to a small error), otherwise go back to the step (5).

2.3. Adding mean back to both ICs and curves

One of the pre-processing steps for ICA is centering the observed data. Centering is subtracting the mean of each observation from its measured data to make it zero mean:

$$\hat{\mathbf{X}} = \mathbf{X} - \mathbf{m}_{\mathbf{x}}$$

where $\mathbf{m}_{\mathbf{x}}$ is a vector containing the mean value of each observation. The effect of this pre-processing step on the estimated ICs is usually accounted for by calculating the contribution of each component in the mean vector $\mathbf{m}_{\mathbf{x}}$ and adding it to the estimated IC:

$$\mathbf{Y} = \mathbf{W}(\hat{\mathbf{X}} + \mathbf{m}_{\mathbf{x}}) \tag{10}$$

but the effect of centering in the temporal domain is usually overlooked. As the inequality constraint in the proposed method is in the temporal domain, this effect has to be accounted for at each iteration and the curve fitting has to be performed on the component curve after removing the effect of centering. The centering effect is removed by first adding the mean to the ICs in spatial domain as shown in (10). Assuming $\mathbf{Y} = \mathbf{S}$ and using (1), the actual component curves are calculated:

$$\mathbf{A} = \mathbf{X}\mathbf{Y}^{+}$$

where $\mathbf{Y}^{+}$ is the pseudo inverse of $\mathbf{Y}$ or estimated ICs.

2.4. Simulation study

A simulation study is conducted to assess the performance of the proposed method and to compare its results to the fast-ICA technique presented by Hyvärinen and Oja (2000). The synthesized dataset simulates the passage of a bolus of dye (contrast agent) through the cerebral microvasculature. A total of seven simulated concentration curves are employed in this simulation. The synthetic image data set containing the seven source images (independent components or ICs) as well as their corresponding component curves are shown in Fig. 1a and b respectively. The contrast concentration changes of the first five components follow the gamma variate function. Components 1 and 2 simulate the entrance of the dye into and its exit from the artery (the artery and the delayed artery) respectively. Components 3 and 4 represent the entrance of the dye into and its exit from the vein (the vein and the delayed vein) respectively. Component 5 is representative of the capillaries that are present in the imaging field of view whose concentration curve has a peak in between that of the artery and vein. Components 6 and 7 are designated to other structures that may be present in the image and may produce signal in the image, therefore their component curves do not follow a gamma variate model. These components are used to demonstrate the ability of the proposed method in separating such structures from the ones that are of interest.

The synthetic dynamic images consist of 80 frames of 64×64 pixels with a temporal resolution of 0.1s. Gaussian additive noise is added to the data to simulate noisy concentration curves. The signal-to-noise-ratios (SNR) of 5, 10, 20 and 30 that are common in DCE imaging are used. SNR is approximated by the ratio of the mean of the signal to standard deviation of the noise signal. The simulation study was carried out in IDL 7 on a Pentium IV PC with 3.00 GHz Core2 CPU and 3 GB of RAM.

2.5. Experimental dynamic contrast enhanced image data set acquisition

Adult male Sprague–Dawley rats were anesthetized with isoflurane, tracheotomized and mechanically ventilated. To enable two photon laser scanning microscopy (2PLSM) imaging of the brain, stereotaxic surgery was done to prepare a small (~5 mm diameter), closed (1% agarose) cranial window over the forelimb representation in the primary somatosensory cortex (~3.5 mm lateral to midline just anterior to bregma).

The tail vein was cannulated to allow administration of the bolus of the fluorescent contrast agent (Texas Red dextran, 70 kDa, 0.05 ml, 25 mg/kg). Muscle relaxant was administered to minimize residual motion. Rectal temperature, tidal pressure of ventilation, arterial blood pressure, and heart rate were monitored and recorded using a BIOPAC MP system (Biopac Systems, Inc., Goleta, CA) throughout the experiments. Two-photon laser scanning microscopy was done using a 20×, 0.95 NA, 2.0 mm working distance objective (Olympus). The bolus tracking experiments employed a 320^2 matrix over ~600 μm^2 field of view, with a 4 μs dwell time, resulting in a temporal resolution of 0.366 s per frame. Fig. 2 shows a sample frame where artery, vein and some capillaries are visible.

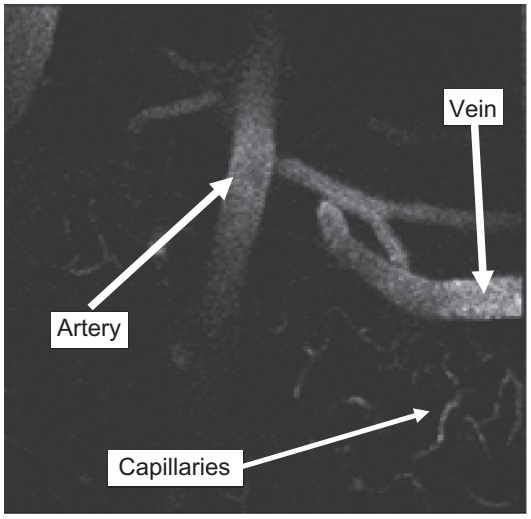

Fig. 2. A sample frame of the image data set where artery, vein and small capillaries are visible.

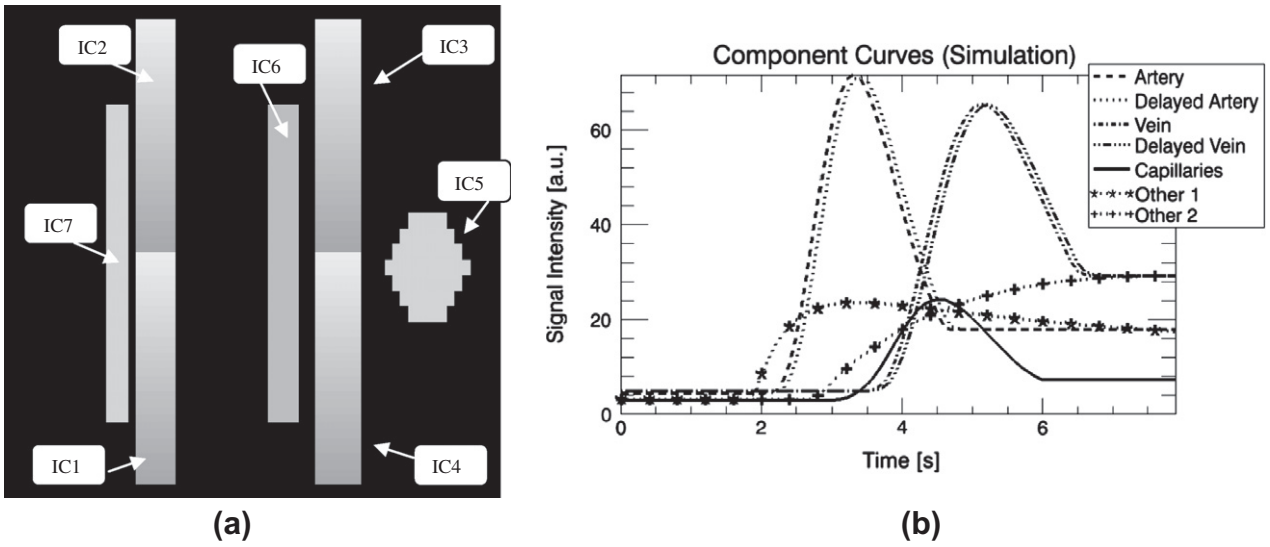

Fig. 1. The noise free synthetic data. (a) The seven source images corresponding to the component images, components 1 and 2 represent the artery and the delayed artery respectively, components 3 and 4 simulate the vein and the delayed vein, component 5 represents the capillaries, components 6 and 7 represent other components named as other 1 and other 2 respectively. (b) The 7 component curves corresponding to the seven component images. The signal intensities of all plots are shown in arbitrary units (a.u.).

