NeuroImage 51 (2010) 183–205

Contents lists available at ScienceDirect

NeuroImage

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

Neuroelectric source imaging using 3SCO: A space coding algorithm based on particle swarm optimization and l_0 norm constraint

Peng Xu [a], Yin Tian [a,b], Xu Lei [a], Dezhong Yao [a,*]

[a] Key Laboratory for NeuroInformation of Ministry of Education, University of Electronic Science and Technology of China, Chengdu, 610054, China
[b] College of Bio-information, ChongQing University of Posts and Telecommunications, ChongQing, 400065, China

ARTICLE INFO

Article history:
Received 9 October 2009
Revised 23 January 2010
Accepted 30 January 2010
Available online 6 February 2010

Keywords:
Underdetermined system
Sparse solution
Particle swarm
Neuroelectric source imaging
Solution space compression

ABSTRACT

The electroencephalogram (EEG) neuroelectric sources inverse problem is usually underdetermined and lacks a unique solution, which is due to both the electromagnetism Helmholtz theorem and the fact that there are fewer observations than the unknown variables. One potential choice to tackle this issue is to solve the underdetermined system for a sparse solution. Aiming to the sparse solution, a novel algorithm termed 3SCO (Solution Space Sparse Coding Optimization) is presented in this paper. In 3SCO, after the solution space is coded with some particles, the particle-coded space is compressed by the evolution of particle swarm optimization algorithm, where an l_0 constrained fitness function is introduced to guarantee the selection of a suitable sparse solution for the underdetermined system. 3SCO was first tested by localizing simulated EEG sources with different configurations on a realistic head model, and the comparisons with minimum norm (MN), LORETA (low resolution electromagnetic tomography), l_1 norm solution and FOCUSS (focal underdetermined system solver) confirmed that a good sparse solution for EEG source imaging could be achieved with 3SCO. Finally, 3SCO was applied to localize the neuroelectric sources in a visual stimuli related experiment and the localized areas were basically consistent with those reported in previous studies.
© 2010 Elsevier Inc. All rights reserved.

Introduction

The scalp electroencephalogram (EEG) represents electrical activity generated by some ensembles of neurons within the brain. Estimating the locations and distributions of the underlying equivalent electric generators based on the scalp EEG is the neuroelectric or EEG inverse problem (Michel et al., 2004). Based on the distributed source assumption that the solution space consists of all the possible source positions, EEG inverse problem can be mathematically stated as (Pascual-Marqui et al., 1994; Silva et al., 2004; Wang et al., 1992),

$$Y = AX + \upsilon \quad (1)$$

where Y is the scalp EEG recordings of $M\times 1$, M is the number of scalp electrodes; A is the lead field matrix of $M\times N$, where N is the dimension size of the solution space Θ; X is the source vector to be estimated and υ is the noise induced in the recordings. In general, X consists of dipoles (Pascual-Marqui et al., 1994; Silva et al., 2004; Wang et al., 1992), yet charges (Yao, 1996, 2003; Yao et al., 2004) or potential (Grave de Peralta Menendez et al., 2000) may be another two possible choices. In this paper, the dipole model was considered. For the general EEG inverse problem, M is usually much smaller than N indicating an underdetermined system. According to the electromagnetism Helmholtz theorem (von Helmholtz, 1853) and the inherent characteristics of an underdetermined system, the EEG inverse problem lacks a unique solution because there are an infinite number of possible source configurations that could explain the measured recordings Y (Pascual-Marqui et al., 1994; Wang et al., 1992). To obtain a physiologically feasible solution, some possible and reasonable constraints are necessary, such as the assumption of a few dipoles in the spatio-temporal dipole fitting procedure (Scherg von Cramon, 1985) and the minimum norm constraint of the solution both utilized in the early effort (Hamalainen and Ilmoniemi, 1984; Wang et al., 1992). However, minimum norm (MN) solution favors the superficial source that the estimated position of a deep source will be shallower than the actual situation. As a step ahead, various weighted minimum norm solutions (WMNS) have been developed, such as the low resolution electromagnetic tomography (LORETA), and many efforts are still being paid to improve the spatial resolution of WMNS to a high level required for neurological research (Ding He, 2008; Pascual-Marqui, 1999; Yao and Dewald, 2005). Recently, the mixed norm constraints have been used in this field, and the reported results reveal that the advantages of these norms can be integrated into the solution by such penalty strategy (Valdes-Sosa et al., 2009; Vega-Hernandez et al., 2008). Besides the above penalty object function based approaches, other approaches like the RAP-MUSIC (Mosher et al., 1992), Markov chain Monte Carlo (MCMC) (Auranen et al., 2005; Nummenmaa et al., 2007) and variational Bayes (VB) (Wipf and Nagarajan, 2009) have been developed to solve EEG inverse problem

* Corresponding author.
E-mail address: dyao@uestc.edu.cn (D. Yao).

1053-8119/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2010.01.106

in other views. RAP-MUSIC is based on the decomposition of the signal and noise subspaces, which needs to involve a time interval of recordings (Mosher et al., 1992). The VB (Wipf and Nagarajan, 2009) and MCMC (Auranen et al., 2005; Nummenmaa et al., 2007) solve inverse problem by assuming the independent distributions of source currents and their variances, whose object function is then subject to maximizing a posteriori (MAP).VB and MCMC have been used to integrate the fMRI priors into the EEG inverse approach (Wipf and Nagarajan, 2009). The main contribution of current work is an extension of penalty object based approach, and we will mainly discuss the comparison among different penalty objects based approaches like MN, LORETA, l_1 norm and FOCUSS.

Studies using fMRI and micro-electrode techniques have demonstrated that the main neural electric activities are sparsely localized in the brain (Olshausen and Field, 1996; Op de Beeck et al., 2008; Simoncelli and Olshausen, 2001), thus a reasonable constraint is to use the sparse information of source distribution (Ding, 2009; Ding and He, 2008; Gorodnitsky and Rao, 1997; Yao and He, 2001). Recently, some researchers have tried to solve the EEG inverse problem for a sparse and local solution with a variety of methods such as the self-coherence enhancement algorithm (Yao and He, 2001), l_1 norm solution (i.e., the least absolute shrinkage selection operator (LASSO)) (Silva et al., 2004; Tibshirani, 1996), and l_p norm iterative sparse source (LPISS) (Xu et al., 2007). However, these methods are not stable and effective for their sensitivities to noise, source configuration and even the initial source distribution (Gorodnitsky and Rao, 1997; Michel et al., 2004; Pascual-Marqui, 1999; Silva et al., 2004).

Sparsity can be regarded as a physiology dependent constraint to the EEG inverse problem (Ding, 2009; Ding and He, 2008; Gorodnitsky and Rao, 1997; Silva et al., 2004). The original and most effective way to measure the sparsity of a signal is the l_0 norm, which is just to count

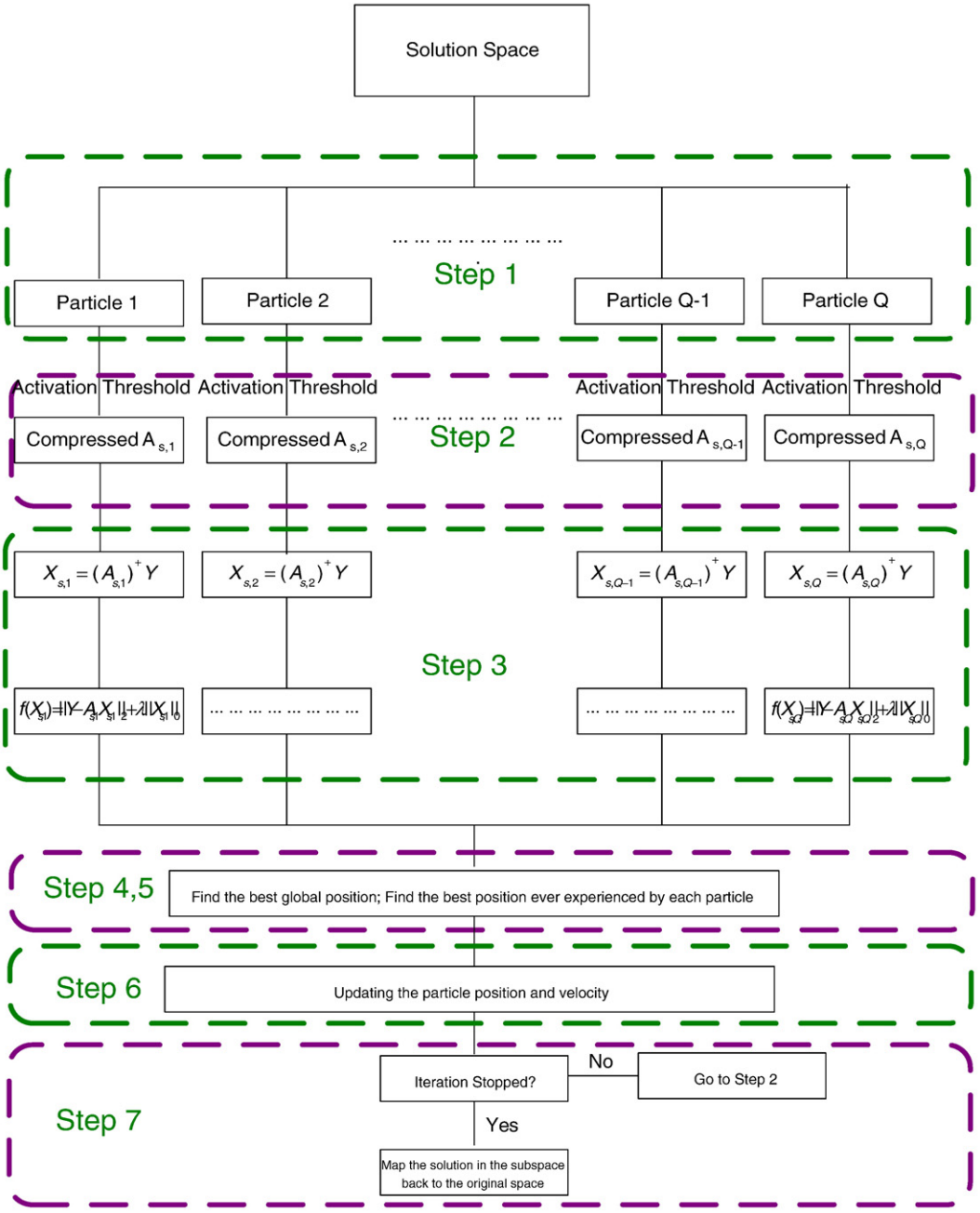
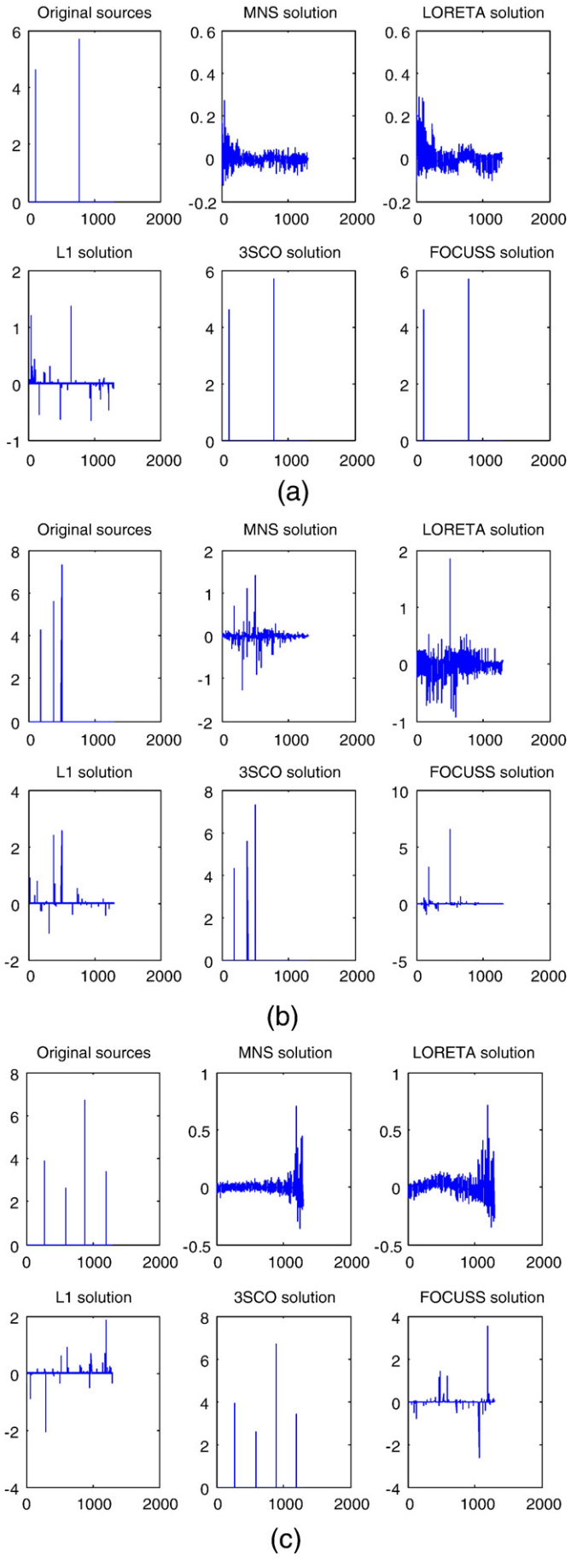


Fig. 1. Iteration diagram of 3SCO. Each dash square corresponds to the step introduced below.



the number of the non-null entries in the signal (Donoho and Elad, 2003; Malioutov et al., 2004). If the l_0 norm of the solution is taken as a constraint, we obtain a Lagrange multiplier expression of the inverse problem given in Equation (1) as (Donoho and Elad, 2003),

$$\min_X \|Y - AX\|_2 + \lambda \|X\|_0 \tag{2}$$

where λ is the sparsity-related penalty parameter. However, Equation (2) is not continuously differentiated and many gradient based effective optimization methods cannot be directly applied to solve this problem, so in practice, $l_p (p \leq 1)$ norm is usually taken to replace l_0 norm, where l_1 norm is mostly adopted for EEG inverse problem. However, the approximation with $p \leq 1$ norm to l_0 norm will lower the sparsity to some degree (Donoho and Elad, 2003; Malioutov et al., 2004). Furthermore, when using this approximation, the source strength is directly introduced into the object function by the term $\lambda \| X \|_p$. Accordingly, a superficial solution with smaller strength is more favorable to minimize the object function than the deep one with larger strength, if both of the solutions can generate the similar scalp potentials. Thus a solution biased toward the surface may still be induced. When using the l_0 norm measurement, the number of sources instead of the strength of sources will contribute to the object function and the bias toward surface can be theoretically reduced.

In Equation (2), the object function consists of two different components, i.e., the residual error between the estimated signal and the observed signal, and the penalty term for the sparsity. In an EEG inverse problem, it is reasonable to assume that the actual number of sources, i.e. the number of non-null entries in the solution X, is much smaller than M, the dimension of the observed signal Y. Therefore, many entries in the solution X are null and have no contribution to the observed signal Y, and accordingly, the corresponding columns in the lead field matrix A are redundant for the sources estimation, which may negatively influence the estimation. If the redundant columns in the lead field matrix can be removed as possible, Equation (1) can be converted to a non-underdetermined system or a system with fewer unknown variables, and the solving of such a system is much easier and more stable than solving of the original under-determined one (Ding and He, 2008; Srebro and Oguz, 1997; Tikhonov and Arsenin, 1977; Yao and He, 2001). So we aim to solve the problem of Equation (1) in an adequately compressed solution space for a sparse solution.

In this paper, suggested is a 3SCO (solution space sparse coding optimization) algorithm, where the evolution strategy of the particle swarm algorithm is adopted to evolve a swarm to hunt for a good compressed solution space of the EEG inverse problem. In the evolution procedure, an l_0 constrained fitness function was introduced to guarantee the algorithm to select a more suitable sparse solution for the original underdetermined system.

