AN ADAPTIVE FIXED-POINT IVA ALGORITHM APPLIED TO MULTI-SUBJECT COMPLEX-VALUED FMRI DATA

Li-Dan Kuang [1], Qiu-Hua Lin [1], Xiao-Feng Gong [1], Fengyu Cong [2,3], and Vince D. Calhoun[4,5]

[1] School of Information and Communication Engineering, Dalian University of Technology, Dalian 116024, China
qhlin@dlut.edu.cn
[2] Department of Biomedical Engineering, Dalian University of Technology, China
[3] Department of Mathematical Information Technology, University of Jyvaskyla, Finland
[4] The Mind Research Network, Albuquerque, NM 87106, USA
[5] Department of Electrical and Computer Engineering, University of New Mexico, Albuquerque, NM 87131, USA

ABSTRACT

Independent vector analysis (IVA) has exhibited great potential for the group analysis of magnitude-only fMRI data, but has rarely been applied to native complex-valued fMRI data. We propose an adaptive fixed-point IVA algorithm by taking into account the extremely noisy nature, large variability of the source component vector (SCV) distribution, and non-circularity of the complex-valued fMRI data. The multivariate generalized Gaussian distribution (MGGD) is exploited to match the SCV distribution based on nonlinearity, the shape parameter of MGGD is estimated using maximum likelihood estimation, and the nonlinearity is updated in the dominant SCV subspace to achieve denoising goal. In addition, the pseudo-covariance matrix is incorporated into the algorithm to represent the non-circularity. Experimental results from simulated and actual fMRI data demonstrate significant improvements of our algorithm over a complex-valued IVA-G algorithm and several circular and noncircular fixed-point IVA variants.

Index Terms—Independent vector analysis (IVA), complex-valued fMRI data, non-circularity, subspace, nonlinearity

1. INTRODUCTION

Independent component analysis (ICA) has been widely applied to the analysis of functional magnitude resonance imaging (fMRI) data. When performing ICA of multi-subject fMRI data, as typically done by group ICA [1], the independence of spatial maps (SMs) is assumed to achieve source separation. Independent vector analysis (IVA), as a kind of joint ICA, exploits the dependence of similar SMs across multi-subject fMRI datasets in addition to the independence of distinct SMs. As a result, IVA can provide superior performance in terms of capturing inter-subject variability [2-4].

IVA was originally proposed by Kim *et al.* in 2006 to solve the permutation problem in frequency-domain blind source separation (BSS) of speech signals [5]. Afterwards, a number of IVA algorithms focused similarly on the frequency-domain solution of convolutive BSS. Typical algorithms include IVA using the multivariate Laplace distribution (IVA-L) [6, 7]; fast fixed-point IVA (FIVA) employing a spherically symmetric, exponential norm distribution (SEND) or spherically symmetric Laplace (SSL) distribution [8]; non-circular FIVA (non-FIVA) utilizing the nonlinear functions to match the SCV distributions [9]; and IVA assuming multivariate generalized Gaussian distribution (MGGD) [10]. In addition, an adaptive IVA algorithm was proposed for separating the convolutively mixed acoustic signals. The multivariate Gaussian mixture model was utilized as the source prior, and the parameters of which was adaptively estimated using the EM algorithm [11]. Note that these algorithms are essentially complex-valued. However, they are tuned to separate the mixed speech signals, thus may obtain unsatisfying results for group analysis of the complex-valued fMRI data.

The first application of IVA to multi-subject fMRI data was presented by Lee *et al.*, a multivariate Laplace distribution was utilized (IVA-L), and magnitude-only fMRI data were analyzed [6, 7]. There are also other candidate algorithms that can be used to analyze magnitude-only fMRI data. For example, IVA using the multivariate Gaussian distribution (IVA-G) [12], combined IVA-G and IVA-L (called IVA-GL, using IVA-G to initialize the demixing matrix and IVA-L to perform the subsequent separation) [12], IVA with the Kotz family of distribution [13], and an adaptive MGGD-based IVA algorithm [14]. Some of these

algorithms have found promising results from magnitude-only fMRI data [2-4, 6, 7].