3. Results

3.1. Simulation results

The proposed CICA as well as the most commonly used classical ICA technique called fast-ICA were applied to the synthesized dynamic data. Fig. 3 shows the estimated ICs of CICA for SNR = 5 and Fig. 4 shows the ICs of fast-ICA for the same data set (SNR = 5). The values of adjustable parameters used in the constrained ICA technique, defined as described in Section 2 are as follows: threshold $\zeta = 0.005$ initial value of Lagrange multiplier $\mu = 0.05$, penalty parameter $\gamma = 0.01$ and learning rate $\eta = 1$. The only adjustable parameter in fast-ICA is the learning rate η whose value is chosen the same as constrained ICA ($\eta = 1$). In these figures, the artery and the delayed artery are extracted as one IC. This also happens for the vein and the delayed vein. This issue has been addressed in Section 4.

To compare the results of data driven techniques with the conventional pixel by pixel model fitting techniques (Ahearn et al., 2005), binary masks have been generated from the time to peak map of the same dataset (SNR = 5). This map was generated by fitting signal intensity curve of each pixel to a gamma variate function and calculating their time to peak parameter. The histogram of this parameter was plotted and three peaks were identified. These three peaks were used to identify three phases and the pixels corresponding to each of these phases are shown in Fig. 5. These binary images are very noisy compared to those obtained using the proposed method.

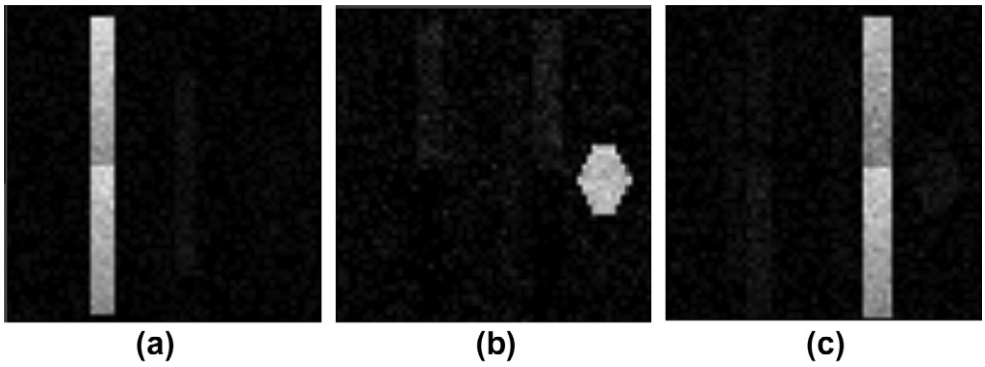

Fig. 3. The estimated independent components (ICs) resulted from applying CICA to the synthetic image data set shown in Fig. 1 in presence of additive Gaussian noise (SNR = 5): (a) artery, (b) capillaries and (c) vein.

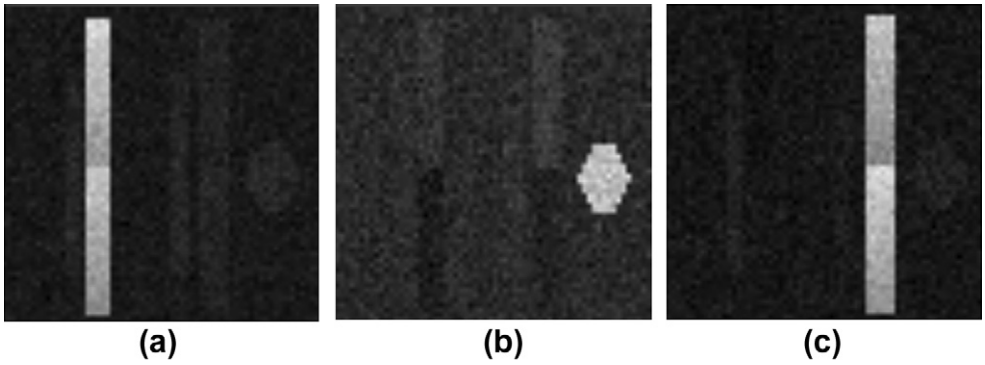

Fig. 4. The estimated independent components (ICs) resulted from applying fast-ICA to the synthetic image data set shown in Fig. 1 in presence of additive noise (SNR = 5): (a) artery, (b) capillaries and (c) vein.

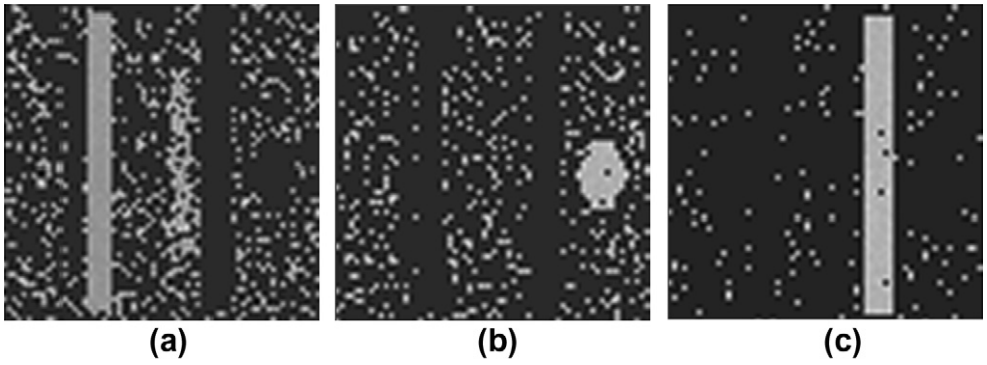

Fig. 5. The binary images generated from the time to peak map obtained using conventional pixel by pixel model fitting of the synthetic image data set shown in Fig. 1 in presence of additive noise (SNR = 5): (a) artery, (b) capillaries and (c) vein.

Figs. 6 and 7 show the curves corresponding to the component curves of CICA and fast-ICA respectively (SNR = 5). In these figures, the true component curve that is used to generate the simulation data as well as the estimated component curve before and after adding the mean back to the data (removing the effect of centering the data) are shown for each component. As shown in the images the extracted curves of CICA follow the gamma variate function better than the results of fast-ICA. They also show that centering the data has significant effect on the extracted curves and it is required to remove the effect of centering in order to have more accurate estimation of the component curves.

Fig. 8 shows the average signal intensity curves for the three components of pixel by pixel model fitting. As can be seen in this figure the difference between average intensity curves and the true curves is more than the ICA based techniques which demonstrates that the ICA based techniques outperform the conventional model fitting techniques.