The remainder of the paper is structured as follows: the 3SCO algorithm is introduced in the section titled *3SCO algorithm of EEG inverse problem*. In the section *Simulation evaluation*, the algorithm is tested by localizing sources on a realistic head model, and the comparison with MN solution, LORETA, l_1 norm solution and FOCUSS under different source configurations and noise conditions is given in this section, too. The localization results for a real event related potential (ERP) recorded in a visual stimuli experiment are given in the section *Real data test*. Finally, discussions and conclusions conclude this paper in the section *Discussion and conclusion*.

Fig. 2. Source localization for different cortex source configurations. (a) Two sources; (b) three sources; (c) four sources. (d) Source mappings on the realistic head for the 2-source configuration in (b). (e) The PSO iterations for the three-source configuration. In panel d, the left panel represents the left hemisphere viewed from $+X$ and $-X$; the right panel represents the right hemisphere viewed from $-X$ and $+X$. The color bar in panel c is the same for all methods.

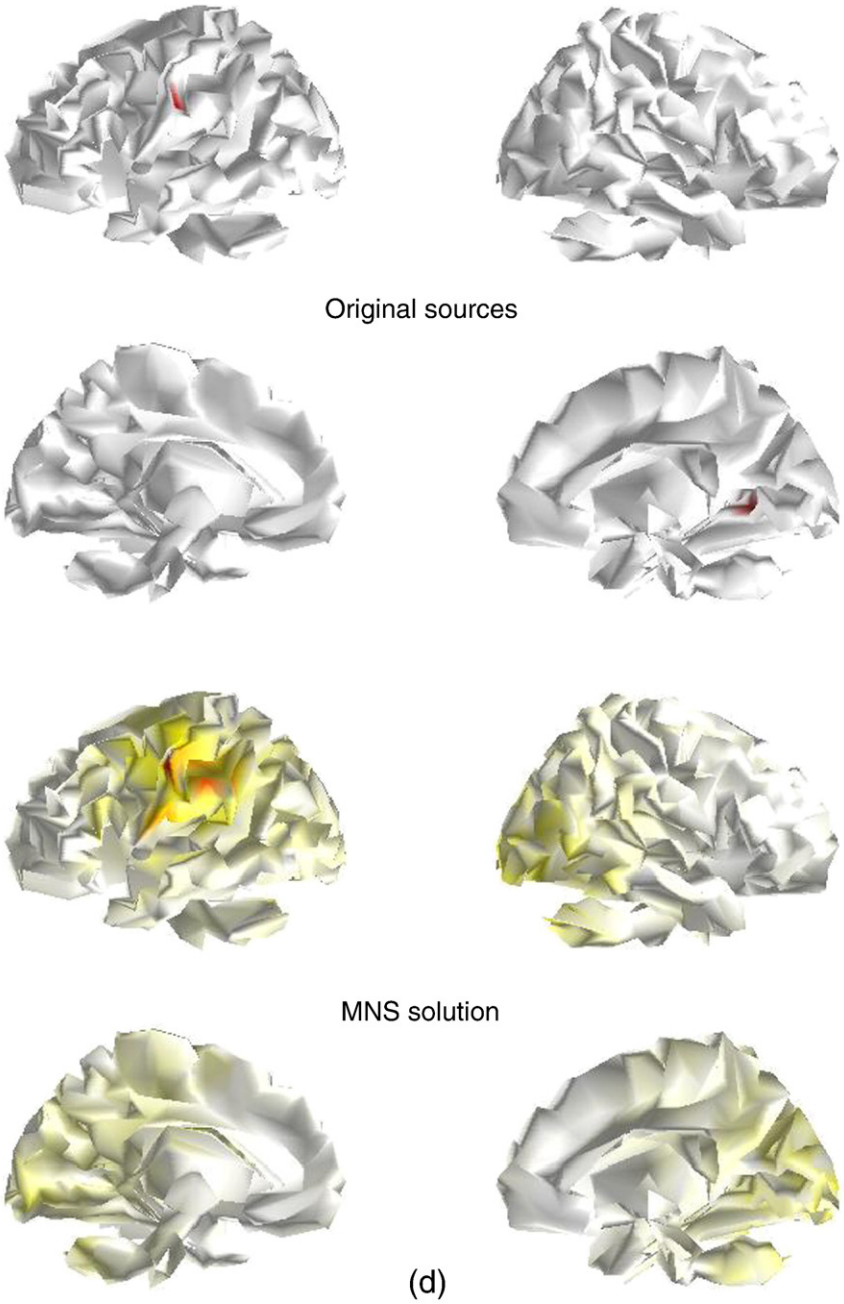


Fig. 2 (*continued*).

3SCO algorithm of EEG inverse problem

Compression of the solution space

As there are an infinite number of solutions satisfying the underdetermined system denoted in Equation (1), the compression of solution space would be a feasible step to reduce the system's underdetermination (Ding, 2009; Ding and He, 2008; Gorodnitsky and Rao, 1997; Grova et al., 2006; Pascual-Marqui, 1999; Pascual-Marqui et al., 1994; Silva et al., 2004; Yao and He, 2001). The intrinsic sparsity of source distribution means that there are only a small number of non-null entries in the solution space while most of the elements in the solution space have no contributions to the observed signal (Malioutov et al., 2004). Based on the sparsity prior, the solution space Θ can be adequately compressed by removing the redundant sub-solution space that has no contribution to the observed signal. Let Θ_s represent the compressed solution space having dimension K, and X_s represent the solution in space Θ_s, where Θ_s is supposed to contain the whole useful spatial information of the neuroelectric sources. The compressed format of Equation (1) in space Θ_s can be expressed as,

$$Y = A_s X_s \tag{3}$$

Based on the sparse assumption of the solution for an EEG inverse problem, it is reasonable to have a compressed space Θ_s with its dimension K smaller than M, and the ideal case is that K, the number of columns in A_s, is exactly equal to the number of sources. When K is smaller than M, the original underdetermined system is transformed to an overdetermined one and for most of the overdetermined systems, the column rank of the matrix A_s is nearly full and the

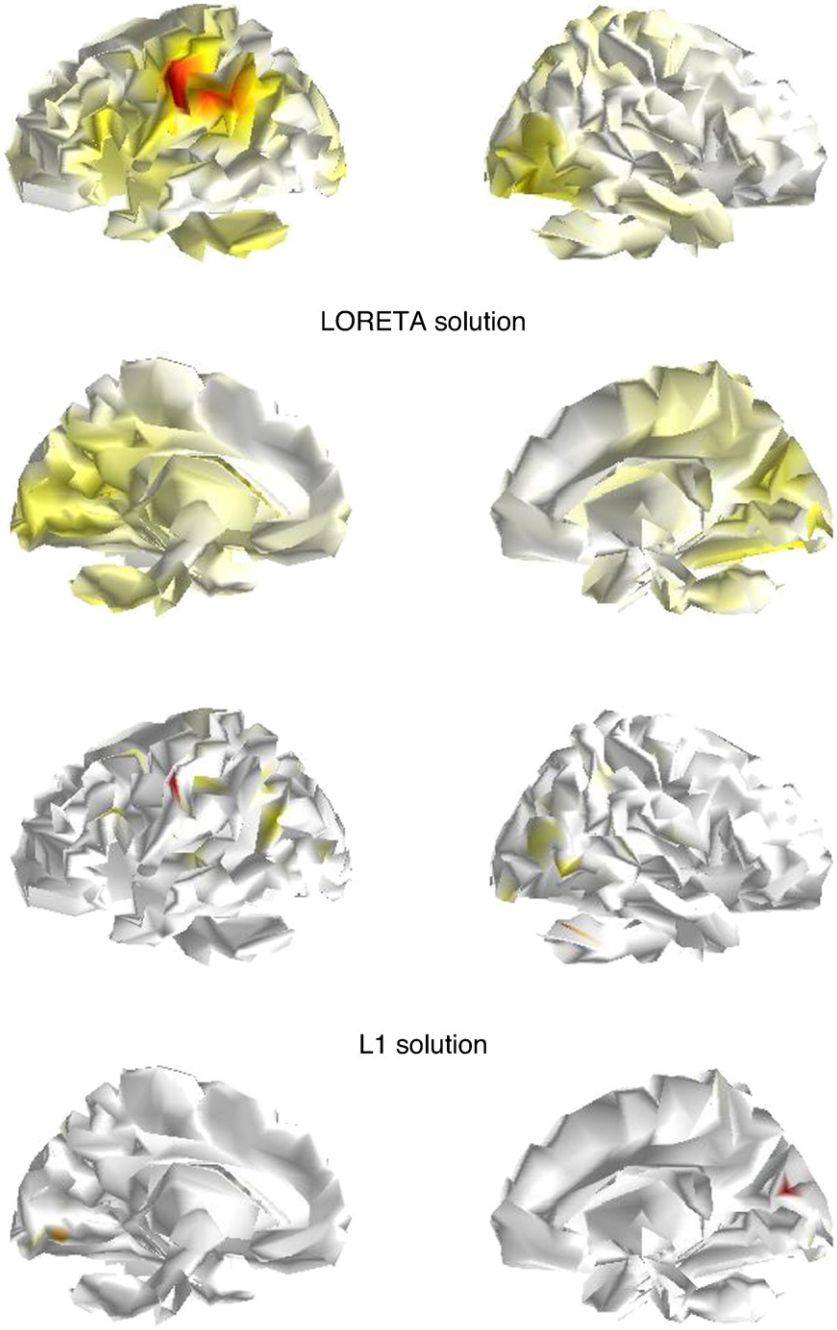


Fig. 2 (*continued*).

solutions are mathematically stable (Ding, 2009; Ding and He, 2008; Srebro and Oguz, 1997; Yao and He, 2001). Thus, in the compressed solution space Θ_s, it is straightforward to estimate the sources as,

$$X_s = (A_s)^{+} Y \tag{4}$$

where $(A_s)^{+}$ denotes the inverse operator of A_s. It should be noted that, other than the pseudo-inverse solution, such solution techniques as l_1 norm and FOCUSS can be used to estimate sources in the compressed space. Finally, the estimated X_s in the compressed space Θ_s is projected back to the original solution space Θ to obtain the possible source positions in the brain for physiological interpretation. Obviously, the compression of solution space can be simultaneously realized in the procedure to remove those redundant columns in the lead field matrix. For a real EEG inverse problem, the prior information about the actual positions of the activated neuroelectric sources is very poor, and if the compressed space Θ_s, a subspace of the intact solution space Θ, is not suitably chosen, the sources may not be correctly estimated (Ding, 2009; Srebro and Oguz, 1997; Xu et al., 2008; Yao and He, 2001). The choice of a proper compressed space is the crucial step of a successful EEG inverse approach, and that is the main point tackled in this work by a particle swarm coding and optimization as shown below.

Coding and compression of the solution space with particles

Particle swarm optimization (PSO) has been proved to be a powerful tool to search for the global optima (Eberhart and Shi, 1998; Hu et al., 2004; Kennedy and Eberhart, 1995; Parsopoulos and Vrahatis, 2002; Qiu et al., 2005; Wachowiak et al., 2004). The working mechanism of standard PSO is provided in Appendix. In this work, a

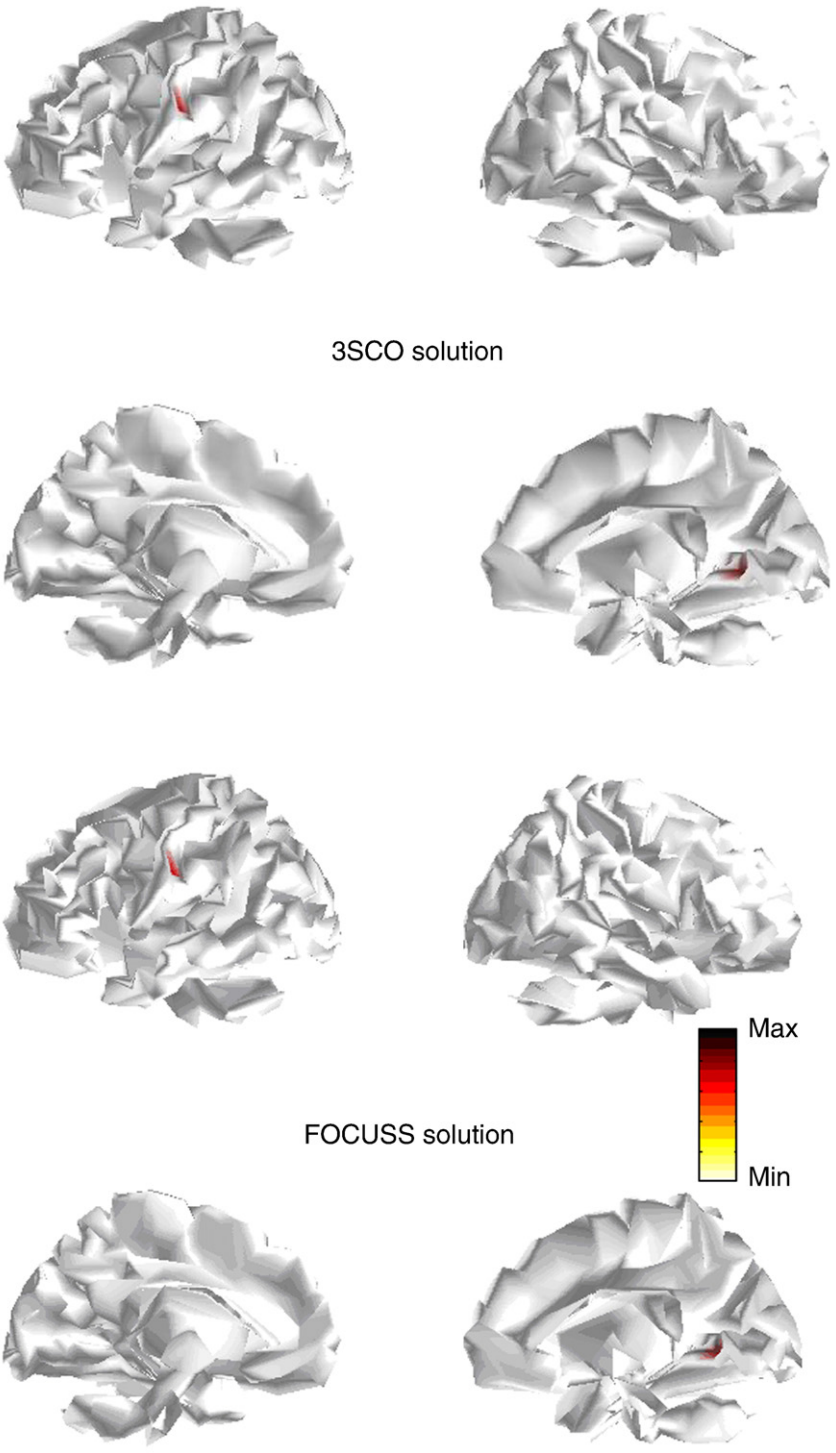


Fig. 2 (*continued*).

particle swarm is adopted to search for a compressed space Θ_s containing the necessary solution information as possible.

For the underdetermined system $Y=AX$, the solution space is encoded by a particle as,

$$\begin{gathered} A = (A(1), A(2), \ldots\ldots, A(N)) \\ \Downarrow \\ E_i = (e_i(1), e_i(2), \ldots\ldots, e_i(N)), V_i = (v_i(1), v_i(2), \ldots\ldots, v_i(N)) \end{gathered} \tag{5}$$

where $A(n)$ $(n=1,..,N)$ is the nth column of the lead field A. E_i with the same dimension as that of the original solution space $(E_i \in \Theta)$ is the ith particle code chain in a particle swarm, and it has the corresponding flying velocity V_i, which has the same dimension as E_i. The element $e_i(n) \in [0, 1]$ $(1 \le n \le N)$ denotes whether the source at the nth position in the original solution space Θ is activated or not in the ith particle, and the element $v_i(n) \in [0, 1]$ $(1 \le n \le N)$ represents the velocity at the nth position of the ith particle. The positions and velocities will be updated during iteration procedure.