Although magnitude-only fMRI data are extensively studied, fMRI data are initially acquired as complex-valued image pairs including magnitude and phase information. Analysis of complex-valued fMRI data can provide additional insights beyond magnitude-only data [15, 16]. However, to the best of our knowledge, there has been no application of IVA to the complex-valued fMRI data. Due to the high noise level and the large variability of the SCV distribution, it is hard to obtain good results by using the IVA algorithms proposed for frequency-domain speech separation, or by using the complex-valued IVA algorithm with a fixed noncircular multivariate Gaussian distribution [17].

As such, we propose an adaptive fixed-point IVA algorithm in an effort to deal with the analysis of complex-valued fMRI data. Since MGGD contains multivariate Gaussian and Laplace distributions, we adaptively learn the SCV distribution using MGGD-based nonlinearity, and estimate the shape parameters using maximum likelihood estimation (MLE). As the complex-valued components of fMRI are naturally noncircular [18, 19], we incorporate the contribution of the pseudo-covariance matrix into the algorithm. Furthermore, motivated by the strategy of subspace denoising [20], we update the MGGD-based nonlinearity in the dominant SCV subspace to address the noise problem. Simulated and actual fMRI data are used to evaluate the proposed algorithm.

2. FIXED-POINT IVA

Assuming there are K subjects, the IVA model for multi-subject fMRI analysis is:

$$\mathbf{x}^{(k)}(m) = \mathbf{A}^{(k)}\mathbf{s}^{(k)}(m), 1 \le k \le K \tag{1}$$

where $\mathbf{s}^{(k)}(m) = [s_1^{(k)}(m),\ldots,s_N^{(k)}(m)]^T \in \mathbb{C}^N$ is a zero-mean and unit-variance SM source vector of subject k, N is the number of SM sources, $1 \le m \le M$, and M is the total number of the in-brain voxels obtained by flattening the volume image data. $\mathbf{x}^{(k)}(m) = [x_1^{(k)}(m),\ldots,x_N^{(k)}(m)]^T \in \mathbb{C}^N$ is the observed fMRI data of subject k, which is generally compressed by PCA and whitened. $\mathbf{A}^{(k)} \in \mathbb{C}^{N\times N}$ is the mixing matrix that contains the time courses (TCs) information. The nth SCV is denoted as $\mathbf{s}_n(m) = [s_n^{(1)}(m),\ldots,s_n^{(K)}(m)]^T \in \mathbb{C}^K$, $n = 1,\ldots,N$. Hereafter, we omit m for simplicity. By learning the demixing matrix $\mathbf{W}^{(k)}$, IVA estimates the nth SM component of subject k as $y_n^{(k)} = (\mathbf{w}_n^{(k)})^H \mathbf{x}^{(k)}$, where $\mathbf{w}_n^{(k)}$ is the nth column of $\mathbf{W}^{(k)}$.

IVA generally minimizes the mutual information among the estimated SCVs [7, 8]:

$$\mathcal{I}_{IVA} = -\sum_{n=1}^{N} \mathrm{E}\left[\log p(\mathbf{y}_n)\right] - \sum_{k=1}^{K} \log\left|\det \mathbf{W}^{(k)}\right| - C \tag{2}$$

where $p(\mathbf{y}_n)$ is the multivariate probability density distribution of SCVs, C is the constant term $H(\mathbf{x}^{(1)},\ldots,\mathbf{x}^{(K)})$.

The FIVA algorithm [8] was extended from the complex-valued fastICA algorithm [21], in which the demixing matrices $\mathbf{W}^{(k)}$ are orthonormal in each iteration, thus the second term of Eq. (2) is zero. Eventually, the IVA objective function is minimized by minimizing the first term of Eq. (2) as follows:

$$\mathcal{I}_{FIVA} = -\sum_{n=1}^{N} \mathrm{E}\left[\log p(\mathbf{y}_n)\right] = \sum_{n=1}^{N} \mathrm{E}\left[G\left(\left|\mathbf{y}_n\right|^2\right)\right] \tag{3}$$

The SCV distribution $p(\mathbf{y}_n)$ is directly related to a real-valued nonlinear function $G(\cdot)$.