Table 1 reports the mean squared error (MSE) between the extracted component curves after adding mean to the data and the true component curves for both fast-ICA and CICA techniques.

Table 2 reports the signal-to-noise-ratio (SNR) of the extracted ICs measured as the ratio of the mean of the signal of estimated ICs to the standard deviation of the background noise. The signal of estimated IC refers to the signal over the entire artery (the artery and the delayed artery components), the entire vein (the vein and the delayed vein components), or the capillaries component. The standard deviation of background noise is calculated by excluding the regions corresponding to the seven components and calculating the standard deviation of the signal in the remaining regions (background). As shown in this table, the SNR of all images are higher for CICA compared to fast-ICA. Signal to interference ratio (SIR) is defined as the ratio of the mean of each estimated IC to the standard deviation of the interference from other ICs in that IC (signal in the areas of the image that correspond to

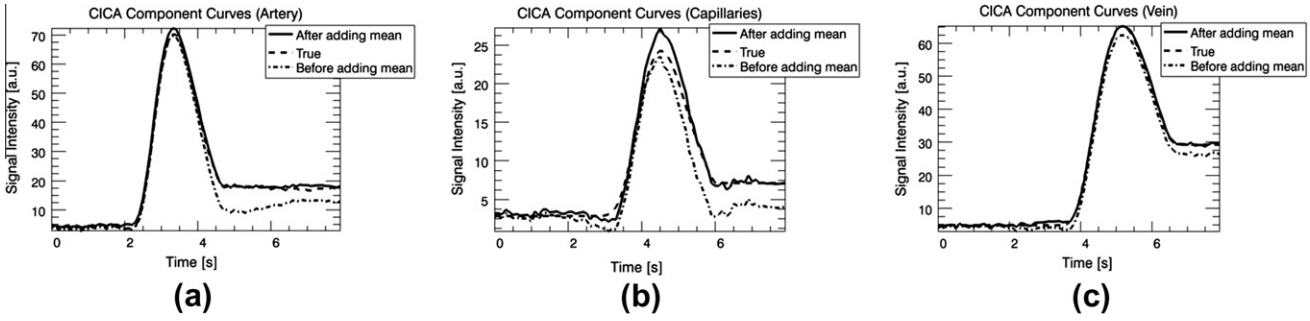

Fig. 6. The estimated component curves corresponding to the model-following components for CICA in presence of additive noise (SNR = 5): (a) artery, (b) capillaries and (c) vein.

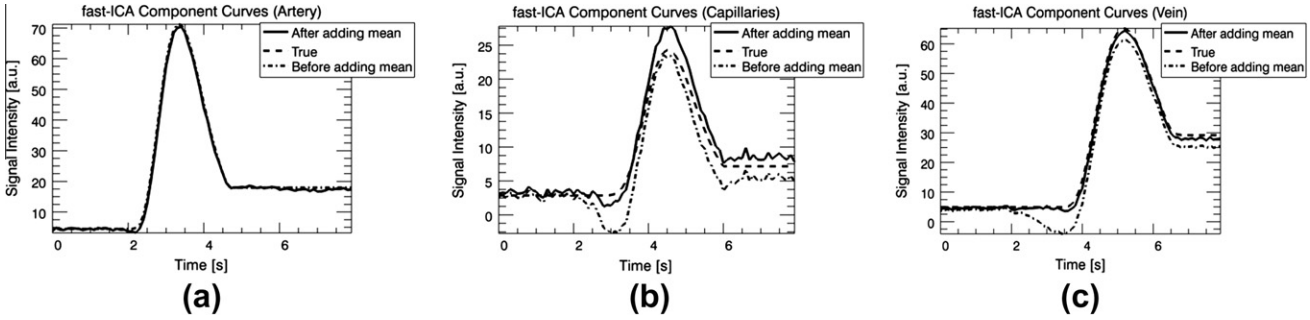

Fig. 7. The estimated component curves corresponding to the model-following components for fast-ICA in presence of additive noise (SNR = 5): (a) artery, (b) capillaries and (c) vein.

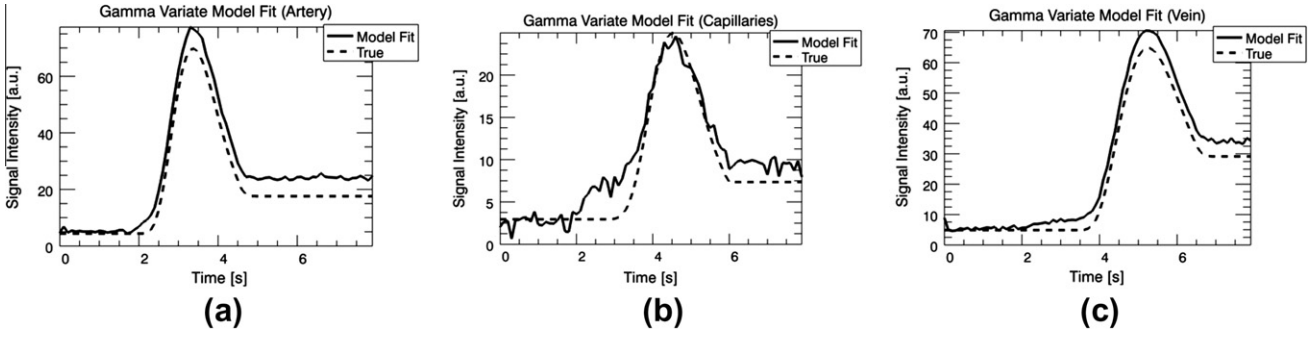

Fig. 8. The average signal intensity curves corresponding to the three components generated from pixel by pixel Gamma variate model fitting in presence of additive noise (SNR = 5): (a) artery, (b) capillaries and (c) vein.

Table 1
The mean squared error (MSE) between the extracted component curves and the true component curves for both fast-ICA and CICA techniques.

	Low SNR: 5		High SNR: 30	
	CICA	Fast-ICA	CICA	Fast-ICA
Artery	0.23	0.64	0.022	0.028
Vein	0.29	1.15	0.15	0.12
Intermediate capillaries	0.97	2.45	0.098	0.96
Mean	0.49	1.41	0.09	0.37

Table 2
The signal-to-noise-ratio (SNR) of the extracted ICs measured as the ratio of the mean of the estimated ICs to the standard deviation of the background noise.

	Low SNR: 5		High SNR: 30	
	CICA	Fast-ICA	CICA	Fast-ICA
Artery	32.4	17.8	63.4	40.2
Vein	25.4	15.3	38.5	36.4
Intermediate capillaries	19.9	11.9	44.8	26.5
Mean	25.9	15	48.9	34.4

Table 3
The signal-to-interference-ratio (SIR) of the extracted ICs measured as the ratio of the mean of each estimated IC to the standard deviation of the interferences from other ICs in that IC.

	Low SNR: 5		High SNR: 30	
	CICA	Fast-ICA	CICA	Fast-ICA
Artery	21.3	17.1	56.8	37.7
Vein	15	11.51	17.9	31.3
Intermediate capillaries	12.3	8.2	18.8	11.4
Mean	16.2	12.3	31.2	26.8

all other ICs). Values of SIR for two noise levels (SNR = 5 and SNR = 30) are reported in Table 3. The SNR and SIR values could not be provided for the model fitting technique as there were no images for the components (only three masks were generated).

