Journal of Neuroscience Methods 216 (2013) 49–61

Contents lists available at SciVerse ScienceDirect

Journal of Neuroscience Methods

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

Computational Neuroscience

SACICA: A sparse approximation coefficient-based ICA model for functional magnetic resonance imaging data analysis

Nizhuan Wang, Weiming Zeng*, Lei Chen

Digital Image and Intelligent Computation Laboratory, Shanghai Maritime University, Shanghai 201306, China

HIGHLIGHTS

- A new proposed analysis model SACICA is based on the mixtures' sparse approximation coefficients and ICA.
- A 1D wavelet packet decomposition and a novel L_p norm are utilized to form the sparse approximation coefficients set.
- SACICA demonstrates stronger spatial source recovery ability than FastICA.
- SACICA has better detection sensitivity performance of functional signal than FastICA.
- SACICA is effective to do group evaluation combining with single-subject based ICA group analysis method.

ARTICLE INFO

Article history:
Received 26 December 2012
Received in revised form 24 February 2013
Accepted 19 March 2013

Keywords:
fMRI
Wavelet packet decomposition
Sparsity
L_p norm
ICA

ABSTRACT

Independent component analysis (ICA) has been widely used in functional magnetic resonance imaging (fMRI) data to evaluate the functional connectivity, which assumes that the sources of functional networks are statistically independent. Recently, many researchers have demonstrated that sparsity is an effective assumption for fMRI signal separation. In this research, we present a sparse approximation coefficient-based ICA (SACICA) model to analyse fMRI data, which is a promising combination model of sparse features and an ICA technique.

The SACICA method consists of three procedures. The wavelet packet decomposition procedure, which decomposes the fMRI data into wavelet tree nodes with different degrees of sparsity, is first. Then, the sparse approximation coefficients set formation procedure, in which an effective L_p norm is proposed to measure the sparse degree of the distinct wavelet tree nodes, is second. The ICA decomposition and reconstruction procedure, which utilises the sparse approximation coefficients set of the fMRI data, is last.

The hybrid data experimental results demonstrated that the SACICA method exhibited the stronger spatial source reconstruction ability with respect to the unsmoothed fMRI data and better detection sensitivity of the functional signal on the smoothed fMRI data than the FastICA method. Furthermore, task-related experiments also revealed that SACICA was not only effective in discovering the functional networks but also exhibited a better detection sensitivity of the visual-related functional signal. In addition, the SACICA combined with Fast-FENICA proposed by Wang et al. (2012) was demonstrated to conduct the group analysis effectively on the resting-state data set.

1. Introduction

Blood-oxygen-level-dependent (BOLD) functional magnetic resonance imaging (fMRI) is a powerful modality that detects functional connectivity among discrete cortical brain regions and can reveal the neural correlates of cognitive processes. Recently, functional connectivity detection using fMRI has received significant attention from many neuroscientists and computer scientists. After extensive study for more than one decade, scientists have developed several methods to detect and analyse functional connectivity. In general, these methods can be classified into two categories: model-based methods and data-driven methods (Li et al., 2009). The straight-forward model-based method is based on the hypothesis of first localising knowledge concerning the seed region, called the region of interest (ROI), which is subjected to the correlation of the time courses between the seed region and other selected brain voxels. Several functional networks have been investigated using the seed-based method such as motor,

---

* Corresponding author at: College of Information Engineering, Shanghai Maritime University, Shanghai 201306, China. Tel.: +86 021 38282873; fax: +86 021 38282800.
E-mail address: zwmcd@yahoo.com (W. Zeng).

0165-0270/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jneumeth.2013.03.014

visual, auditory and even certain cognitive networks (Biswal et al., 1995, 1997; Cordes et al., 2001). However, the accuracy of the patterns demonstrated in this model-based method is susceptible to the prior neuroscience knowledge or experience, which is a limitation in detecting all the resting-state networks (RSNs) of resting-state fMRI data without specific priors. In contrast, the data-driven methods have overcome the prior restrictions involved with ROIs. Generally, the data-driven methods can be divided into two types: decomposition-based methods, such as principal component analysis (Baumgartner et al., 2000; Kiviniemi et al., 2000) such as singular value decomposition (Worsley et al., 2005) and independent component analysis (ICA) (McKeown et al., 1998; Biswal and Ulmer, 1999), and clustering-based methods such as fuzzy clustering analysis (Scarth et al., 1995; Golay et al., 1998; Fadili et al., 2000) and hierarchical clustering analysis (Cordes et al., 2002). Among these data-driven methods, ICA became a popular method for detecting functional connectivity without any prior spatial or temporal pattern since McKeown et al. (1998) introduced the ICA method to fMRI analysis. ICA can be further divided into spatial ICA (sICA) (McKeown et al., 1998) and temporal ICA (tICA) (Biswal and Ulmer, 1999) based on spatial independence and temporal independence, respectively. However, no method is perfect. First, the criterion of utilising the sICA or tICA appears to be task-dependent. Pekar et al. (2001) reported that sICA and tICA could have diverging results depending on the characteristics of the underlying signals to be estimated. Second, both types of ICA methods only considered the spatial or temporal independence and ignored the additional property of the sources such as spatial regularity (Valente et al., 2009) and sparsity. Despite certain drawbacks, great developments have been achieved in ICA in subsequent years. Beckmann and Smith (2004) proposed a probabilistic independent component analysis method as a further development, which performed probabilistic modelling and achieved an asymptotically unique decomposition of fMRI data. To investigate the commonality of the functional connectivity across a group of subjects, several ICA-based group analysis methods have been developed such as the multi-session temporal concatenated group ICA method (Calhoun et al., 2001), fully exploratory network ICA (FENICA) method (Schöpf et al., 2010) and Fast-FENICA method (Wang et al., 2012).