3. PROPOSED ALGORITHM

The subspace method is typical for denoising. For the complex-valued fMRI data, the first eigenvalue of the covariance matrices $\mathbf{C}_n = \frac{1}{M}\sum_{m=1}^{M} \left|\mathbf{y}_n(m)\right|\left|\mathbf{y}_n^H(m)\right|$ of each SCV $\mathbf{y}_n$ is much larger than the other eigenvalues (referring to Fig. 1). This indicates that the nonlinearity $G(\cdot)$ in Eq. (3) can be learned within the one-dimensional subspace spanned by the dominant eigenvector [20]. Assume the dominant eigenvalue and eigenvector are λ_n and $\mathbf{v}_n = [v_{1n},\ldots,v_{Kn}]^T$, and let $q_n = \lambda_n\left(\mathbf{v}_n^T\left|\mathbf{y}_n\right|\right)^2 = \lambda_n\left(\sum_{k=1}^{K} v_{kn}\left|y_n^{(k)}\right|\right)^2$, we use the following nonlinear function for Eq. (3):

$$G(q_n) = (q_n)^{\beta_n} = \left[\lambda_n\left(\sum_{k=1}^{K} v_{kn}\left|y_n^{(k)}\right|\right)^2\right]^{\beta_n} \tag{4}$$

and the MGGD-based SCV distribution [22, 23]:

$$p(\mathbf{y}_n) = \mathcal{K}_n\left|\mathbf{\Sigma}_n\right|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}\left(\left|\mathbf{y}_n\right|^T \mathbf{\Sigma}_n^{-1}\left|\mathbf{y}_n\right|\right)^{\beta_n}\right\} \tag{5}$$

where $\mathcal{K}_n = \frac{K\Gamma(K/2)}{\pi^{K/2}\Gamma(1+K/(2\beta_n))2^{1+K/(2\beta_n)}}$, $\Gamma(\cdot)$ is gamma function; β_n is the shape parameter, $\beta_n = 1$ is the multivariate Gaussian distribution and $\beta_n = 0.5$ corresponds to the multivariate Laplace distribution; and $\mathbf{\Sigma}_n$ is the symmetric positive definite matrix.

We estimate the shape parameter β_n using the MLE method with Newton-Raphson optimization [23] at each iteration:

$$\beta_n \leftarrow \beta_n - \frac{\log L\left(\beta_n; \mathbf{y}_n(1),\ldots,\mathbf{y}_n(M)\right)}{\partial \log L\left(\beta_n; \mathbf{y}_n(1),\ldots,\mathbf{y}_n(M)\right)/\partial\beta_n} \tag{6}$$

where the log-likelihood function of β_n is given by

$$\begin{aligned}&\log L\left(\beta_n; \mathbf{y}_n(1),\ldots,\mathbf{y}_n(M)\right)\\ &= M\log\mathcal{K}_n - \frac{M}{2}\log\left|\mathbf{\Sigma}_n\right| - \frac{1}{2}\sum_{m=1}^{M}\left[\left|\mathbf{y}_n(m)\right|^T \mathbf{\Sigma}_n^{-1}\left|\mathbf{y}_n(m)\right|\right]^{\beta_n}\end{aligned} \tag{7}$$

and $\mathbf{\Sigma}_n$ is specifically defined to incorporate the subspace characteristics:

$$\mathbf{\Sigma}_n^{-1} = \lambda_n \begin{bmatrix} v_{1n}^2 & v_{1n}v_{2n} & \cdots & v_{1n}v_{Kn} \\ v_{1n}v_{2n} & v_{2n}^2 & \cdots & v_{2n}v_{Kn} \\ \vdots & \vdots & \ddots & \vdots \\ v_{1n}v_{Kn} & v_{2n}v_{Kn} & \cdots & v_{Kn}^2 \end{bmatrix} \tag{8}$$