The performance of source separation is also evaluated by a performance index (PI) of permutation error

$$\mathrm{PI} = \frac{1}{m}\left(\sum_{i=1}^{m} rPI_i + \sum_{j=1}^{m} cPI_j\right)$$

where $rPI_i = \sum_{j=1}^{m} |p_{ij}| / \max |p_{ik}| - 1$, $cPI_j = \sum_{i=1}^{m} |p_{ij}| / \max k|p_{kj}| - 1$ in which p_{ij} is an element of the permutation matrix $\mathbf{P} = \mathbf{WA}$ ($\mathbf{W}$ is the unmixing matrix and $\mathbf{A}$ is the mixing matrix). PI is commonly used in the ICA literature, particularly in the context of constrained ICA, to assess the performance of ICA techniques in source separation. The term cPI_j measures the degree of the source c_j appearing multiple times at the output and the term rPI_i gives the error of the separation of the output component y_i with respect to the sources. The lower the PI, the better is the performance of the algorithm. Table 4 reports the values of PI for different SNR's for CICA and fast-ICA and pixel by pixel model fitting.

Table 4
The values of PI for different SNR's for both CICA and fast-ICA algorithms.

	SNR = 5	SNR = 10	SNR = 20	SNR = 30
PI (CICA)	0.058	0.047	0.037	0.02
PI (fast-ICA)	0.064	0.056	0.078	0.04
Pixel by Pixel fit	0.22	0.21	0.175	0.16

3.2. Experimental results

The rat brain images are fed to the CICA algorithm and the main ICs and component curves are extracted. To illustrate the performance of the method the results of one data set containing 50 images of the cerebral microvasculature is presented. The imaging field of view is such that artery and vein are clearly visible. There are also some small vessels and capillaries in the field of view. The imaging scans the interval starting before injection of the contrast agent until it washes out and the signal intensity plateaus. The ICs and their corresponding component curves resulting from applying the CICA to the dataset are shown in Fig. 9a–d respectively.

In order to evaluate the performance of the method we also applied fast-ICA to this data set as well as extracting parametric maps which is used in model fitting techniques. The extracted ICs for fast-ICA technique and their corresponding component curves are shown in Fig. 10a–e. The values of adjustable parameters used in the constrained ICA technique that are defined as described in Section 2 are as follows: threshold $\zeta = 1$ initial value of Lagrange multiplier $\mu = 0.01$, penalty parameter $\gamma = 0.001$ and learning rate $\eta = 1$. The only adjustable parameter in fast-ICA is the learning rate η whose value is chosen the same as constrained ICA ($\eta = 1$).

In Fig. 11 binary images and their corresponding curves have been generated from the time to peak map obtained using conventional pixel by pixel model fitting (Ahearn et al., 2005). The histogram of the time to peak map was plotted and 3 peaks were identified; these peaks were used to identify three phases and the pixels corresponding to each of these phases are shown in Fig. 11. These binary images are very noisy compared to those obtained using the proposed method.

As ICA is capable of extracting ICs up to a scaling and permutation, multiplying the ICs or their corresponding curves with a scalar value does not change the results. Thus all of the extracted component curves (both for fast-ICA and CICA) are normalized with respect to the maximum of the artery curves in model fitting technique. This scaling is required to be able to compare the results of different techniques. Table 5 reports the mean squared distance of the normalized component curves from their gamma variate fit for both CICA and fast-ICA techniques.

Table 6 reports the time to peak and Table 7 gives the onset time of all three components of different methods which shows the results of CICA match with the model fitting data better than the results of fast-ICA.

4. Discussions

4.1. Simulation study

The simulation results show that the artery and the delayed artery (the vein and the delayed vein) ICs are extracted as one IC in both fast-ICA and CICA techniques. This is due to high correlation between the two components in the temporal domain. As the IC estimation process requires orthogonalization of the unmixing matrix, if two components are highly correlated they cannot be separated in ICA techniques. Therefore, the artery and the delayed artery (the vein and the delayed vein) components are mixed and extracted as one component.

In the synthesized data, there is relatively high degree of correlation between the components in the temporal domain. This simulates the cerebral microcirculation in which the contrast uptake in the capillaries and veins are highly correlated with that of the arteries. In addition, there is minimal difference in contrast uptake of different portions inside an artery or a vein. The components corresponding to an artery or a vein are highly correlated and

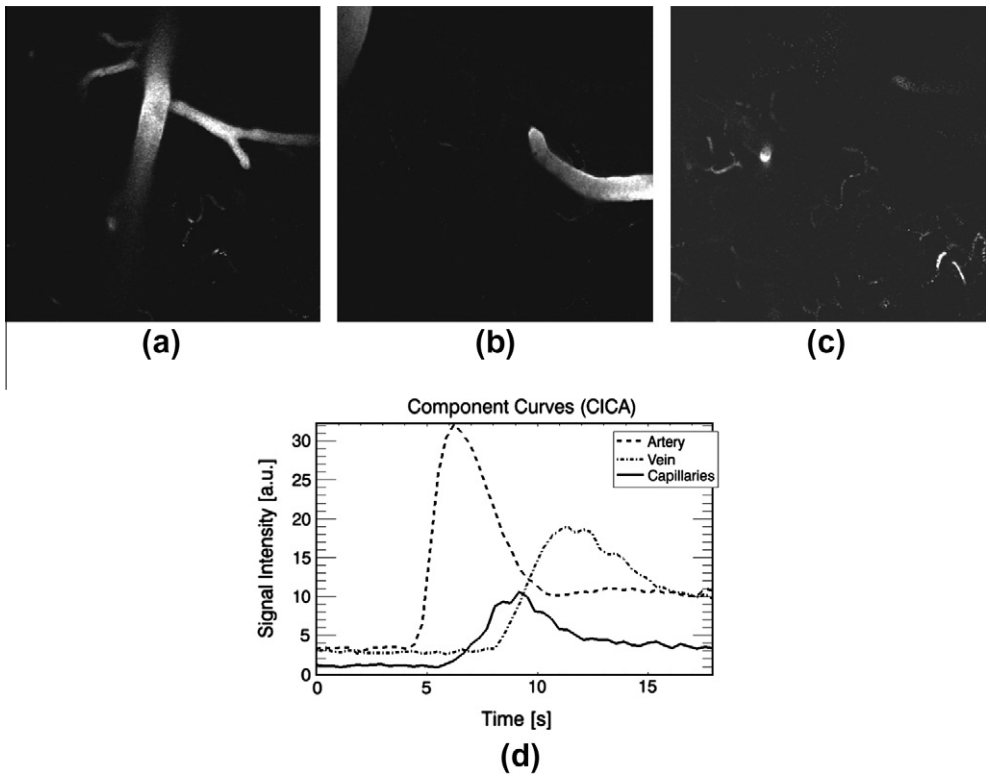

Fig. 9. The ICs and their corresponding component curves resulting from applying the CICA to the rat brain dataset: (a) artery, (b) vein, (c) capillaries and (d) component curves corresponding to the 3 ICs. The signal intensities are shown in arbitrary units (a.u.).