For a given source activation threshold ε within range [0, 1], if $\| e_i(n) \| \ge \varepsilon$, the nth position is assumed to have active source. By scanning each element in the code chain E_i from the beginning to the end, all the positions with $\| e_i(n) \| \ge \varepsilon$ are taken out to construct the

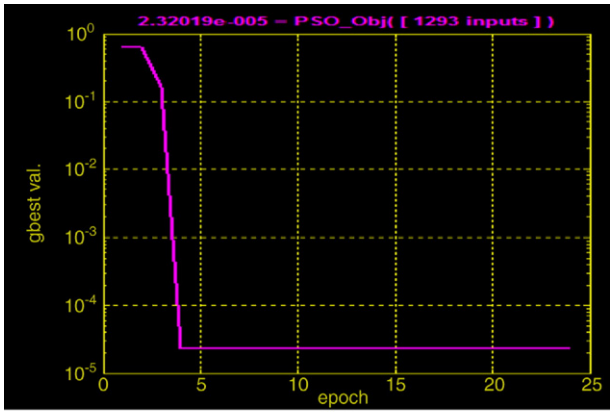


Fig. 2 (*continued*).

compressed space $\Theta_{s,i}$, and the corresponding columns in lead field matrix $A \in \Theta$ are used to form the sub-lead field matrix $A_{si} \in \Theta_{s,i}$. For convenience to describe the algorithm, the procedure to determine the possible source positions and form the subspace is represented by an operator $\mathrm{Com}(\cdot,\cdot,\cdot)$ as,

$$A_{s,i} = \mathrm{Com}(A, \quad E_i, \quad \varepsilon) \tag{6}$$

Based on the compressed space $\Theta_{s,i}$ and the corresponding lead field matrix $A_{s,i}$, a solution $X_{s,i}$ can be readily obtained by Equation (4). Apparently, the dimension K of the subspace $\Theta_{s,i}$ is not a constant, but a dynamic variable along with the evolution of the particle state in the iterations.

To get a good compressed subspace with the desired information remained and the redundant information removed as much as possible, a particle swarm having Q particles is adopted to encode the solution space. Let $\{E_1,E_2,\ldots\ldots,E_Q\}$ denote the particle swarm with each $E_i(1 \le i \le Q)$ representing an individual coding of the solution space, and the corresponding particle velocities are $\{V_1,V_2,\ldots\ldots,V_Q\}$. For each coding chain E_i and its corresponding solution $X_{s,i}$, the l_0 norm constrained fitness function is given by,

$$f\left(X_{s,i}\right) = \|Y - A_{s,i}X_{s,i}\|_2 + \lambda\|X_{s,i}\|_0 \tag{7}$$

This function is adopted to evaluate the quality of the compressed subspace $\Theta_{s,i}$ represented by current particle. The updating strategy of PSO integrates both the local optima and the global optima reached in previous iteration into current iteration, and can drive the particles to move toward the global optima, which provides PSO a powerful capability to find the global optima. A more detailed description of PSO is given in Appendix.

Adaptive threshold function to reduce the noise effect in Θ_s

In practice, the observed signal Y is usually corrupted by noise, and various regularization techniques such as the truncated singular value decomposition, L-curve, CRESO and Tikhonov regularization have been adopted to suppress the noise effect in current inverse techniques (Michel et al., 2004; Pascual-Marqui, 1999; Pascual-Marqui et al., 1994; Silva et al., 2004; Tikhonov and Arsenin, 1977). In this work, we assume that there is no reliable prior information about the scalp data and the solution (Hauk, 2004). In the compressed space, the EEG inverse problem is transformed to a more robust system, which may be solved by the popular generalized inverse operator as shown in Equation (4). As the matrix A_s is almost in a normal state with a high rank in column, the inverse calculation of the matrix is relatively stable and the effect of noise is suppressed with only a few weak fake sources remained. In this work, to further tackle those relatively weak fake sources induced by the noise in the estimated source X_s in the compressed space (Tikhonov and Arsenin, 1977; Yao and Dewald, 2005), we introduced an adaptive threshold similar to the hard threshold strategy adopted in wavelet denoising (Donoho and Johnstone, 1994). In our method, we assume that the threshold is adaptively varied with the maximum source power in current iteration. Let J be the maximum absolute strength in X_s, the adaptive noise effect reduction function is defined as,

$$R(x(k)) = \begin{cases} x(k), |x(k)| > \gamma J \\ 0, |x(k)| \le \gamma J \end{cases} \tag{8}$$

where $x(k)$ is the kth element in X_s, and γ is the noise reduction factor, which may vary with noise power induced in recordings.

Procedure of 3SCO algorithm

Suppose the dimension of the original solution space Θ is N and the swarm consists of Q particles, set the source activation threshold ε, the maximum iteration number $G_{\max}$, the termination error δ, the toleration iteration steps D, the sparsity penalty term λ and the noise reduction factor γ. The specific values of these parameters are listed in the following simulation studies. The procedure of 3SCO for EEG inverse problem can be intuitively depicted by below Fig. 1.

The detailed implementation of Fig. 1 can be summarized as follows,

Step 1. Particle initialization: initialize the position $P_i(1 \le i \le Q)$ of each particle in swarm $\{E_1, E_2,\ldots\ldots E_Q\}$ with N random values uniformly distributed within [0, 1], respectively; initialize the velocity $V_i(1 \le i \le Q)$ of each particle by N random values uniformly distributed within [0, 1].

Step 2. Compress the solution space represented by each particle using Equation (6). For the ith particle E_i, the compression is $A_{s,i} = Com(A, E_i, \varepsilon)$, $1 \le i \le Q$.

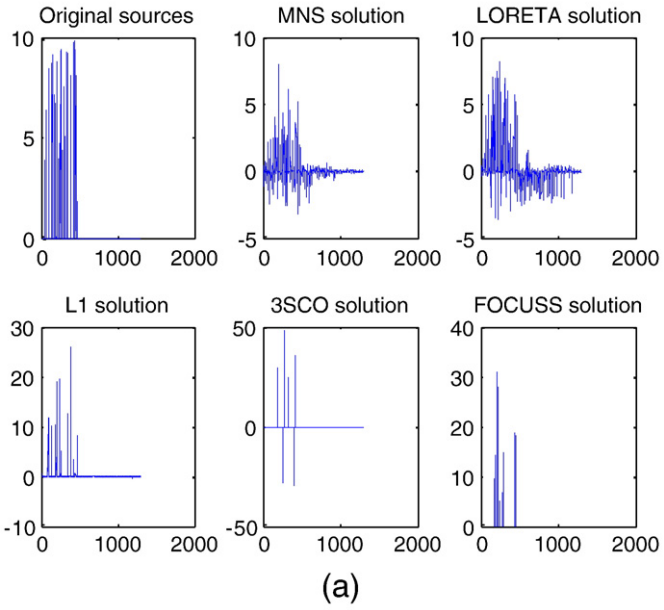


Fig. 3. Localization for extended sources on cortex. (a) Sources in the indexed solution space. (b) Source mappings on the realistic head for the extended source configuration in panel a. In panel b, the left panel represents the left hemisphere viewed from $+X$ and $-X$; the right panel represents the right hemisphere viewed from $-X$ and $+X$. The color bar in panel b is the same for all methods.

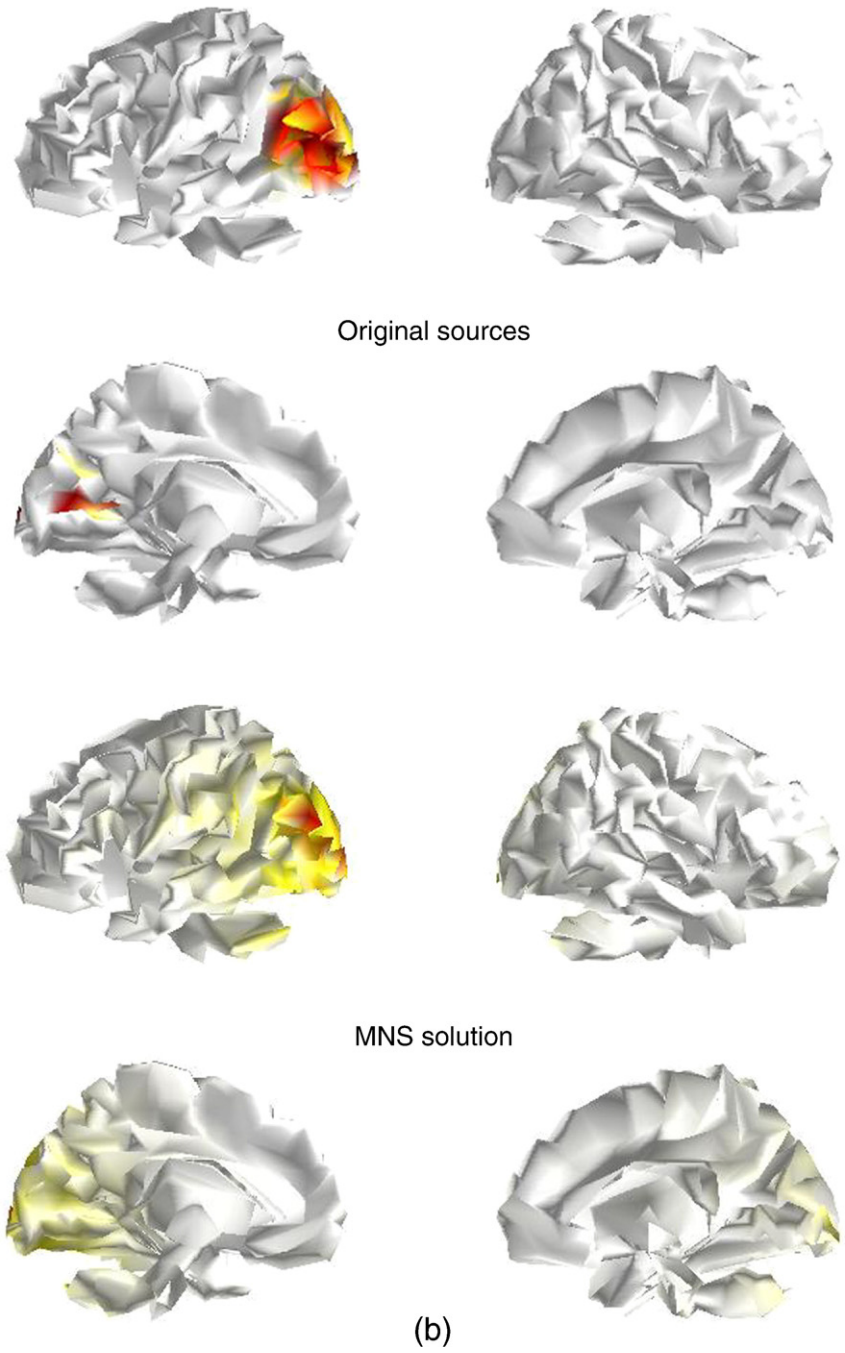


Fig. 3 (*continued*).

Step 3. Estimate the sources in the subspace as $X_{s,i} = (A_{s,i})^{+}Y$ and find the maximum strength of sources in $X_{s,i}$, followed by the elimination of the noise effect by Equation (8). Finally, $A_{s,i}$ and $X_{s,i}$ are used to calculate the fitness value as: $f(X_{s,i}) = \| Y - A_{s,i} X_{s,i}\|_2 + \lambda\|X_{s,i}\|_0$ for the ith particle.

Step 4. Update the best position of each particle: compare the fitness value $f(i)$ of the ith particle at current position (code chain E_i) with the best fitness value ever achieved by this particle at the position P_i, if $_f(i)$ is better, P_i will be replaced by the current code chain E_i, otherwise P_i will be remained, where $1 \le i \le Q$.

Step 5. Update P_g, the best position ever achieved by all the particles: compare the updated best fitness value of the ith particle at position P_i with the global optimal fitness value at P_g, if the value at P_i is better, P_g will be replaced by P_i, otherwise P_g will remain unchanged.

Step 6. Update the velocity and position of each particle as,

$$\begin{cases} v_i(n) = wv_i(n) + c_1 r_1 (p_i(n) - e_i(n)) + c_2 r_2 \left(p_g(n) - e_i(n)\right) \\ e_i(n) = e_i(n) + v_i(n) \end{cases} \quad (9)$$

where $1 \le n \le N$, $1 \le i \le Q$. $e_i(n)$ and $v_i(n)$ are the nth coding element and the nth velocity element of the ith particle, respectively; w, c_1 and c_2 are the same as those in the standard PSO. In our method, the particle velocity and the element in a code chain are constrained

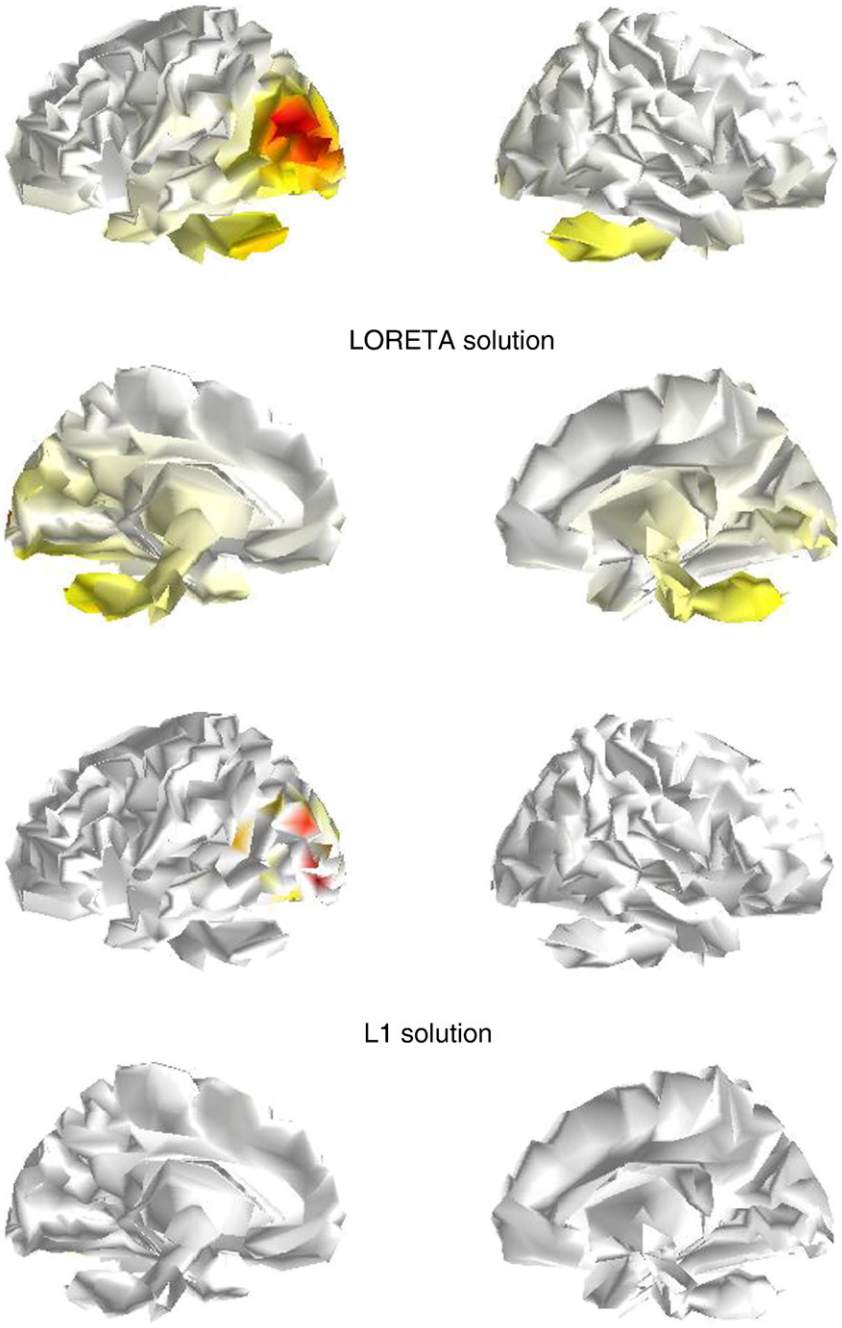


Fig. 3 (*continued*).

within range [−1, 1] as: If $v_i(n)>1$, $v_i(n)=1$; if $v_i(n)<-1$, $v_i(n)=-1$; if $e_i(n)>1$, $e_i(n)=1$; if $e_i(n)<-1$, $e_i(n)=-1$.