In addition to the aforementioned analysis methods, wavelet decomposition, which simultaneously considers the frequency–time characteristics in fMRI data analysis, has been developed as a useful tool. Bullmore et al. (2003, 2004) suggested that the wavelet domain was a rich source of new concepts and techniques to enhance the power of statistical analysis of human fMRI data. This domain was based on wavelet coefficients resampling, a wavelet-based linear estimation model for signals and errors, wavelet shrinkage in frequency and Bayesian frameworks for multiple hypothesis testing. To provide an efficient representation of heterogeneous brain structures, Long et al. (2004) presented a spatiotemporal wavelet model combining a stimulus-convolved haemodynamic signal with correlated noise, which resulted in a reduced mean-squared error and more localised fMRI activation maps compared with the use of standard wavelets or a smoothing technique. The temporal clustering analysis of fMRI data combined with the wavelet domain has been developed in recent years. For example, Sato et al. (2007) introduced a combination of the discrete wavelet transform (DWT) and the classification expectation maximisation (CEM) algorithm as an effective method of fMRI analysis when the general linear model (GLM) assumptions were questionable or the signal-to-noise ratio (SNR) was low. Costafreda et al. (2009) further proposed an integrated framework for Bayesian estimation and regularisation in wavelet space that allowed for voxel-wise hypothesis testing, which suggested that the wavelet-based approach was superior to Gaussian smoothing and preserved more accurate morphological information concerning the fMRI source signal. Recently, Khullar et al. (2011a) presented a wavelet-domain based framework (w-ICA) for noise removal of fMRI data at the single subject level, in which a multidirectional denoising scheme with a 3-dimensional stationary discrete wavelet transform was utilised. Compared with Gaussian smoothing, this method demonstrated certain significant improvements and preserved a more accurate signal shape. As a further development of w-ICA, Khullar et al. (2011b) used this framework to de-noise multi-group (healthy/patient) fMRI data, gaining higher specificity and increasing shape accuracy compared with Gaussian smoothing. The above studies imply that wavelet decomposition is a promising tool for analysing fMRI data.

The sparse property is a general assumption as previously mentioned, and many scholars have performed studies on the sparsity of the fMRI signal to help decode the functional connectivity between the cortical regions of the human brain. Alexander and Baumgartner (2001) demonstrated that optimal sparse representation could significantly benefit the de-noising of fMRI time courses. Georgiev et al. (2007) proposed the sparse component analysis (SCA) method, which revealed a stronger potential ability of dependent fMRI sources separation compared with ICA. In the same year, Flandin and Penny (2007) proposed a type of Bayesian fMRI data analysis with sparse spatial basis functions, which first projected the original to wavelet space, followed by threshing the small coefficients to increase the sparsity of the projected-back data. Recently, Ye et al. (2011) presented a data-driven sparse geo-statistical analysis clustering method and achieved better performance compared with GLM analysis. These studies have demonstrated that the sparsity of the fMRI signal is a useful feature for source separation. In addition to the fMRI signal separation field, other field studies (Zibulevsky and Pearlmutter, 2001; Zibulevsky and Zeevi, 2002; Kisilev et al., 2003; Bronstein et al., 2005) have demonstrated that the assumption of sparseness is also very powerful and could significantly improve the signal separation accuracy and the computation efficiency of existing ICA algorithms.

Based on the heuristic of the informed studies, in this study, an effective sparse approximation coefficient-based ICA (SACICA) model is presented, which is a novel model combining the sparse feature and ICA technique in fMRI data analysis. The remainder of this paper is organised as follows: the theory and method of SACICA will first be presented followed by the hybrid data and task-related data experimental tests, which were designed to evaluate the signal source recovery ability and detection sensitivity performance of the functional signal of SACICA compared with FastICA. Finally, SACICA combined with Fast-FENICA is performed on group evaluation on resting-state data set; the results and analysis will be presented together with interpretations and conclusions related to the advantages and limitations of this new data analysis model.

2. Theory and method

In this section, we will first present an overview of the theory concerning the sparse approximation coefficients of the mixture followed by the framework of SACICA. The details of the framework of SACICA include three steps: wavelet packet decomposition of the fMRI data, the sparse approximation coefficients set formation procedure and the ICA decomposition and reconstruction procedure. The sparse approximation coefficients set formation procedure of the fMRI data includes sparsity measurement and optimal sparse wavelet packet tree nodes selection.

2.1. Sparse approximation coefficients of the mixture

The functional connectivity detection problem of human brain activities has always been modelled as a type of blind source