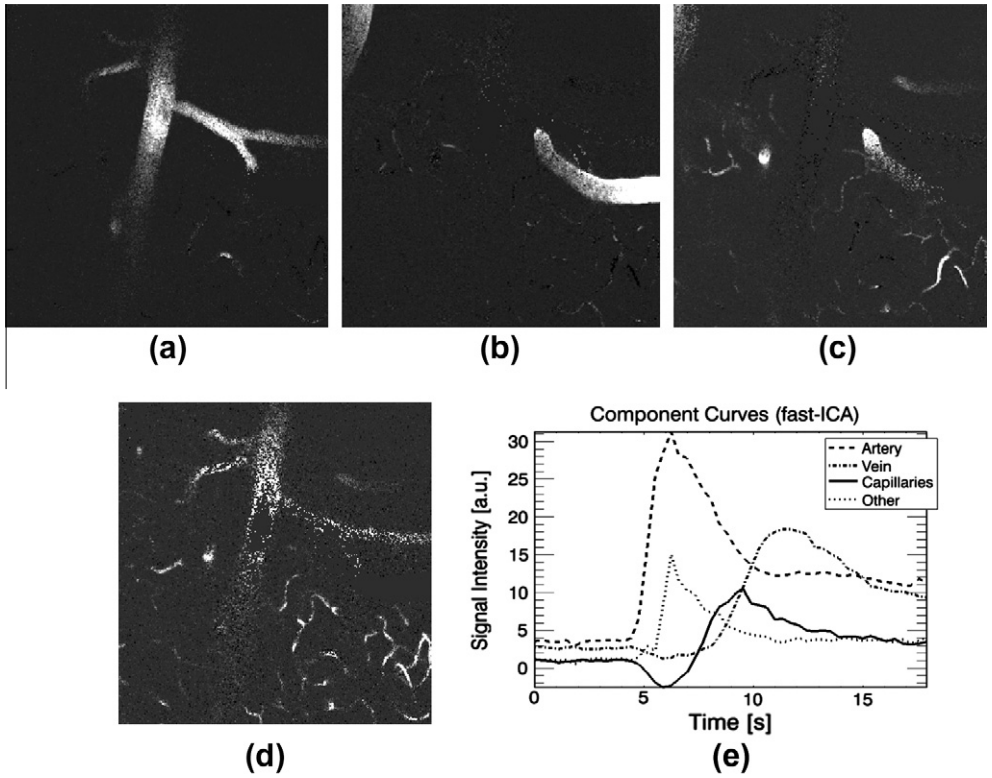

Fig. 10. The ICs and their corresponding component curves resulting from applying the fast-ICA to the rat brain dataset: (a) artery, (b) vein, (c) capillaries, (d) other artifacts and (e) component curves corresponding to the four ICs.

therefore, they are always extracted as one IC, this is shown in Figs. 3–5.

The proposed CICA, fast-ICA and conventional pixel by pixel model fitting were applied to the synthesized dynamic data. Since

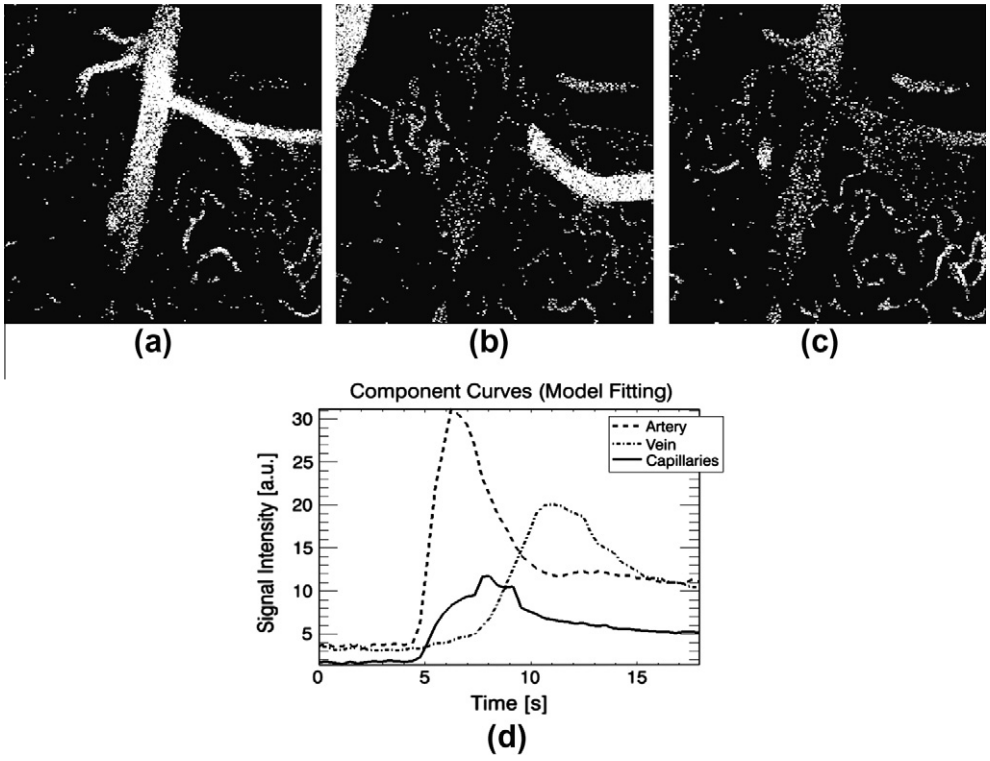

Fig. 11. The binary images and their corresponding component curves generated from the time to peak map obtained using conventional pixel by pixel model fitting to the rat brain dataset: (a) artery, (b) vein, (c) capillaries and (d) component curves corresponding to the three parametric maps.

Table 5
The mean squared distance of the normalized component curves from their gamma variate fit for both CICA and fast-ICA techniques.

	First component (artery)	Second component (capillaries)	Third component (vein)
CICA	0.0038	0.0077	0.0038
Fast-ICA	0.0085	0.0083	0.0048

Table 6
The time to peak extracted from the results of the three methods for each component curve.

	First component (artery)	Second component (capillaries)	Third component (vein)
Model fitting	6.89 (s)	8.42 (s)	11.35 (s)
Fast-ICA	6.22 (s)	9.52 (s)	11.71 (s)
CICA	6.22 (s)	8.42 (s)	11.35 (s)

Table 7
The onset time extracted from the results of the three methods for each component curve.

	First component (artery)	Second component (capillaries)	Third component (vein)
Model Fitting	4.76 (s)	4.03 (s)	6.22 (s)
Fast-ICA	4.39 (s)	7.32 (s)	8.42 (s)
CICA	4.39 (s)	6.22 (s)	7.32 (s)

ICA is a stochastic approach, it returns different results each time depending on the starting point. Thus; in fast-ICA each IC may be extracted in an arbitrary step depending on the starting point and to extract all the components that follow the model it is required to extract all ICs and then manually select the ones that follow the model. Due to presence of the constraints, this is not the case in CICA and the ICs that follow the model are separated from others. As expected the CICA successfully extracted the 3 ICs that followed the gamma variate model (artery, vein and capillaries) in the first 3 ICs, and then stopped the algorithm as there were no model-following components. This demonstrates the capability of the CICA in separating model-following ICs from artifact without requirement for user intervention. Whereas fast-ICA required extraction of up to 6 ICs to estimate the model-following ICs and manual selection of the model-following curves was required to separate them from artifacts. Moreover, Figs. 6 and 7 show that ignoring the effect of centering the data on the estimated component curves results in less accurate curves. Adding the mean effect to the curves show great improvement in the accuracy of the extracted curves and as can be seen in these figures, extracted curves follow the original curves better after adding mean to the data. Fig. 8 shows that the conventional pixel by pixel model fitting does not provide accurate results and in presence of noise ICA based techniques have better performance.