Step 7. Evaluate the stopping criteria: If the number of generation is larger than the predefined number $G_{\max}$ or the decreasing of P_g has been less than the termination error δ in the last continuous D iterations, the iterations will be stopped and the solution corresponding to the global optimal position (code chain) P_g will be taken as the final solution for the underdetermined system, else return to Step 2 and go on.

The above iteration procedure shows that 3SCO will iteratively approximate the desired subspace using the PSO updating mechanism, and then the inverse problem will be alternatively solved in the compressed solution space. The proved searching ability of PSO can provide a large opportunity to find an optimal subspace facilitating the final EEG source estimation.

Simulation evaluation

Head model

In this work, simulations were conducted on a realistic head model with current sources on the cortical gray matter, where each current source was constrained to be perpendicular to the cortical sheet (Dale and Sereno, 1993). The tissue conductivities for brain cortex, skull and scalp were 1.0, 0.0125, and $1.0\Omega^{-1}m^{-1}$, respectively. The solution

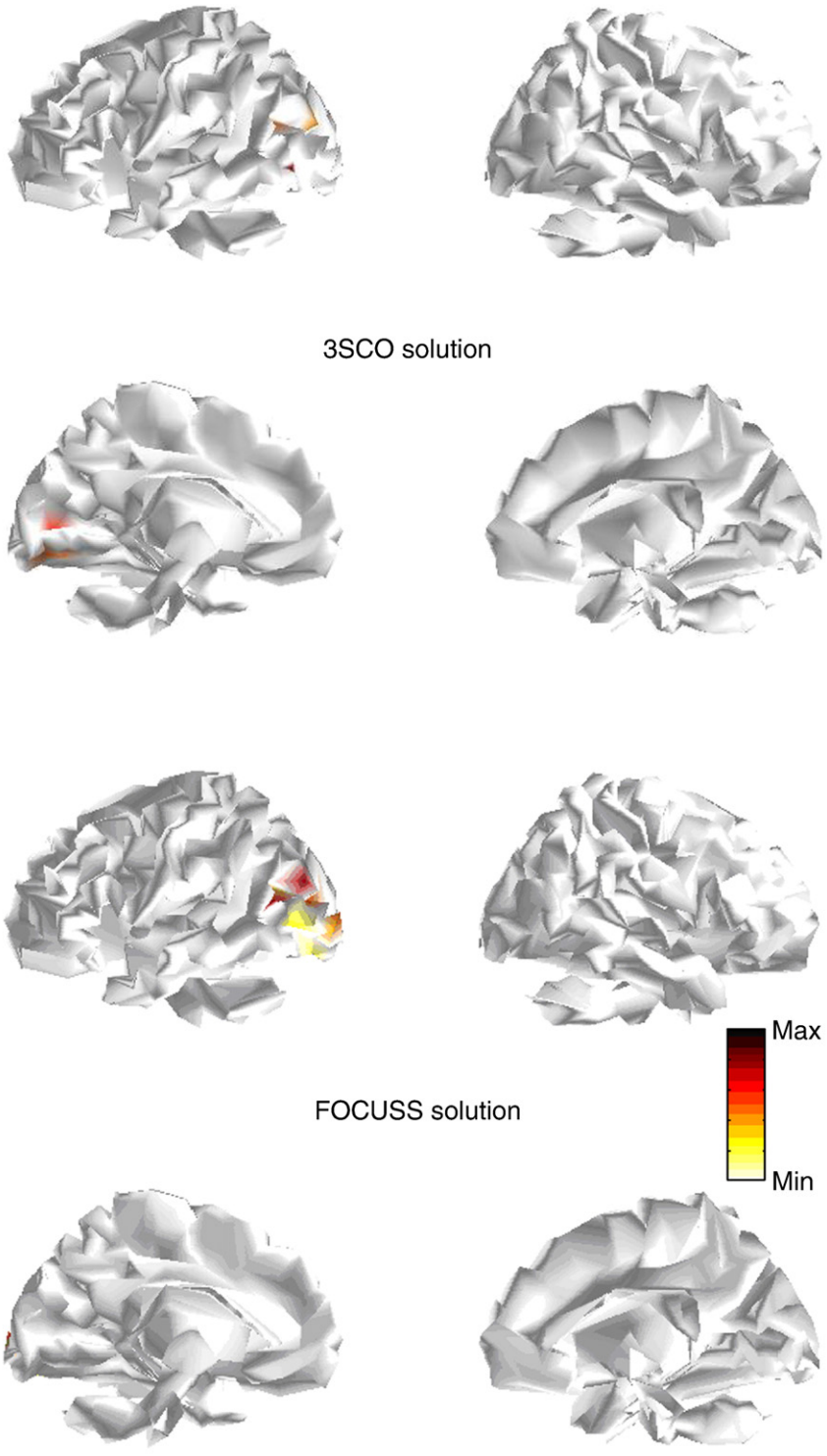


Fig. 3 (*continued*).

space consisted of 1293 nodes. The lead field matrix with dimension of 128×1293 was calculated by the dipole model using boundary element method (BEM) (Fuchs and Drenckhahn, 1998), where the dipole was assumed to be perpendicular to the cortex surface. The origin of the coordinate system was defined as the midpoint between the left and right preauriculars, and the directed line from the origin through the nasion defines the +*X*-axis, the +*Y*-axis was the directed line from the origin through the left preauricular. Finally, the +*Z*-axis was the line from the origin toward the top of the head (through electrode Cz). The 128 electrodes used for measurements were registered on the scalp.

Evaluation protocol

We will first report the possible effect of source configurations on inversion approach in the section *Localization for different source*

configurations on cortex and section *Localization for extended sources*. In the section *Effect of noise on source localization*, we will also report the source estimation under noise conditions. The section *Effect of parameter setups on 3SCO* will study the effect of 3SCO parameters on source estimation. The section *Quantitative performance analysis of localization methods* will provide the quantitative comparison among different approaches.

Localization for different source configurations on cortex

In this section, to investigate the possible effect of different source configurations on the inverse approach, we evaluated and compared the localization abilities of multiple methods for various source configurations having different source number and source position. The tested three different source configurations are of two, three and four sources, respectively. For the two-source configuration, the two dipoles with strengths of 4.6 and 5.7 were placed at positions (−47.0 mm, −18.0 mm, 28.1 mm) and (17.2 mm, −62.0 mm, −1.0 mm); for the three-source case, the three dipoles with strengths of 4.3, 5.6, and 7.3 were placed at positions (−27.7 mm, −1.0 mm, 55.0 mm), (−23.0 mm, 16.2 mm, −18.0 mm) and (−12.0 mm, 2.1 mm, 64.0 mm); and for the four-source case, the four dipoles with strengths of 3.9, 2.6, 6.7 and 3.4 were placed at positions (−31.8 mm, −9.0 mm, −8.0 mm), (−8.8 mm, −40.0 mm, 41.0 mm), (20.0 mm, −52.3 mm, 16.0 mm) and (56.1 mm, −64.0 mm, 11.0 mm). The parameters of 3SCO were initialized as follows: the swarm size (particle number) Q was 300; the sparsity penalty factor λ was 0.01; the number of generation $G_{\max}$ was 100; the source activation threshold ε was 0.8; the inertia weight was updated as $w(u) = 0.9 - \frac{0.4}{G_{max}} \times u$, where u was the current generation number (Shi and Eberhart, 1998); the velocity constants, c_1 and c_2, were both 2.0; the noise reduction factor γ was 0.1, the tolerance error δ was set to 10^{-6} and D was 20.

The localization results were shown in Fig. 2. When 3SCO converged, the iteration numbers for two, three and four sources were 53, 23 and 67, respectively. 3SCO showed better localization ability than the other tested methods for all the tested configurations, where 3SCO was able to exactly recovered all the simulated source configurations.

Localization for extended sources

The previous simulations tested a few isolated sources configurations. In this section, we would test the localization capability of 3SCO for extended sources. The neighboring cortex patches enclosed in a 30 mm-radius-sphere centered at (−42.0 mm, −83.4 mm, 16.0 mm) were assumed to be the activated areas. Those neighboring patches consisted of an extended source area approximately in the occipital area. The amplitudes of those sources were randomly varied within [0, 10]. The parameters of 3SCO were the same as those in the Section *Localization for different source configurations on cortex*. Fig. 3 showed the sources estimated when different methods were applied to the extended sources. The results in Fig. 3 revealed that a blurred distribution was obtained by MN and LORETA, and a sparser distribution was estimated by other three methods, i.e., l_1 norm, FOCUSS and 3SCO. The 3SCO procedure converged to a solution at iteration number 79.

Effect of noise on source localization

In this section, we will briefly report the source localization with 20 dB noise contaminated for a three-source configuration, where the positions for the three sources are (−30.0 mm, −98.9 mm, −2.0 mm), (25,0 mm, −89.0 mm, 1.9 mm) and (41.0 mm, 39.0 mm, 27.0 mm), respectively. Here, the SVD truncations for MN, LORETA and FOCUSS were 7%, 2% and 5% of the corresponding maximum singular values, respectively. The regularization parameter for l_1 norm was 0.05. The parameters for 3SCO were as follows: the swarm size (population size) Q was 300; the regularization parameter λ was 0.1; the number of generation $G_{\max}$ was 100; the source activation threshold ε was 0.8; the inertia weight was updated as $w(u) = 0.9 - \frac{0.4}{G_{max}} \times u$, where u was the current generation number (Kennedy, 1998; Kennedy and Eberhart, 1995; Shi and Eberhart, 1998); the velocity constants, c_1 and c_2, were both 2.0; the noise reduction factor γ was 0.3; the tolerance error δ is 10^{-6} and D was 20. After 93 iterations, 3SCO converged to a solution. As seen in Fig. 4, the performance of each method was decreased to some degree compared to the noiseless simulation. However, 3SCO again showed superior performance in dealing with noise.

Effect of parameter setups on 3SCO

3SCO involves those parameters like swarm size Q, regularization parameter λ, activation threshold ε and the noise reduction factor γ, and they will potentially influence the performance of 3SCO. We adopted the three-source configuration used in the section *Localization for different source configurations on cortex* to study the possible effects of those parameters on 3SCO. In the following four simulations, except for the specific parameter to be studied, other parameters were kept same as their corresponding setups in *Localization for different source configurations on cortex*.

Effect of particle number Q

The particle numbers were varied as: 60, 90, 120, 150, 210, 270, 300 and 360. For each particle number, we ran 3SCO for source estimation and checked if the source configuration is correctly recovered. The results are listed in Table 1.

Effect of activation threshold ε

This study was implemented by varying ε as: 0.50, 0.55, 0.60, 0.65, 0.70, 0.75, 0.80, and 0.85, and the similar protocol was applied (Table 2).

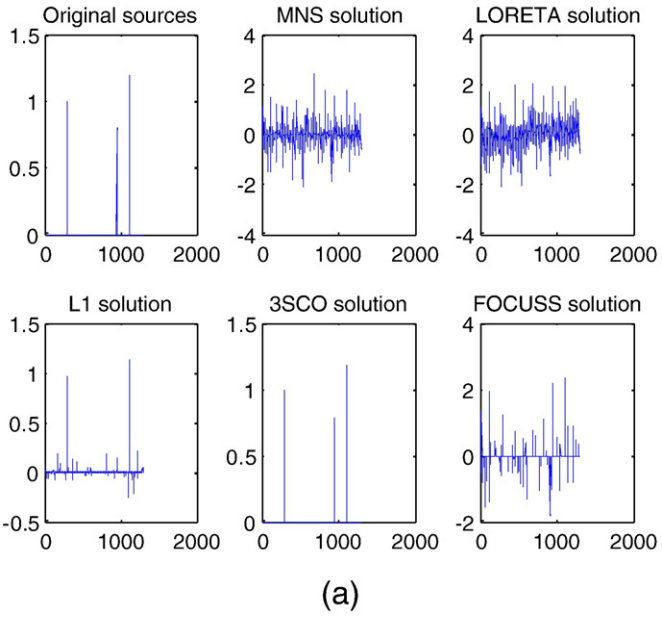


Fig. 4. Localization under 20 dB noise condition. (a) Sources in the indexed solution space. (b) Source mappings on the realistic head for the three-source configuration in panel a. In panel b, the left panel represents the left hemisphere viewed from $+X$ and $-X$; the right panel represents the right hemisphere viewed from $-X$ and $+X$. The color bar in panel b is the same for all methods.

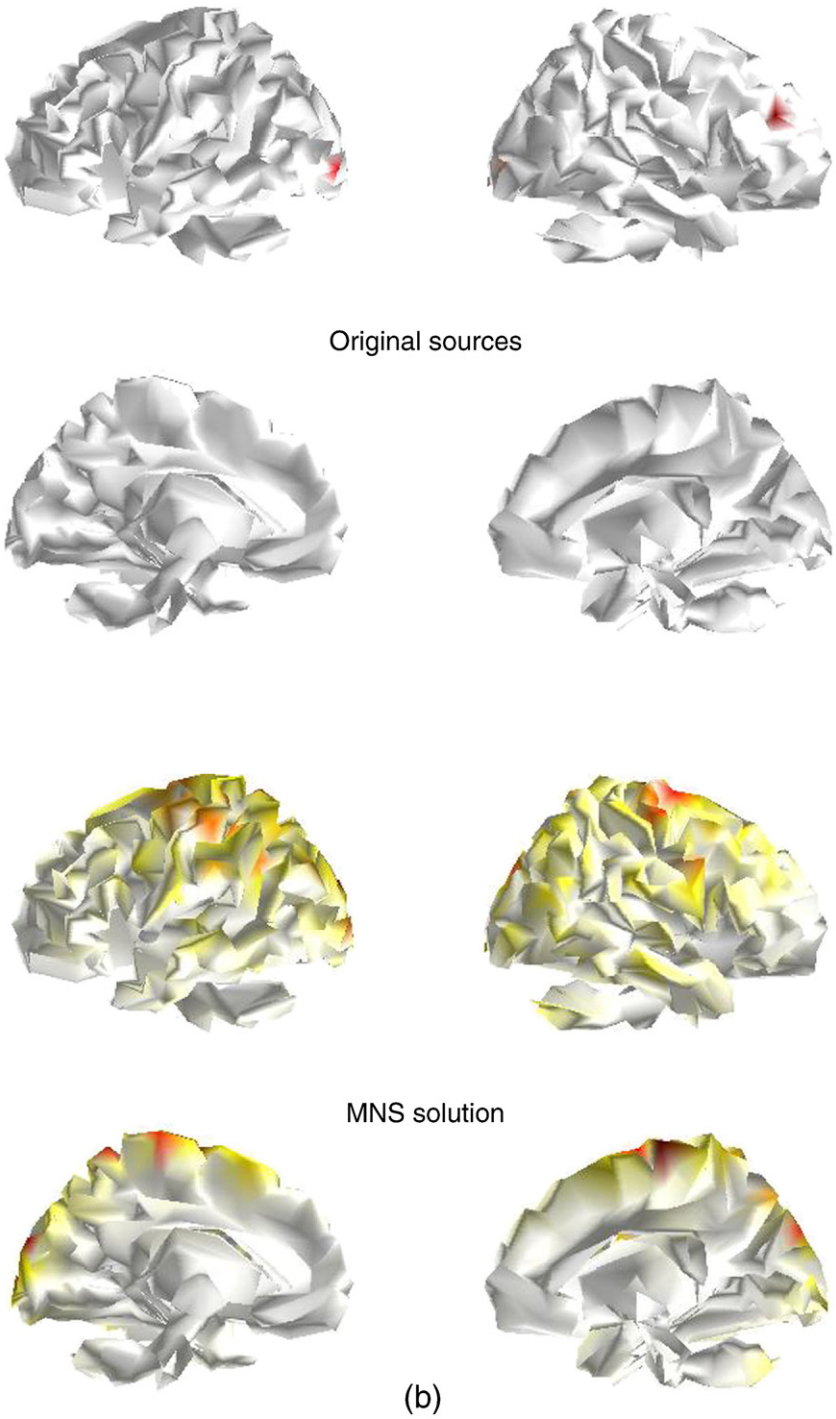


Fig. 4 (*continued*).