It can be seen in the images given in Figs. 3 and 4 that the components extracted using CICA are less noisy than those obtained using fast-ICA, and that there is better separation of arteries, veins and capillaries using CICA. In the fast-ICA results capillaries can also be seen in the ICs that represent the artery and vein, and more interference from the arterial and venous components can be seen in the capillary image. These interferences have been reduced in CICA as can be seen in Fig. 3 and 4 as well as Table 3. Fig. 5 shows the components of model fitting technique are very noisy and large areas of the background are also identified as the vessels. This poor performance is also reflected in the high PI (performance index) of the model fitting curves.

Another advantage of incorporating constraints to the ICA is that the curves extracted by CICA follow the gamma variate function better and consequently provide physiologically more meaningful curves compared to the results of fast-ICA. This can be

seen visually in Figs. 6 and 7 and also, from Table 1, the MSE between the estimated and the original curves are smaller in case of CICA for all ICs. Furthermore, as reported in Table 2, the SNR of all images are higher for CICA compared to fast-ICA and according to Table 3 the SIR of most of the ICs are higher for CICA and the mean SIR is higher for CICA compared to fast-ICA. The simulation results also show that using CICA improves (smaller PI) the separation performance of ICA and segments the image sequence into its main components more accurate than ICA without constraint (larger PI).

4.2. Experimental study

The results of applying CICA to the DCE images of rat brain microcirculation showed improvement in the performance of the algorithm compared to other existing algorithms. Similar to the simulation study, the algorithm extracted 3 ICs corresponding to the artery, vein and the capillaries and then stopped. As can be seen in Fig. 9 the curves follow a gamma variate function and the field of view is well segmented. Fast-ICA however extracted 4 ICs (Fig. 10) showing that it was unable to distinguish between ICs that follow the model from other ICs.

Moreover, comparing the time to peak (Table 6) and the onset time (Table 7) of all three components of the three different methods (CICA, fast-ICA and model fitting) shows the results of CICA match with the model fitting data better than the results of fast-ICA. In addition the better performance of CICA in following a gamma variate model can be seen visually in Figs. 9d and 10e. This has also been quantified and reported in Table 5 by measuring the mean squared error (MSE) between the extracted component curves of each method and their gamma variate function fit. As such, the component curves extracted using CICA provide components that are physiologically more meaningful compared to fast-ICA. Furthermore, the segmented ICs of CICA are less noisy and show less interference from other ICs. In Figs. 9–11 the order of ICs showing up in the algorithms are not preserved.

5. Conclusions

Separating artery, vein and capillaries in a DCE image dataset followed by extracting their intensity-time curves is essential in understanding the hemodynamics of cerebral microvasculature as the blood passes through artery to small vessels and capillaries and finally to the vein. The initial stage to extract this information is detecting artery and vein and also separating them from smaller capillaries in the field of view. Dynamic information such as onset time and time to peak can be used to detect vessel type and also extract information about microvasculature hemodynamics. This information can be inferred from the intensity-time curves of contrast agent uptake in each vessel.

In this study a constrained ICA (CICA) technique was developed to separate different vessels (artery, vein and capillaries) in the image sequence and extract their intensity-time (component) curves. The new method combines the two existing analysis techniques, the model based technique and the data driven technique. The model based technique used here is pixel by pixel gamma variate model fitting. This model provides *a priori* information that is used to set the constraint in CICA. The data driven technique is independent component analysis (ICA). We have chosen to use the fast-ICA algorithm with negentropy (Hyvärinen and Oja, 2000) here as it is the most commonly used method and its implementation is very stable (Hyvärinen et al., 2001). In future work we will investigate other approaches and cost functions to determine the optimal algorithm for this application. Although we are using a non-negativity constraint our method is not the same as non-negative matrix factorization (NMF) and there are significant differences in the two approaches.

Both experimental and simulation studies were conducted to assess the performance of CICA compared to the two available methods. CICA was demonstrated to perform better than the most commonly used classical ICA algorithm (fast-ICA) technique in terms of providing better ICs as the signal to noise ratio of the extracted images was higher (higher SNR), it also extracted physiologically more meaningful component curves as the error between the extracted curves and their gamma variate fits were smaller and the segmented images showed less interference from other ICs (higher SIR). In addition, the MSE between the extracted and actual curves in simulation study was smaller.

CICA showed better performance compared to model based technique (generating parametric maps with pixel by pixel curve fitting) in handing noise. CICA, as shown in Fig. 9, better segmented the artery, vein and capillaries compared to the model based technique shown in Fig. 11. Furthermore, the time to peak and onset time values extracted from CICA were closer to those of pixel by pixel curve fitting compared to results of fast-ICA technique.

Acknowledgement

The authors would like to thank the Natural Sciences and Engineering Research Council of Canada (NSERC) for supporting this work.

Appendix. The non-negativity constraint

We call a source **s** non-negative if it holds the following two properties:

1. All elements of $\boldsymbol{s_i}$ are greater than zero: $\boldsymbol{P}(\mathbf{s}_i < 0) = 0$ where $\boldsymbol{P}(\cdot)$ is the probability density function.
2. The source $\mathbf{s}_i$ is well grounded: $\{\forall \boldsymbol{\varepsilon} > 0 | \boldsymbol{P}(\mathbf{s}_i < \varepsilon) > 0\}$, i.e. the pdf of $\mathbf{s}_i$ is non-zero for all positive values of $\boldsymbol{\varepsilon}$.

Suppose that we have **Y**, a mixture of the source signals, such that $\mathbf{Y} = \mathbf{US}$, where $\mathbf{S} = [\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_M]^T$, a matrix of M non-negative well grounded independent unit-variance sources and **U** is a orthonormal mixing matrix i.e. $\mathbf{UU}^T = \mathbf{I}$. It was proved in Plumbley (2002) that **U** is a permutation matrix if and only if **Y** is non-negative. Thus, to estimate the independent sources, considering the fact that the source signals in ICA can be estimated up to a scaling and permutation, it is enough to find any orthonormal transformation $\mathbf{Y} = \mathbf{WZ}$ such that **Y** is non-negative (**Z** is the whitened but not centered data).