Effect of noise reduction factor γ

This simulation was realized by varying γ as: 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, and 0.8. Using the similar evaluation protocol, the result is listed in Table 3.

Effect of regularization parameter λ

In this simulation, Table 4 below was achieved by varying λ as: 0.01, 0.05, 0.10, 0.20, 0.60, 1.00, 1.40, 1.60, 1.80, and 2.00.

This regularization parameter λ plays a very important role for the penalty based inversion approaches, and Bayesian Inference Criteria (BIC) can be used to determine the optimal regularization parameter. In this section, we also calculated the corresponding BIC for those different λ. A scatter plot of λ and BIC was shown in Fig. 5, where the smallest λ is 0.01 and the largest λ is 2.00.

Quantitative performance analysis of localization methods

Generally, the head model, source position, dipole orientation and even the source distribution etc. can affect the performance of a localization algorithm (Michel et al., 2004; Pascual-Marqui, 1999; Yao and Dewald, 2005). The position error and strength error are two useful metrics used to evaluate the performance of an EEG localization

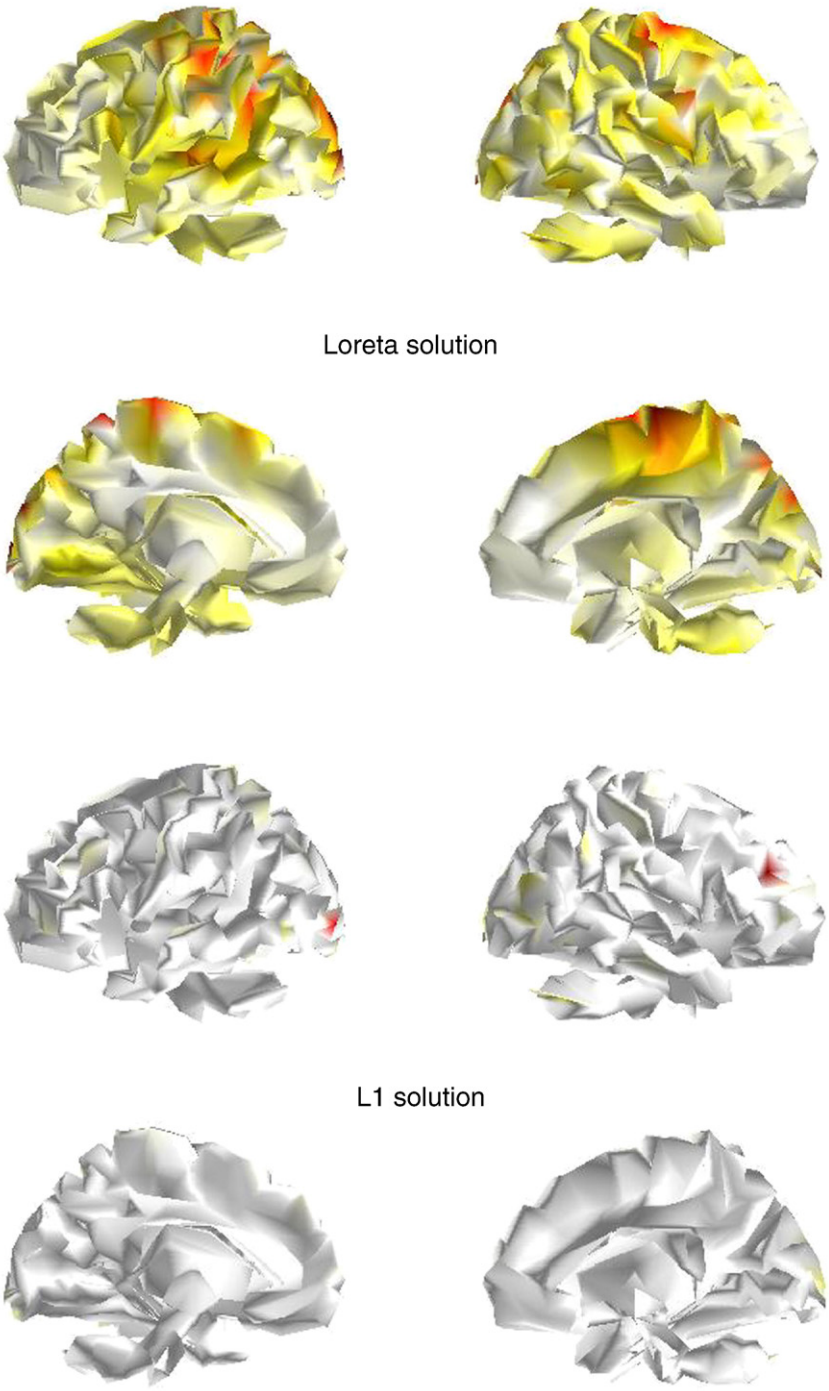


Fig. 4 (*continued*).

algorithm (Liu and Gao, 2004; Michel et al., 2004; Pascual-Marqui, 1999; Yao and Dewald, 2005). The localization position error is defined as $E_{\text{localization}} = ||p_{\text{est}} - p_{\text{simu}}||$, where p_{est} and p_{simu} are the position vectors of the estimated source and the desired (simulated) source; the strength error is defined as $E_{\text{energy}} = \frac{||J_{\text{simu}} - J_{\text{est}}||}{||J_{\text{simu}}||} \times 100\%$, where J_{est} and J_{simu} are the moments (strengths) of the estimated source and the assumed (simulation) source, respectively. As for the one-source case, the source having the maximum power can be directly regarded as the estimated source (Liu and Gao, 2004; Yao and Dewald, 2005). When more complex source configurations consisting of multiple sources areconsidered, how to assign the estimated sources to the corresponding simulated sources is another difficult problem to be addressed. As for the multi-source configuration, the earth mover distance (EMD) is a useful measurement to evaluate the source difference without having to match the estimated sources with the simulated sources. EMD metric provides a meaningful measure for arbitrary types of source distributions, where a small EMD denotes a good match between two distributions (Haufe et al., 2008). In this paper, EMD was calculated with the code provided by Rubner et al. (2000). We calculated the corresponding EMD for both the one-source configuration and the two-source configuration. In this section, we implemented two simulations to investigate the algorithm performance from different views.

(a) The aim of this simulation was to generate the error distribution across the whole solution space. In this simulation, after

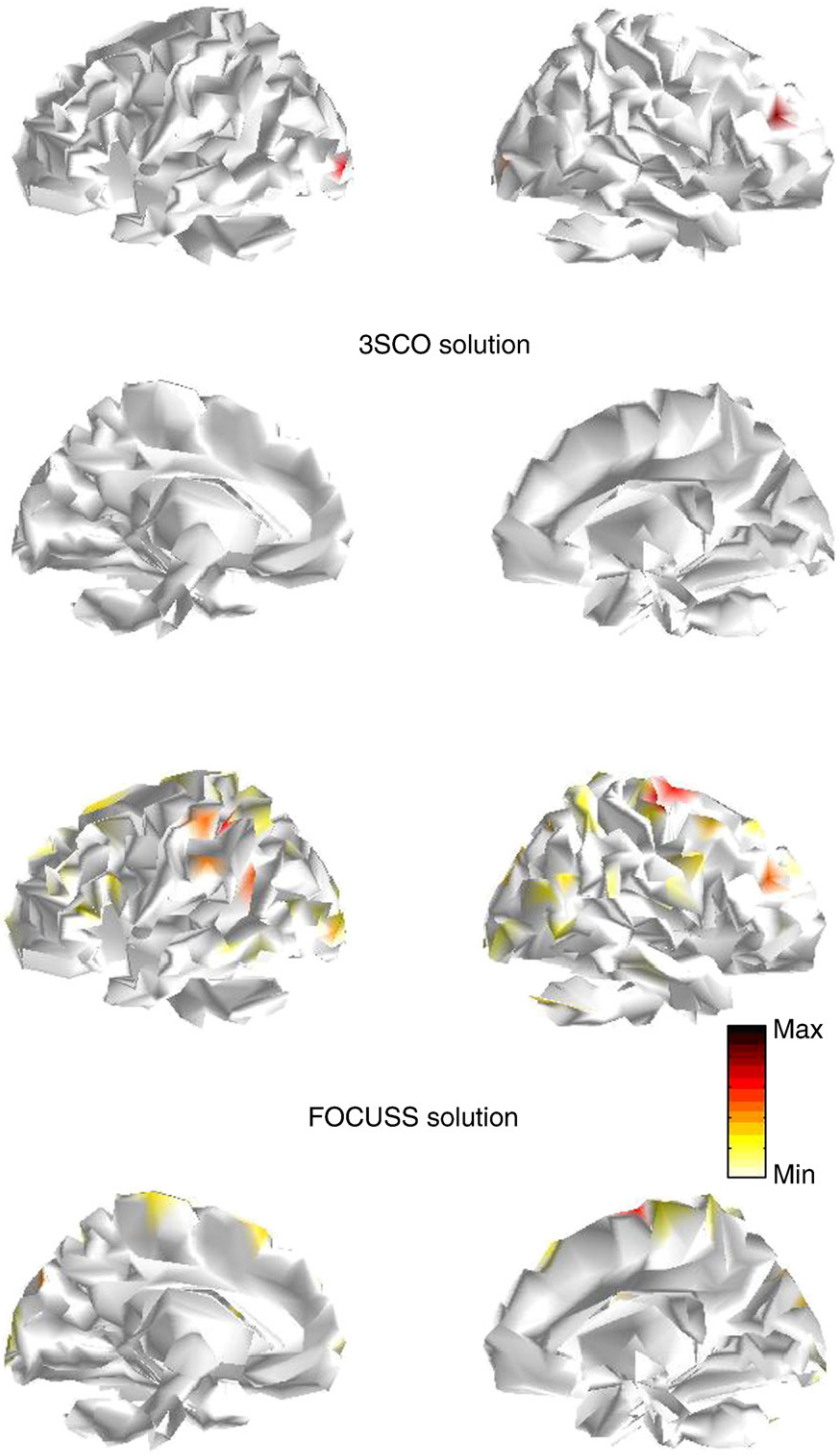


Fig. 4 (*continued*).

a unit source was placed on each solution node to generate certain scalp potential, this scalp potential was corrupted with 20 dB SNR noise. The five inversion approaches were then used to estimate the sources for the noise-free and noisy scalp potentials, respectively. After the calculation of $E_{\text{localization}}$, E_{energy} and EMD for each node, the mean and standard deviation (SD) of $E_{\text{localization}}$, E_{energy} and EMD were calculated over all nodes for the noise-free and noisy cases, respectively. Those parameters setup of 3SCO in the section *Localization for different source configurations on cortex* and the section *Effect of noise on source localization* were used for the noise-free and 20 dB SNR cases, respectively. The error distribution on the head model for 20 dB noise case was shown in Fig. 6. The mean iteration numbers of 3SCO before convergence

Table 1
The effect of particle number on source localization.

Q	60	90	120	150	210	270	300	360
Source recovered?	No	No	Yes	No	Yes	Yes	Yes	Yes

Table 2
The effect of activation threshold on source localization.

ε	0.50	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95
Source recovered?	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	No

Table 4
The effect of λ on source localization.

λ	0.01	0.05	0.10	0.2	0.6	1.0	1.4	1.6	1.80	2.00
Source recovered?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No

were 43 and 49 for the noiseless and 20 dB noise conditions, respectively.

(b) This Monte Carlo simulation was conducted to evaluate the performance of different inversion approaches when the spatial positions and strengths of sources were randomly varied. The simulation was implemented by generating the scalp potential using two sources with positions randomly distributed in the solution space and the strengths uniformly distributed within [0, 10]. After generating the scalp potential, the five approaches were used to solve the inverse problem under the noise-free and 20 dB noise conditions, respectively. A total of 200 runs were conducted. For each run, the EMD between the estimated sources and the simulated sources was calculated. The mean EMD error across the 200 runs was listed in Table 6. The mean iteration number of 3SCO across the 200 runs before convergence was 61 for the noiseless case and 76 for the 20 dB noise case.

Real data test

Experiment protocol

A fixation cross ($0.5^\circ \times 0.5^\circ$) at the center of the monitor and two peripheral boxes positioned as illustrated in Fig. 7 were displayed throughout the entire block of trials. Each trial began with the fixation point flashing for 50 ms. After a duration of 750 ms, a stimulus, i.e., a long white bar ($0.75^\circ \times 0.25^\circ$), appeared randomly in the left visual field (LVF) or right visual field (RVF), with its center 5° off and 2.5° above the fixation cross. The stimulus was designed to appear with equal probability at the left or right visual field. The duration of stimuli was 100 ms. The intertrial interval (ITI) ranged randomly between 1000 and 1200 ms. Subjects were required to fixate on the cross and minimize eye blinks and body motion during all experimental blocks. Subjects were instructed to press the key "1" with their left thumb for stimuli appearing in LVF and key "2" with their right thumb for the stimuli in RVF. Response accuracy and speed were emphasized equally. The experiment consisted of a total of 200 trials per participant, separated into two blocks of 100 trials each. Short breaks were allowed between blocks. The experiment flow diagram was shown in Fig. 7. Twelve subjects aging from 20 to 24 years participated in the experiment and the MRI scanning was performed for those subjects. The used cortex head model was wrapped from the averaged subjects' MRIs, which consisted of 1232 nodes.

Source analysis with 3SCO

The grand averaging waveform for the LVF was shown in Fig. 8a. We focused on ERPs with latency of 152 ms after stimulus onset, at which the ERPs including contralateral P1, ipsilateral P1, and early N1 could be observed. The source localization results of 3SCO for this moment showed that the contralateral P1 was localized in the right occipital cortex. The parameter setup of 3SCO was the same as that employed in the section *Effect of noise on source localization*. After 87 iterations, 3SCO stopped. The ipsilateral P1 was localized in the left occipital cortex, and the early N1 was distributed in the right occipito-parietal, occipito-temporal cortex/regions. These activations were related to the primary visual sensory processing (Di Russo et al., 2003), running with a bottom-up model (Corbetta et al., 2002; Corbetta and Shulman, 2002), and were consistent with the previous study related to the visual early ERPs (Fu et al., 2001). Moreover, because the left stimulus was used, the strongest activations were observed in the right occipital, which was denoted by the much darker red color.

Table 3
The effect of γ on source localization.

γ	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8
Source recovered?	Yes	Yes	Yes	Yes	No	No	No	No

Discussion and conclusion

In Fig. 2, whatever for the two-, three- or four-source configuration, the patterns estimated by MN and LORETA are scattered and blurred, which are rather far from the actual one. Compared to MN and LORETA, the solutions of FOCUSS and l_1 norm were more focal and had relatively higher spatial resolution for the sparse sources. In the section *Localization for different source configurations on cortex*, when the source configuration became more and more complex with source number increased from 2 to 4, the simulation demonstrated that the source configuration actually influenced the estimations. As for the relatively simple two-source configuration, FOCUSS, 3SCO and even l_1 norm were able to reconstruct the source patterns well. However, for the more complex three- and four-source configurations, some sources in the tested configurations could not be correctly recovered with l_1 norm and FOCUSS. Moreover, as shown in Fig. 2, some fake sources are introduced by both l_1 norm and FOCUSS for the tested three-source and four-source configurations. Whereas, 3SCO was able to recover the sources at their actual positions.

As shown in Fig. 3, for extended sources, l_1 norm, FOCUSS and 3SCO localized a focal and smaller activation area than the actual activation area. The activated area localized with MN and LORETA were much closer to the actual extended distribution. Certainly, the ideal result was that the simulated distributed sources were recovered correctly, and such a situation is rather difficult to achieve for EEG inverse problem. In practice, many inverse methods just localized an equivalent source or an equivalent distribution with some difference from the true configuration. Generally, if most of the estimated sources are located within or near the simulated area, the localization could be thought to be successful (Grova et al., 2006; Michel et al., 2004). By verifying the estimated source locations, we found that all the five algorithms were able to localize sources within or near the true source area. Due to the emphasis on sparsity of sources in l_1 norm, FOCUSS and 3SCO, they resulted in source configurations having fewer sources than was the actual case. In

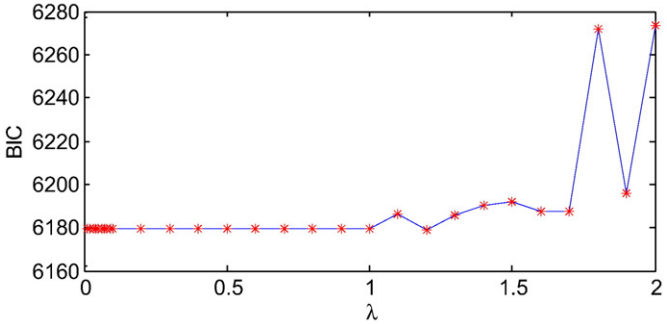


Fig. 5. Scatter plot of λ and BIC.

other words, the results showed that all the five methods obtained their reasonable equivalent source estimations, and the main difference was that l_1 norm, FOCUSS and 3SCO localized a more focal distribution than that estimated by MN or LORETA.

Fig. 4 reveals that the noise had great effect on the source estimation, where the relatively strong artifact sources could be observed for MNS, LORETA, l_1 norm and FOCUSS. However, 3SCO still showed good performance. We attribute this to two aspects of 3SCO: one is the solving of the EEG inverse problem in the compressed subspace, which may reduce the effect of the noise disturbance, and the other is the hard threshold strategy used to further remove the small artifacts induced by noise in the subspace.

Those parameters like swarm size Q, regularization parameter λ, activation threshold ε and the noise reduction factor γ need to be set before using 3SCO, and their effects on 3SCO were shown in the section *Effect of parameter setups on 3SCO*. In the current work, Q is set to be 300. It has been proved that the ability of 3SCO to search for the global optima is largely dependent on the swarm size Q (Eberhart and Shi, 1998; Shi and Eberhart, 1998; Trelea, 2003; Wachowiak et al., 2004). From Table 1, we can see that when particle number is less than 150, the estimation is not reliable, but when particle number becamelarger than 210, 3SCO was able to correctly localize the sources. In our inverse problem, the solution space consisted of 1293 nodes and the table shows that utilization of more particles would facilitate the finding of a better solution just as what had been

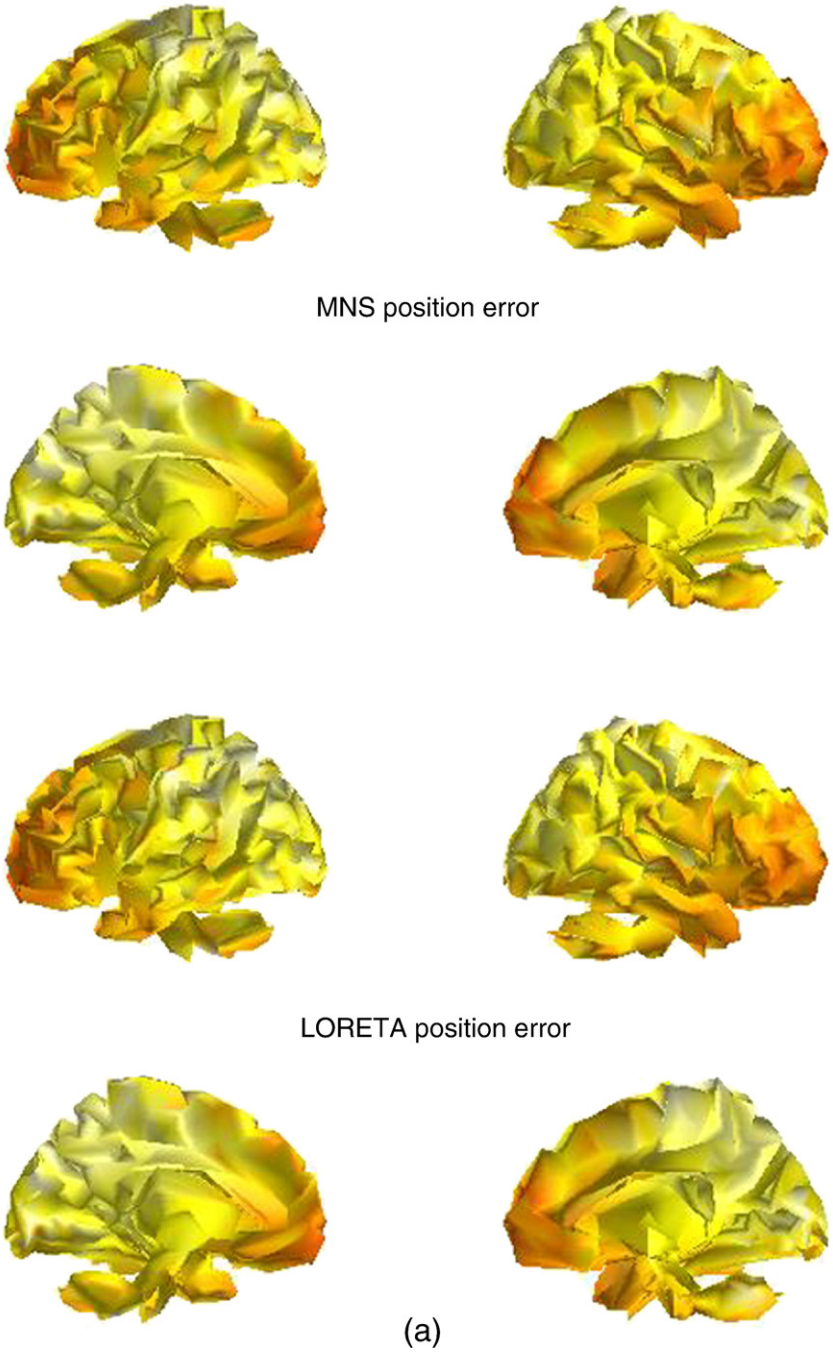


Fig. 6. Error distribution in the solution space for the different methods when 20 dB noise is considered. (a) Position errors; (b) energy errors. The color bar is the same for all methods.

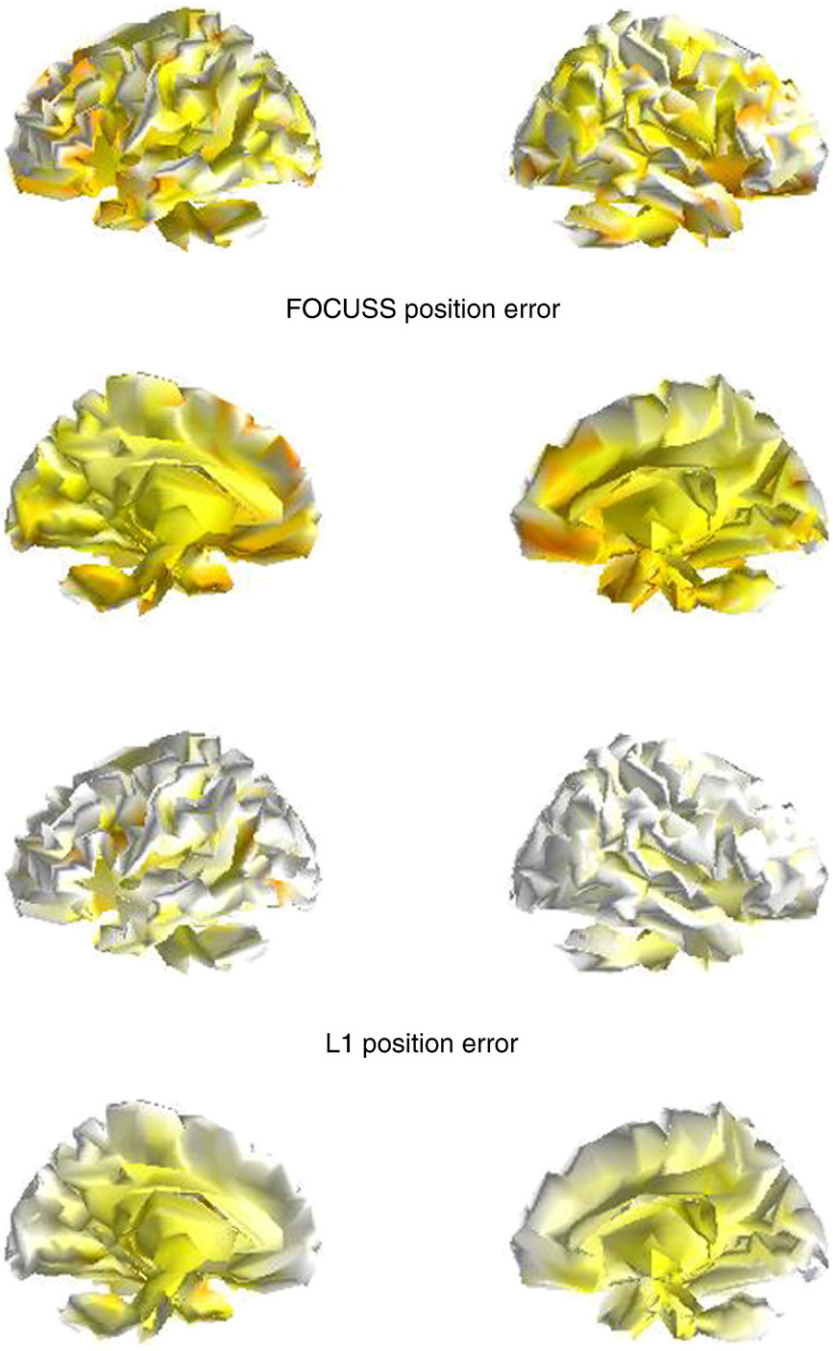


Fig. 6 (*continued*).

reported in other PSO related papers (Qiu et al., 2005; Trelea, 2003; Wachowiak et al., 2004). However, the increased number of particles will accordingly cause 3SCO to require more time to solve an inverse problem. Table 2 gives effect of ε on source localization. We can see that the selection of ε actually influences the estimation. A too smaller or too larger ε will distort the source estimation because a smaller ε will not facilitate the removing of redundant spaces while a larger ε will lead to the loss of the useful space information. In the current work, we set ε to be 0.80. The effect of γ was listed in Table 3, and the table shows that γ also influenced the source estimation. If γ is too large, the weak sources may be truncated; otherwise, the noise artifact cannot be effectively compressed when noise is induced into the recordings. Like other EEG inversion approaches, the regularization parameter λ determines the tradeoff between residual error and the l_0 norm penalty. Generally, the determination of λ is still an open problem though such techniques as L-curve, GCV and BIC, etc. have been proposed. As shown in Table 4, 3SCO did not show too much sensitivity to the regularization parameter λ, where when λ was smaller than 1.60, 3SCO produced the source estimation consistent with the original distribution. However, the BIC plot given in Fig. 5 shows that when λ is within 0.01 and 1.00, BIC keeps a small value and when λ becomes larger than 1.0, the corresponding BIC is obviously increased. From the above results, we can see that 3SCO is stable across a relatively wide range of those parameters. Although much more simulations are needed to further consolidate the result, the conducted study on this three-source configuration does provide

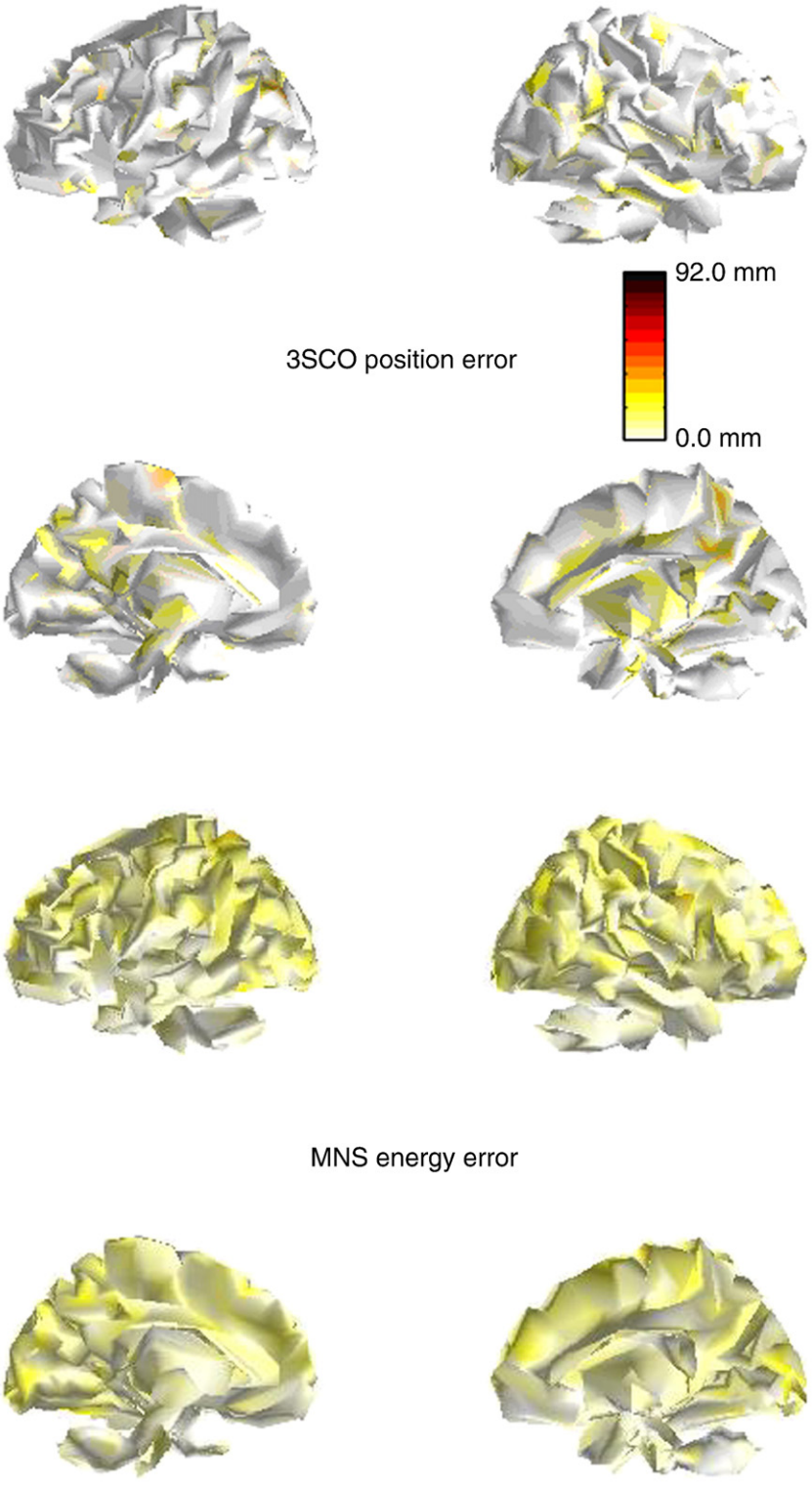


Fig. 6 (*continued*).

confidence in the robustness of the parameter λ in general case. However, the reported effects were derived from one-source configuration, and we acknowledge that the accurate and complete determination of them still needs to perform such analysis like calculating receiver operator curves (ROC) for each parameter.