This problem can be related to non-linear PCA and the independent sources can be estimated through non-linear PCA rules (Plumbley, 2003; Plumbley and Oja, 2004). If we write the MSE criterion for non-linear PCA (Xu, 1993), for the whitened mixture **Z** we have:

$$e_{EMS} = \mathbf{E}\|\mathbf{Z} - \mathbf{W}^T f(\mathbf{WZ})\|^2$$

where **W** is the unmixing matrix and $\boldsymbol{f}(\cdot)$ is a non-linear function. Since **W** is orthonormal (data is whitened), the criterion can be rewritten as:

$$e_{EMS} = \mathbf{E}\|\mathbf{Y} - f(\mathbf{Y})\|^2 \tag{A.1}$$

where $\mathbf{Y} = \mathbf{WZ}$. Eq. (A.1) can be rewritten as (A.2) (Oja, 1999) which has been related to some forms of ICA cost functions in Oja (1999):

$$e_{EMS} = \sum_{i=1}^{N} \mathbf{E}\{[\mathbf{y_i} - \boldsymbol{f}(\mathbf{y}_i)]^2\} \tag{A.2}$$

If we use rectification function as the non-linear function $\boldsymbol{f}(\cdot)$ we have:

$$e_{EMS} = \sum_{(i=1)}^{N} \mathbf{E}\left\{[\mathbf{y_i} - \boldsymbol{f}_+(\mathbf{y_i})]^2\right\} = \sum_{(i=1)}^{N} \mathbf{E}\{[\mathbf{y_i^2}|\mathbf{y_i} > 0\} \tag{A.3}$$

where $\boldsymbol{f}_+(\mathbf{y_i}) = \max\{\mathbf{y_i}, 0\}$. The value of e_{EMS} in Eq. (A.3) is clearly zero if $\mathbf{y_i}$ is non-negative with probability one (Plumbley and Oja, 2004). However in the constrained ICA algorithm that is used in this article the cost function is:

$$\begin{array}{ll} \text{maximize} & \boldsymbol{J}(\mathbf{y}) \approx \rho[E\{\boldsymbol{G}(\mathbf{y})\} - E\{\boldsymbol{G}(\mathbf{y_{Gauss}})\}]^2 \\ \text{Subject to} & \mathbf{y}'(\mathrm{i}) \geqslant 0 \quad \text{(non-negativity)} \end{array} \tag{A.4}$$

where ρ is a positive constant, $\boldsymbol{G}(\cdot)$ is a non-quadratic function. The gamma variate constraint part of the cost function is not shown here as it is added to the system using Lagrange multipliers and thus the non-negativity constraint can be proved without loss of generality.

At every iteration, the non-negativity constraint is applied by first finding $\Delta\mathbf{w}_k$ using the Newton-like maximization approach such that it maximizes the first term in the ICA cost function. Let us call this updated unmixing vector $\mathbf{w_k}$. Then the corresponding IC ($\mathbf{y'_k}$) is calculated by multiplying the updated unmixing vector ($\mathbf{w_k}$) by the original mixed data or $\mathbf{X}$ (not whitened and not centered). The rectification function is applied to ($\mathbf{y'_k}$) and then ($\mathbf{y'_{+k}}$) is multiplied by the pseudo inverse of $\mathbf{X}$ to give the non-negative unmixing vector $\mathbf{w'_k}$. This updated vector is then normalized and orthogonalized with respect to previously estimated unmixing vectors and is used in the next iteration. The major importance of this method is that $\boldsymbol{J}(\mathbf{y_{+k}})$ is twice differentiable with respect to $\mathbf{w'_k}$ while is it not differentiable with respect to $\mathbf{w_k}$.

Thus the update rule of the optimization algorithm ($\mathbf{w_{k+1}} = \mathbf{w}_k + \Delta\mathbf{w}_k$) is modified to the following

$$\mathbf{w}_{k+1} = \mathbf{w}'_k + \Delta\mathbf{w}_k \tag{A.5}$$

therefore; we need to show that the two update rules converge. The update rule for the Newton method clearly optimizes the cost function and it suffices to show that the modified update rule (A.5) optimizes the cost function too. $\mathbf{w'_k}$ is clearly the same as $\mathbf{w_k}$ if all elements of $\mathbf{y'_k}$ are non-negative ($\mathbf{y'_k} \geqslant 0$). Thus suffices to check the optimization for $\mathbf{y'_k} < 0$.

$$\{\forall \mathbf{y}_k(i) | \cosh(\mathbf{y}_k(i)) \geqslant 1\}$$

$$\Rightarrow \{\forall \mathbf{y}_k(i) | \log(\cosh(\mathbf{y}_k(i))) \geqslant 0\}$$

$$\Rightarrow \mathbf{E}\{\log(\cosh(\mathbf{y}_{+k}(i)))\} \leqslant \mathbf{E}\{\log(\cosh(\mathbf{y}_k(i)))\}$$

Taking into account that for zero mean and equal variances, Negentropy takes its largest value for a Gaussian signal i.e. $\mathbf{E}\{\boldsymbol{G}(\mathbf{y_{Gauss}})\} \geqslant \mathbf{E}\{\boldsymbol{G}(\mathbf{y})\}$, we have;

$$(\mathbf{E}\{\boldsymbol{G}(\mathbf{y}_{+k})\} - \mathbf{E}\{\boldsymbol{G}(\mathbf{y_{Gauss}})\})^2 \geq (\mathbf{E}\{\boldsymbol{G}(\mathbf{y}_k)\} - \mathbf{E}\{\boldsymbol{G}(\mathbf{y_{Gauss}})\})^2$$
$$\text{thus} : \boldsymbol{J}(\mathbf{y}_{+k}) \geq \boldsymbol{J}(\mathbf{y}_k)$$

Thus any $\Delta\mathbf{w_k}$ that maximizes $\boldsymbol{J}(\mathbf{y}_k)$ also maximizes ($\boldsymbol{J}(\mathbf{y}_{+k})$and since at the end of iterative approach $\mathbf{y}'_\infty$ is supposed to be non-negative ($\mathbf{y}'_\infty \geqslant 0$) the two update rules converge towards each other. Thus (A.5) maximizes the Negentropy and thus its results are as independent as possible.

References

Ahearn, T.S., Staff, R.T., Redpath, T.W., Semple, S.I.K., 2005. The use of the Levenberg–Marquardt curve-fitting algorithm in pharmacokinetic modelling of DCE-MRI data. Physics in Medicine and Biology 50 (9), N85–N92.

Almeida, L.B., 2004. Linear and nonlinear ICA based on mutual information – the MISEP method. Signal Processing 84 (2), 231–245.

Araujo, A., Giné, E., 1980. The Central Limit Theorem for Real and Banach Valued Random Variables. Wiley, New York.

Barber, D.C., 1980. The use of principal components in the quantitative analysis of gamma camera dynamic studies. Physics in Medicine and Biology 25 (2), 283–292.

Bertsekas, C.P., 1982. Constrained Optimization and Lagrange Multiplier Methods. Academic, New York.

Biswal, B.B., Ulmer, J.L., 1999. Blind source separation of multiple signal sources of fMRI data sets using independent component analysis. Journal of Computer Assisted Tomography 23 (2), 265–271.

Buvat, J., Benali, H., Frouin, F., Basin, J.P., Di Paola, R., 1993. Target apex-seeking in factor analysis of medical image sequences. Physics in Medicine and Biology 38 (1), 123–137.

Calamante, F., Gadian, D.G., Connelly, A., 2000. Delay and dispersion effects in dynamic susceptibility contrast MRI: simulations using singular value decomposition. Magnetic Resonance in Medicine 44 (3), 466–473.