As shown in Tables 5 and 6, 3SCO showed a much better localization performance than the other four methods. In particular, there existed nearly no bias for 3SCO in the whole solution space when it was used in the one-source configuration under noise-free condition. When noise was considered, the performances of all methods were somewhat lowered compared with the noise-free condition, but 3SCO still had the smallest $E_{\text{localization}}$, E_{energy} and EMD among the five approaches. Moreover, Fig. 6 reveals that the errors of MNS and LORETA were almost distributed across the entire solution

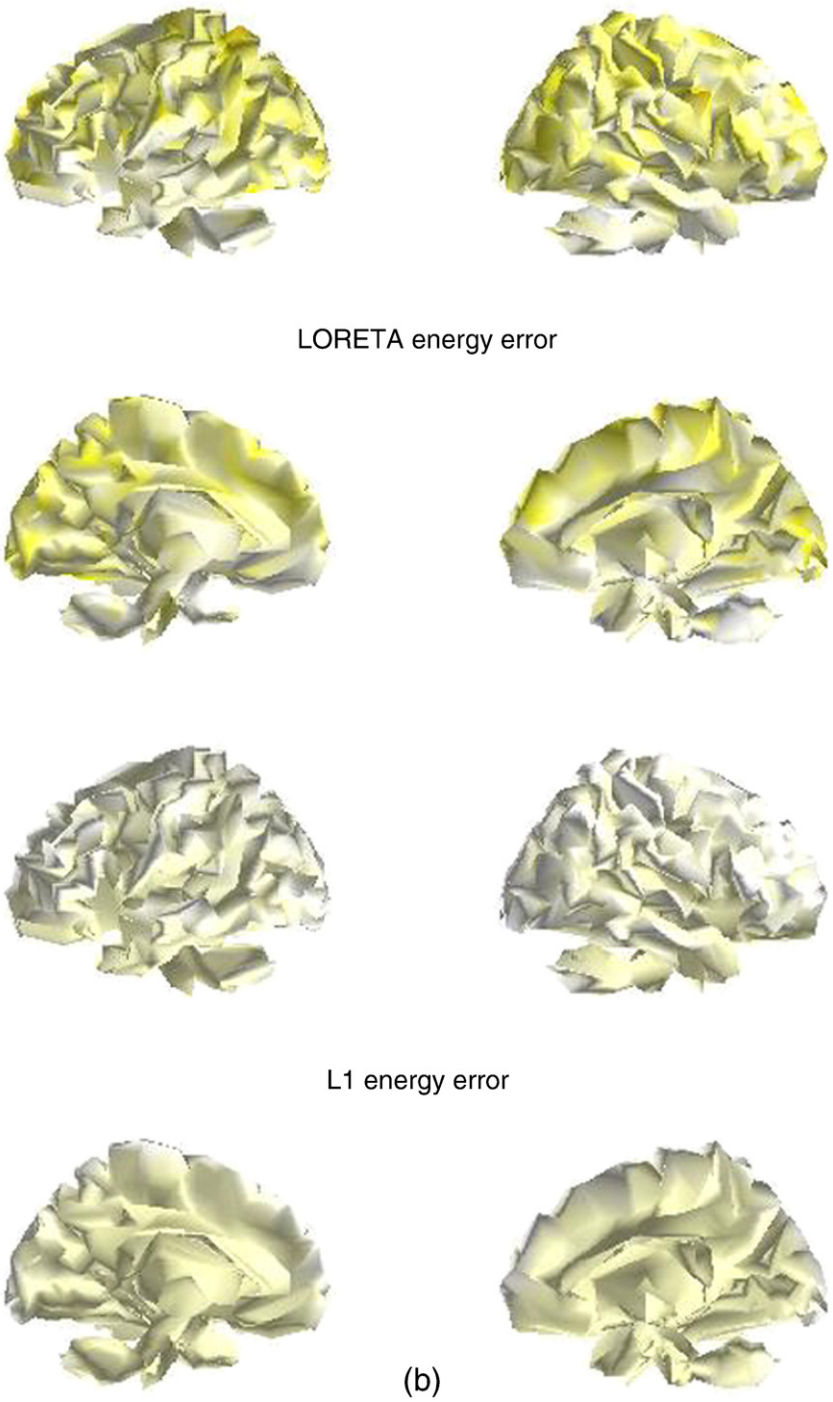


Fig. 6 (*continued*).

space, and l_1 norm and FOCUSS had larger errors in the relatively deep nodes. The findings were consistent with the above simulations where LORETA and MN had larger localization errors than l_1 norm, FOCUSS and 3SCO when used in the sparse sources. Because the simulated source was the isolated source distributed on each node, it was reasonable to conclude that MN and LORETA were not as competitive as l_1 norm, FOCUSS and 3SCO. Table 6 shows that when the more complex source configuration having more than one source was considered, the performances of all the inversion approaches were lowered compared to the corresponding unique source simulation. As with other simulations, 3SCO again gave the best performance.

For the real left visual stimuli data, using 3SCO, activations in the occipitals and the right occipital–parietal area were localized, which were consistent with the findings in previous studies (Corbetta and Shulman, 2002; Di Russo et al., 2003; Fu et al., 2001).

The above results consistently show that 3SCO is robust for isolated source estimation. For all the tested simulations, 3SCO can localize the position of sources with high accuracy. Compared to other localization methods, the improvement of 3SCO is mainly attributed to the solving of EEG inverse problem in the subspace and the using of coding strategy to compress the solution space by the swarm particles. The powerful searching capability of particle swarm ensures to find an optimal compressed sub-solution space with high

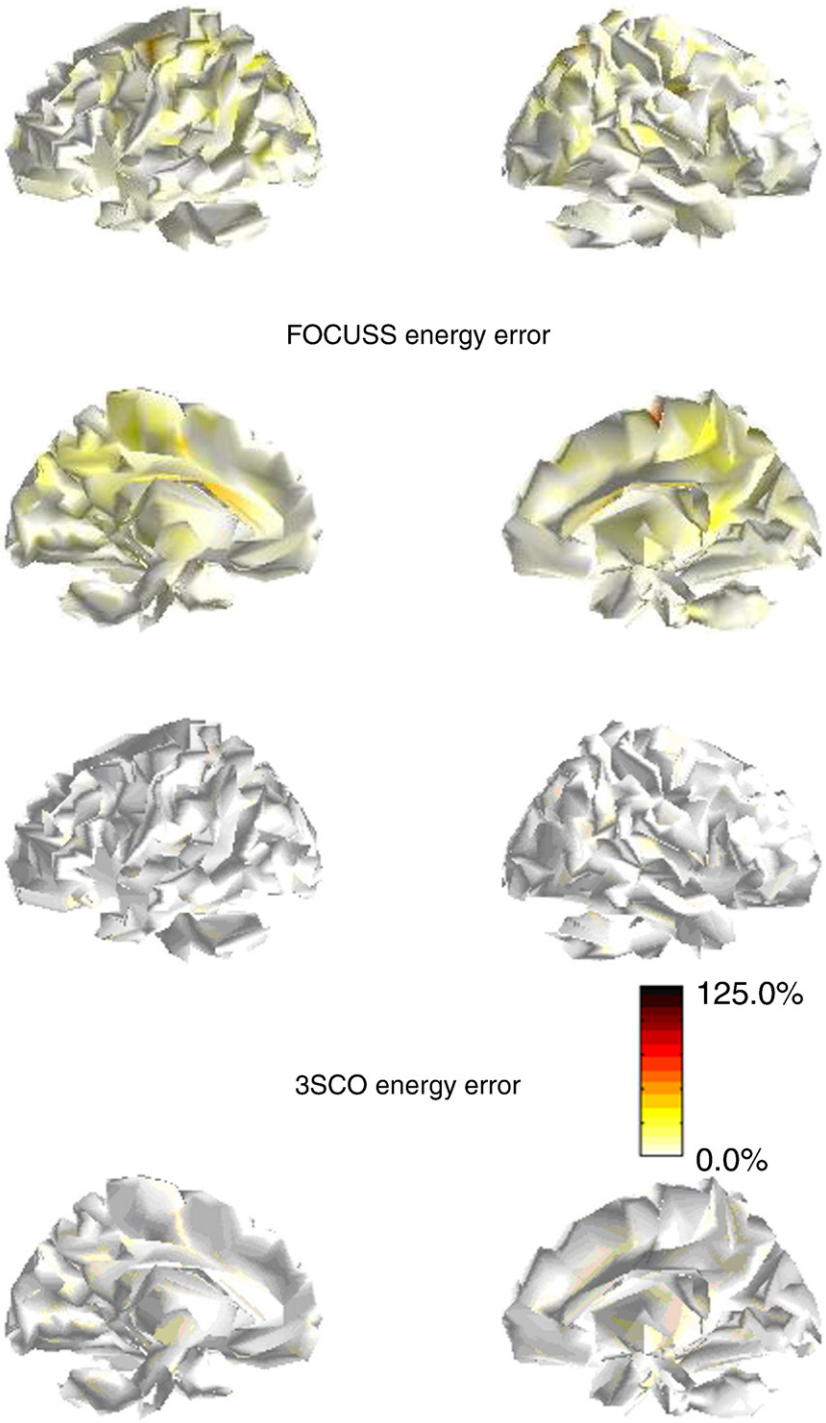


Fig. 6 (*continued*).

probability, and in the optimal compressed solution space, the mathematical performance of the compressed lead field matrix is improved to have a nearly full column rank, so the inverse problem is usually converted to a non-underdetermined one that can be solved in the sub-solution space in a relatively stable manner. In addition, the new adopted fitness function using l_0 norm measure makes the solution sparse and robust compared to the traditional minimum norm strategy. In essence, the usual bias toward the surface in current minimum norm based method is mostly due to the imposed minimum norm strategy as the power of solution partly contributes to the object function. When using the strategy of l_0 norm constraint, the number of sources instead of the source power is contributed to the object function, and thus 3SCO can theoretically modify the bias.

Certainly, 3SCO is mainly used to localize the sparse configuration sources. Whereas for the extended sources, 3SCO may converge to an equivalent sparse solution, and other methods like MNS and LORETA may be more suitable for the estimation of the extensive sources. Vega-Hernandez et al. (2008) have developed a framework to combine multiple norm penalties, and this framework can adapt to the sparse or the extended sources by adjusting the corresponding penalty parameters. In essence, such norm constrained

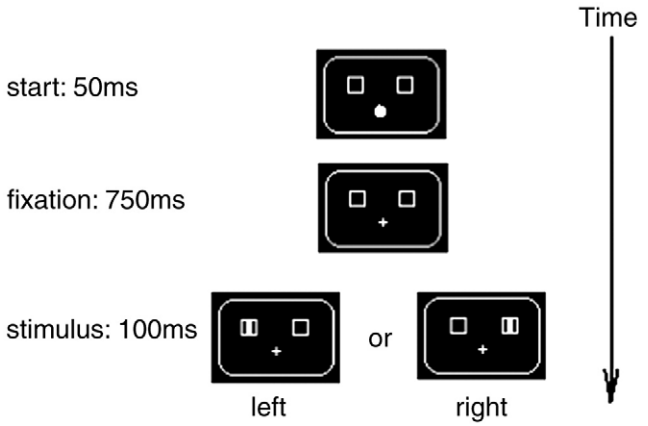


Fig. 7. The experiment flow diagram.

solutions like LASSO (i.e., l_1 norm) (Tibshirani, 1996), LASSO Fusion (LFusion) (Land and Friedman, 1996), LASSO Fused (LFUSED) (Tibshirani et al., 2005), Ridge Fused (RFUSED) and Elastic Net (ENET) (Zou and Hastie, 2005) can be integrated into this framework. Those mixed norm constraints have been proven to be capable of capturing the advantages of corresponding norms (Ou et al., 2009; Valdes-Sosa et al., 2009; Vega-Hernandez et al., 2008). It is a natural extension to combine those different norm constraints into the fitness function of 3SCO to expand it flexible to both isolated and extended source configurations.

In current work, considering the algorithm efficiency, the pseudo-inverse was used to solve EEG inverse problem in the compressed space, and other approaches like FOCUSS and l_1 norm can be used as well. Moreover, it needs to be pointed out that the SVD truncation can still be considered to deal with noise disturbances in the compressed solution space like that used in the MNS and LORETA solution. The integration of EEG and fMRI can generate the source estimations having both high temporal and spatial resolutions (Babiloni et al., 2003). Actually, the fMRI-derived priors can be easily combined into the 3SCO procedure by coding those possible solution spaces revealed by fMRI with some larger values. In this paper, the focus is the EEG inverse problem, but according to the principle of the algorithm, the method can be applied to solve other underdetermined system for a sparse solution, too. In the current version, those parameters such as the source activation threshold, ε the sparsity penalty factor λ and the noise reduction factor γ are mainly determined by a simple grid search and experience. In future, systemic studies such as BIC, GCV and L-curve, ROC curve etc. are needed to precisely determine them. The current version of 3SCO was developed to deal with the instantaneous sources, and the comparison with some spatio-temporal based methods like RAP-MUSIC was not conducted in

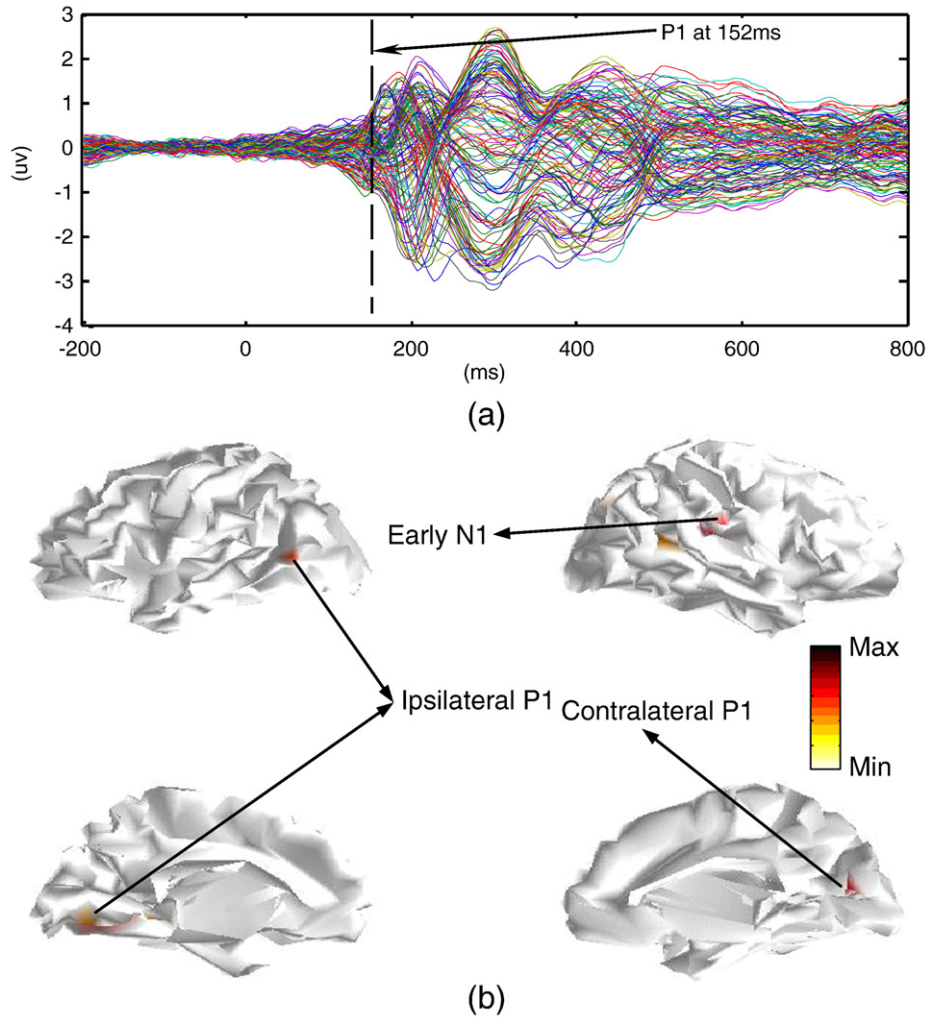


Fig. 8. Sources estimated by 3SCO. (a) The LVR waveform; (b) the activated areas at 152 ms. In panel c, the left panel represents the left hemisphere viewed from $+X$ and $-X$; the right panel represents the right hemisphere viewed from $-X$ and $+X$.

Table 5
Performance of the five localization methods for the single source configuration.