Calhoun, V.D., Adali, T., 2006. Unmixing fMRI with independent component analysis. IEEE Engineering in Medicine and Biology Magazine 25 (2), 79–90.

Chaigneau, E., Oheim, M., Audinat, E., Charpak, S., 2003. Two-photon imaging of capillary blood flow in olfactory bulb glomeruli. Proceedings of the National Academy of Sciences of the United States of America 100 (22), 13081–13086.

Chen, Y., Rege, M., Dong, M., Hua, J., 2008. Non-negative matrix factorization for semi-supervised data clustering. Knowledge and Information Systems 17 (3), 355–379.

Comon, P., 1994. Independent component analysis: a new concept? Signal Processing 36 (3), 287–314.

Denk, W., Strickler, J.H., Webb, W.W., 1990. Two-photon laser scanning fluorescence microscopy. Science 248 (4951), 73–76.

Erdogmus, D., Hild II, K.E., Rao, Y.N., Príncipe, J.C., 2004. Minimax mutual information approach for independent component analysis. Neural Computation 16 (6), 1235–1252.

Fenstermacher, J., Nakata, H., Tajima, A., Lin, S., Otsuka, T., Acuff, V., et al., 1991. Functional variations in parenchymal microvascular systems within the brain. Magnetic Resonance in Medicine 19 (2), 217–220.

Helmchen, F., Fee, M.S., Tank, D.W., Denk, W., 2001. A miniature head-mounted two-photon microscope: high-resolution brain imaging in freely moving animals. Neuron 31 (6), 903–912.

Hesse, C.W., James, C.J., 2006. On semi-blind source separation using spatial constraints with applications in EEG analysis. IEEE Transactions on Biomedical Engineering 53 (12), 2525–2534.

Hutchinson, E.B., Stefanovic, B., Koretsky, A.P., Silva, A.C., 2006. Spatial flow-volume dissociation of the cerebral microcirculatory response to mild hypercapnia. NeuroImage 32 (2), 520–530.

Hyvärinen, A., 1999a. Survey on independent component analysis. Neural Computing Surveys 2, 94–128.

Hyvärinen, A., 1999b. Fast and robust fixed-point algorithms for independent component analysis. IEEE Transactions on Neural Networks 10 (3), 626–634.

Hyvärinen, A., Oja, E., 2000. Independent component analysis: algorithms and applications. Neural Networks 13 (4–5), 411–430.

Hyvärinen, A., Karhunen, J., Oja, E., 2001. Independent Component Analysis. Wiley Interscience.

Jang, G., Lee, T., 2004. A maximum likelihood approach to single-channel source separation. Journal of Machine Learning Research 4 (7–8), 1365–1392.

Kim, J., Mueller, C.W., 1978. Methods of Extracting Initial Factors. Factor Analysis: Statistical Methods and Practical Issues. SAGE Publications, Inc.

Kleinfeld, D., Mitra, P.P., Helmchen, F., Denk, W., 1998. Fluctuations and stimulus-induced changes in blood flow observed in individual capillaries in layers 2 through 4 of rat neocortex. Proceedings of the National Academy of Sciences of the United States of America 95 (26), 15741–15746.

Koenig, M., Kraus, M., Theek, C., Klotz, E., Gehlen, W., Heuser, L., 2001. Quantitative assessment of the ischemic brain by means of perfusion-related parameters derived from perfusion CT. Stroke 32 (2), 431–437.

Koh, T.S., Yang, X., Bisdas, S., Lim, C.C.T., 2006. Independent component analysis of dynamic contrast-enhanced computed tomography images. Physics in Medicine and Biology 51 (19), N339–N348.

Lee, D.D., Seung, H.S., 1999. Learning the parts of objects by non-negative matrix factorization. Nature 401 (6755), 788–791.

Lin, Q., Zheng, Y., Yin, F., Liang, H., Calhoun, V.D., 2007. A fast algorithm for one-unit ICA-R. Information Sciences 177, 1265–1275.

Lu, W., Rajapakse, J.C., 2005. Approach and applications of constrained ICA. IEEE Transactions on Neural Networks 16 (1), 203–212.

Lu, W., Rajapakse, J.C., 2006. ICA with reference. Neurocomputing 69 (16–18), 2244–2257.

Martel, A.L., Moody, A.R., Allder, S.J., Delay, G.S., Morgan, P.S., 2001. Extracting parametric images from dynamic contrast-enhanced MRI studies of the brain using factor analysis. Medical Image Analysis 5 (1), 29–39.

McKeown, M.J., Sejnowski, T.J., 1998. Independent component analysis of fMRI data: examining the assumptions. Human Brain Mapping 6 (5–6), 368–372.

Mehrabian, H., Lindvere, L., Stefanovic, B., Martel, A.L., 2010. A temporally constrained ICA (TCICA) technique for artery–vein separation of cerebral microvasculature. Proceedings of SPIE, 7626.

Nijran, K.S., Barber, D.C., 1986. A completely automatic method if processing 131I-labelled Rose Bengal dynamic liver studies. Physics in Medicine and Biology 31 (5), 563–570.

Novey, M., Adali, T., 2005. ICA by maximization of nongaussianity using complex functions. Machine Learning for Signal Processing, 21–26.

Oja, E., (1999). Nonlinear PCA criterion and maximum likelihood in independent component analysis. In: Proceedings of Intenranational Workshop Independent Component Analysis and Signal Separation (1CA'99), France, pp. 143–148.

Patlak, C.S., Blasberg, R.G., Fenstermacher, J.D., 1984. An evaluation of errors in the determination of blood flow by the indicator fractionation and tissue equilibration (kety) methods. Journal of Cerebral Blood Flow and Metabolism 4 (1), 47–60.

Plumbley, M., 2002. Conditions for nonnegative independent component analysis. IEEE Signal Processing Letters 9 (6), 177–180.

Plumbley, M., 2003. Algorithms for nonnegative independent component analysis. IEEE Transactions on Neural Networks 14 (3), 534–543.

Plumbley, M., Oja, E., 2004. A nonnegative PCA algorithm for independent component analysis. IEEE Transactions on Neural Networks 15 (1), 66–76.

Rosen, B.R., Belliveau, J.W., Buchbinder, B.R., McKinstry, R.C., Porkka, L.M., Kennedy, D.N., et al., 1991. Contrast agents and cerebral hemodynamics. Magnetic Resonance in Medicine 19 (2), 285–292.

Thompson Jr., H.K., Stramer, C.F., Whalen, R.E., Mcintosh, H.D., 1964. Indicator transit time considered as a gamma variate. Circulation Research 14 (6), 502–515.

Wu, X.Y., Liu, G.R., 2006. Independent Component Analysis of Dynamic Contrast-Enhanced Images: The Number of Components. Computational Methods. Springer, Netherlands, pp. 1111–1116.

Wu, X.Y., Liu, G.R., 2007. Application of independent component analysis to dynamic contrast-enhanced imaging for assessment of cerebral blood perfusion. Medical image analysis 11 (3), 254–265.

Xu, L., 1993. Least mean square error reconstruction principle for self-organizing neural-nets. Neural Networks 6 (5), 627–648.