SNR		MNS	LORETA	l_1 norm	3SCO	FOCUSS
Noise-free	$E_{localization}$ (mm)	20.61 ± 16.59	18.96 ± 15.76	0.86 ± 5.01	0.00 ± 0.00	2.26 ± 6.34
	E_{energy} (%)	88.64 ± 13.71	87.86 ± 13.16	22.14 ± 28.95	0.00 ± 0.00	9.02 ± 23.57
	EMD	0.53 ± 4.43	0.41 ± 2.17	0.05 ± 0.43	0.00 ± 0.00	0.01 ± 0.09
20 dB	$E_{localization}$ (mm)	29.78 ± 22.19	26.52 ± 25.75	16.85 ± 24.63	7.78 ± 19.57	19.72 ± 36.72
	E_{energy} (%)	95.80 ± 18.72	91.80 ± 17.90	60.81 ± 28.50	11.17 ± 22.44	66.15 ± 40.76
	EMD	1.62 ± 17.66	1.42 ± 5.97	0.22 ± 0.79	0.09 ± 0.58	0.86 ± 11.77

the current work. Moreover, other approaches including MCMC and VB have been developed for EEG source imaging in view of the posterior probability, and they have attracted wide attention. In essence, 3SCO is still the traditional penalty based inversion approach, therefore, in this paper, we restricted our comparison to other classical penalty based approaches like MN, LORETA, l_1 norm and FOCUSS. Because the searching capability of PSO is largely dependent on the particle number (swarm size), and the solution space for EEG inverse problem is large, usually more than thousands, it needs to use a large number of particles to guarantee a relatively good optima, which alternatively increases the time needed by 3SCO to solve the EEG inverse problem. Like other optimization techniques, it must be pointed out that PSO cannot guarantee to find the global optima for each run, but the used particles and the updating strategy can give PSO large probability to obtain the global optima. In the current work, when 300 particles are involved, approximately 30 minutes is needed for one run on a PC with 2G RAM and a 2.33 GHz CPU. However, with the rapid development of computer hardware, the calculation time can be expected to be dramatically reduced in the near future.

Acknowledgments

This work was supported by NSFC 60701015, 30525030, 60736029 and 30800242, and the 863 project 2009AA02Z301. We thank the anonymous reviewers for their constructive criticisms to improve the manuscript. Thanks are given to Hu JieHui for his assistance to help us polish the writings.

Appendix. Standard particle swarm algorithm

Particle swarm optimization (PSO) was originally designed and developed by Eberhart and Kennedy which was motivated from the simulation of social behavior (Eberhart and Kennedy, 1995; Eberhart and Shi, 1998; Kennedy, 1998; Kennedy and Eberhart, 1995). In recent years, PSO has been successfully adopted to solve many difficult problems (Hu et al., 2004; Parsopoulos and Vrahatis, 2002; Wachowiak et al., 2004). In PSO, instead of using genetic operators, each particle (individual) adjusts its 'flying' according to its own flying experience and its companions' flying experience (Kennedy, 1998; Shi and Eberhart, 1998). Each particle is treated as a point in an N-dimensional space, where N is usually the dimension of the solution space.

The ith particle is represented as $X_i = (x_i(1), x_i(2), \ldots\ldots, x_i(N))$. The best previous position (the position giving the best fitness value) of the ith particle is recorded and represented as $P_i = (p_i(1), p_i(2), \ldots\ldots, p_i(N))$. The best position ever achieved by the all particles is denoted by P_g, which also has dimension N. The velocity of the ith particle is represented as $V_i = (v_i(1), v_i(2), \ldots\ldots, v_i(N))$. In each generation, the velocity and position information of the ith particle is updated as (Kennedy and Eberhart, 1995),

$$\begin{cases} v_i(n) = wv_i(n) + c_1 r_1 (p_i(n) - x_i(n)) + c_2 r_2 \left(p_g(n) - x_i(n)\right) \\ x_i(n) = x_i(n) + v_i(n) \end{cases} \quad \text{(A1)}$$

where $1 \le n \le N$, w is the inertia weight, c_1 and c_2 are two positive velocity constants, r_1 and r_2 are two random numbers in the range [0,1]. In Equation (A1), it can be seen that the current velocity of nth particle consists of three components: a. the contribution of previous velocity wv_i; b. the contribution of the best position ever achieved by current particle $c_1 r_1 (p_i(n) - x_i(n))$; c. the contribution of the best swarm position $c_1 r_1 (p_g(n) - x_i(n))$. Actually, the component $c_1 r_1 (p_i(n) - x_i(n))$ denotes the local effect of the particles and the third component $c_1 r_1 (p_g(n) - x_i(n))$ means that the global swarm effect is considered. The two random parameters r_1 and r_2 can allow the optimization to escape from the local optima. It is the updating scheme given in Equation (A1) that provides PSO powerful ability to search for the global optima. PSO used a iteration procedure to solve the optimization problem, and further details of PSO algorithm can be found in the relevant literature (Kennedy, 1998; Kennedy and Eberhart, 1995; Shi and Eberhart, 1998; Trelea, 2003).

Table 6
EMD for the 200-run Monte Carlo simulation.

SNR	MNS	LORETA	l_1 norm	3SCO	FOCUSS
Noise-free	0.51 ± 1.82	0.62 ± 3.46	0.07 ± 0.14	0.02 ± 0.09	0.04 ± 0.20
20 dB	1.86 ± 2.15	1.45 ± 1.17	0.31 ± 1.13	0.17 ± 0.76	1.01 ± 3.45

References

Auranen, T., Nummenmaa, T.A., Hamalainen, M.S., Jaaskelainen, I.P., Lampinen, J., Vehtari, A., Samsa, M., 2005. Bayesian analysis of the neuromagnetic inverse problem with l-p norm priors. NeuroImage 26 (3), 870–884.
Babiloni, F., Babiloni, C., Carducci, F., Romani, G.L., Rossini, P.M., Angelone, L.M., Cincotti, F., 2003. Multimodal integration of high-resolution EEG and functional magnetic resonance imaging data: a simulation study. NeuroImage 19 (1), 1–15.
Corbetta, M., Shulman, G.L., 2002. Control of goal-directed and stimulus-driven attention in the brain. Nat. Rev. Neurosci. 3 (3), 201–215.
Corbetta, M., Akbudak, E., Conturo, T.E., Snyder, A.Z., Ollinger, J.M., Drury, H.A., Lineweber, M.R., Petersen, S.E., 2002. A common network of functional areas for attention and eye movements. Neuron 21 (4), 761–773.
Dale, A.M., Sereno, M.I., 1993. Improved localization of cortical activity by combining EEG and MEG with MRI cortical surface reconstruction: a linear approach. J. Cogn. Neurosci. 5, 162–176.
Di Russo, F., Martinez, A., Hillyard, S.A., 2003. Source analysis of event-related cortical activity during visuo-spatial attention. Cereb. Cortex 13 (5), 486–499.
Ding, L., 2009. Reconstructing cortical current density by exploring sparseness in the transform domain. Phys. Med. Biol. 54 (9), 2683–2697.
Ding, L., He, B., 2008. Sparse source imaging in EEG with accurate field modeling. Hum. Brain Mapp. 29 (9), 1053–1067.
Donoho, D.L., Elad, M., 2003. Optimally sparse representation in general (nonorthogonal) dictionaries via l(1) minimization. Proc. Natl. Acad. Sci. U. S. A. 100 (5), 2197–2202.
Donoho, D.L., Johnstone, I.M., 1994. Ideal spatial adaptation by wavelet shrinkage. Biometrika 81 (3), 425–455.
Eberhart, R.C., Kennedy, J., 1995. A new optimizer using particle swarm theory. Paper presented at: IEEE Proceedings of the sixth international symposium on micro machine and human science (Nagoya, Japan, IEEE).
Eberhart, R.C., Shi, Y., 1998. Comparison between genetic algorithms and particle swarm optimization. Paper presented at: Proc 7th Annual Conf on Evolutionary Programming (Berlin, Springer).
Fu, S., Fan, S., Chen, L., Zhuo, Y., 2001. The attentional effects of peripheral cueing as revealed by two event-related potential studies. Clin. Neurophysiol. 112 (1), 172–185.
Fuchs, M., Drenckhahn, R., 1998. An improved boundary element method for realistic volume-conductor modeling. IEEE Trans Biomed. Eng. 45 (8), 980–997.

Gorodnitsky, I.F., Rao, B.D., 1997. Sparse signal reconstruction from limited data using FOCUSS: a re-weighted minimum norm algorithm. IEEE Trans. SP 45 (3), 600–616.
Grave de Peralta Menendez, R., Gonzalez Andino, S.L., Morand, S., Michel, C.M., Landis, T., 2000. Imaging the electrical activity of the brain: ELECTRA. Hum. Brain Mapp. 9 (1), 1–12.
Grova, C., Daunizeau, J., Lina, J.-M., Benar, C.G., Benali, H., Gotman, J., 2006. Evaluation of EEG localization methods using realistic simulations of interictal spikes. NeuroImage 29 (3), 734–753.
Hamalainen, M.S., Ilmoniemi, R.J., 1984. Interpreting measured magnetic fields of the brain: estimates of current distributions. In Technical Report TKK-F-A559 (Helsinki University of Technology).
Haufe, S., Nikulin, V., Ziehe, A., Muller, K., Nolte, G., 2008. Combining sparsity and rotational invariance in EEG/MEG source reconstruction. NeuroImage 42 (2), 726–738.
Hauk, O., 2004. Keep it simple: a case for using classical minimum norm estimation in the analysis of EEG and MEG data. NeuroImage 21 (4), 1612–1621.
Hu, X., Shi, Y., Eberhart, R.C., 2004. Recent advances in particle swarm. Paper presented at: Proceedings of the IEEE International Conference on Evolutionary Computation (IEEE).
Kennedy, J., 1998. The behavior of particles. Paper presented at: Proc 7th Annual Conf on Evolutionary Programming (Berlin, Springer).
Kennedy, J., Eberhart, R.C., 1995. Particle swarm optimization. Paper presented at: Proc IEEE Int Conf on Neural Networks, (Perth).
Land, S., Friedman, J., 1996. Variable fusion: a new method of adaptive signal regression. Technical Report Department of Statistics, Stanford University, Stanford.
Liu, H.S., Gao, X.R., 2004. A recursive algorithm for the three-dimensional imaging of brain electric activity: shrinking LORETA-FOCUSS. IEEE Trans. Biomed. Eng. 51 (10), 1794–1802.
Malioutov, D.M., Cetin, M., Willsky, A.S., 2004. Optimal sparse representations in general overcomplete bases. Paper presented at: IEEE International Conference on Acoustics, Speech, and Signal Processing (IEEE).
Michel, C.M., Murray, M., Lantz, G., Gonzalez, S., Spinelli, L., Grave de Peralta, R., 2004. EEG source imaging. Clin. Neurophysiol. 115 (10), 2195–2222.
Mosher, J.C., Lewis, P.S., Leahy, R.M., 1992. Multiple dipole modeling and localization from spatio-temporal MEG data. IEEE Trans. Biomed. Eng. 39 (6), 541–547.
Nummenmaa, A., Auranen, T., Hamalainen, M.S., Jaaskelainen, I.P., Lampinen, J., Sams, M., Vehtaria, A., 2007. Hierarchical Bayesian estimates of distributed MEG sources: theoretical aspects and comparison of variational and MCMC methods. NeuroImage 35 (2), 669–685.
Olshausen, B.A., Field, D.J., 1996. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. Nature 381 (6583), 607–609.
Op de Beeck, H.P., Haushofer, J., Kanwisher, N.G., 2008. Interpreting fMRI data: maps, modules and dimensions. Nat. Rev. Neurosci. 9 (2), 123–135.
Ou, W., Hamalainen, M.S., Golland, P., 2009. A distributed spatio-temporal EEG/MEG inverse solver. NeuroImage 44 (3), 932–946.
Parsopoulos, K.E., Vrahatis, M.N., 2002. Recent approaches to global optimization problems through particle swarm optimization. Natural Computing 1, 235–306.
Pascual-Marqui, R.D., 1999. Review of methods for solving the EEG inverse problem. Int. J. Bioelectromagn. 1 (1), 75–86.
Pascual-Marqui, R.D., Michel, C.M., Lehmann, D., 1994. Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain. Int. J. Psychophysiol. 18 (1), 49–65.
Qiu, L., Li, Y., Yao, D., 2005. A feasibility study of EEG dipole source localization using particle swarm optimization. Paper presented at: Proceedings of the IEEE International Conference on Evolutionary Computation (IEEE).
Rubner, Y., Tomasi, G., Guibas, L., 2000. The earth mover's distance as a metric for image retrieval. Int. J. Comput. Vis. 40 (2), 99–121.
Scherg, M., von Cramon, D., 1985. Two bilateral sources of the late AEP as identified by a spatio-temporal dipole model. Electroencephalogr. Clin. Neurophysiol. 62 (1), 32–44.
Shi, Y., Eberhart, R.C., 1998. Parameter selection in particle swarm optimization. Paper presented at: Evolutionary Programming VII: Proc. EP (New York, Springer-Verlag).
Silva, C., Maltez, J.C., Trindade, E., 2004. Evaluation of L1 and L2 minimum norm performances on EEG localizations. Clin. Neurophysiol. 115 (7), 1657–1668.
Simoncelli, E.P., Olshausen, B.A., 2001. Natural image statistics and neural representation. Annu. Rev. Neurosci. 24 (1), 193–216.
Srebro, R., Oguz, R.M., 1997. Estimating cortical activity from VEPs with the shrinking ellipsoid inverse. EEG Clin. Neurophysiol. 102 (4), 343–355.
Tibshirani, R., 1996. Regression shrinkage and selection via the Lasso. J. R. Stat. Soc. B 58 (1), 267–288.
Tibshirani, R., Saunders, M., Rosset, S., Zhu, J., Knight, K., 2005. Sparsity and smoothness via the fused lasso. J. R. Stat. Soc. Ser. B 67, 91–108.
Tikhonov, A.N., Arsenin, V.Y., 1977. Solutions of Ill-Posed Problems (trans. from Russian). Wiley, New York.
Trelea, I.C., 2003. The particle swarm optimization algorithm: convergence analysis and parameter selection. Inf. Process. Lett. 85 (6), 317–325.
Valdes-Sosa, P.A., Vega-Hernandez, M., Sanchez-Bornot, J.M., Martinez-Montes, E., Bobes, M.A., 2009. EEG source imaging with spatio-temporal tomographic nonnegative independent component analysis. Hum. Brain Mapp. 30 (6), 1898–1910.
Vega-Hernandez, M., Martynez-Montes, E., Sanchez-Bornot, J.M., Lage-Castellanos, A., Valdes-Sosa, P.A., 2008. Penalized least squares methods for solving the EEG inverse problem. Stat. Sin. 18, 1535–1551.
von Helmholtz, H., 1853. Ueber einige Gesetze der Vertheilung elektrischer Strome in korperlichen Leitern, mit Anwendung auf die thierischelektrischen Versuche. Ann. Phys. Chem. 89, 211–233 353-357.
Wachowiak, M.P., Smolikova, R., Zheng, Y.F., 2004. An approach to multimodal biomedical image registration utilizing particle swarm optimization. IEEE Trans. Evol. Comput. 8 (3), 289–301.
Wang, J.Z., Williamson, S.J., Kaufman, L., 1992. Magnetic source images determined by a lead-field analysis: the unique minimum-norm least-squares estimation. IEEE Trans. Biomed. Eng. 39 (7), 665–675.
Wipf, D., Nagarajan, S., 2009. A unified Bayesian framework for MEG/EEG source imaging. NeuroImage 44 (3), 947–966.
Xu, P., Tian, Y., Chen, H., Yao, D., 2007. Lp norm iterative sparse solution for EEG source localization. IEEE Trans. BME 54 (3), 400–409.
Xu, P., Tian, Y., Lei, X., Hu, X., Yao, D., 2008. Equivalent charge source model based iterative maximum neighbor weight for sparse EEG source localization. Ann. Biomed. Eng. 36 (12), 2051–2067.
Yao, D., 1996. The equivalent source technique and cortical imaging. Electroencephalogr. Clin. Neurophysiol. 98 (6), 478–483.
Yao, D., 2003. High-resolution EEG mapping: an equivalent charge-layer approach. Phys. Med. Biol. 48 (13), 1997–2011.
Yao, D., He, B., 2001. A self-coherence enhancement algorithm and its application to enhancing 3D source estimation from EEGs. Ann. Biomed. Eng. 29 (11), 1019–1027.
Yao, D., Wang, L., Dremstrup, K.N., Arendt-Nielsen, L., Chen, C.N., 2004. Cortical mapping of EEG alpha power using a charge layer model. Brain Topogr. 17 (2), 65–71.
Yao, J., Dewald, P.A., 2005. Evaluation of different cortical source localization methods using simulated and experimental EEG data. NeuroImage 25 (2), 369–382.
Zou, H., Hastie, T., 2005. Regularization and variable selection via the elastic net. J. R. Stat. Soc. Ser. B 67, 301–320.