Finally, by exploiting the noncircularity of the complex-valued fMRI data, we explicitly utilize the pseudo-covariance matrix $\mathrm{E}\{\mathbf{x}^{(k)}(\mathbf{x}^{(k)})^T\}$ in the learning rule of $\mathbf{w}_n^{(k)}$ [9]:

$$\begin{aligned} \mathbf{w}_n^{(k)} \leftarrow & -\mathrm{E}\{y_n^{(k)*} G'(q_n)\mathbf{x}^{(k)}\} \\ & + \mathrm{E}\{G'(q_n) + |y_n^{(k)}|^2 G''(q_n)\}\mathbf{w}_n^{(k)} \\ & + \mathrm{E}\{\mathbf{x}^{(k)}(\mathbf{x}^{(k)})^T\}\mathrm{E}\{(y_n^{(k)*})^2 G''(q_n)\}\mathbf{w}_n^{(k)*} \end{aligned} \tag{9}$$

where $G'(\cdot)$ and $G''(\cdot)$ are the first and second derivatives of the nonlinear function $G(\cdot)$:

$$G'(q_n) = \beta_n (q_n)^{\beta_n - 1} \tag{10}$$

$$G''(q_n) = \beta_n (\beta_n - 1)(q_n)^{\beta_n - 2} \tag{11}$$

After updating $\mathbf{W}^{(k)}$, a decorrelation step is followed [8]:

$$\mathbf{W}^{(k)} \leftarrow (\mathbf{W}^{(k)}(\mathbf{W}^{(k)})^H)^{-1/2}\mathbf{W}^{(k)} \tag{12}$$

4. EXPERIMENTS AND RESULTS

We generated ten simulated complex-valued fMRI datasets based on a real single-subject simulated fMRI data with eight components (http://mlsp.umbc.edu). The phase ranges of each SM and each TC are uniformly distributed from $-\pi/18$ to $\pi/18$, thus all components are non-circular. For different subjects, the SM voxels are randomly decreased by 20%, 50% and 80% with respect to the original ones (denoted as Δs = 20%, 50% and 80%) to simulate the inter-subject variability. The complex-valued Gaussian noise was added to the mixed data at CNR = -10dB, -5dB, -3dB, 0dB, 5dB and 10dB, respectively. The actual fMRI data are from 16 subjects performing a finger-tapping motor task (referring to [16] for details), and were filtered by a 10~80mHz band-pass filter. Fig. 1 shows the eigenvalues of the covariance matrices of all estimated SCVs from a case of simulated fMRI data (Δs=50%, CNR=5dB) and the actual fMRI data. The first eigenvalue of each SCV is much larger than the other eigenvalues for both datasets. This verifies the feasibility of the subspace nonlinearity.

We compare our algorithm (called AFIVA) with five IVA algorithms including the complex-value IVA-G with a fixed noncircular multivariate Gaussian distribution [17], FIVA [8], non-FIVA [9], FIVAs and non-FIVAs that we added the subspace strategy into FIVA and non-FIVA. The nonlinearity based on SSL distribution was used for these four fixed-point algorithms. We used the number of components N = 8 for the simulated data and N = 40 for the actual fMRI data.

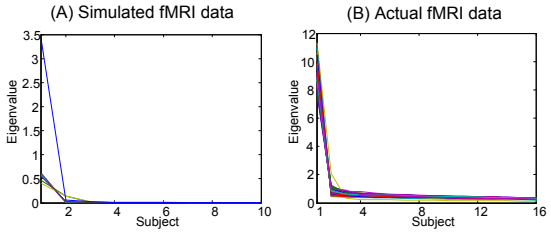

Fig. 1. The eigenvalues of the covariance matrices of all estimated SCVs from a case of simulated fMRI data (Δs=50%, CNR=5dB) (A) and the actual fMRI data (B).

4.1. Simulated fMRI Data

We used two indices to test the performance: (1) the joint normalized inter-symbol-interference (JISI) ranging from 0 to 1, and 0 indicates an ideal separation [17]; (2) an error rate defined as the ratio of the number of wrong components to the number of subjects (i.e., K) within an SCV. A wrong component is determined if its Pearson correlation with a specific ground truth is not maximal while most of the other subject-specific components reach the maximum. Here we calculated an average error rate over eight SCVs to present a concise demonstration. The results were averaged over 20 runs. Fig. 2 includes the results of the mean and standard deviation of JISI and the average error rate from all six algorithms.

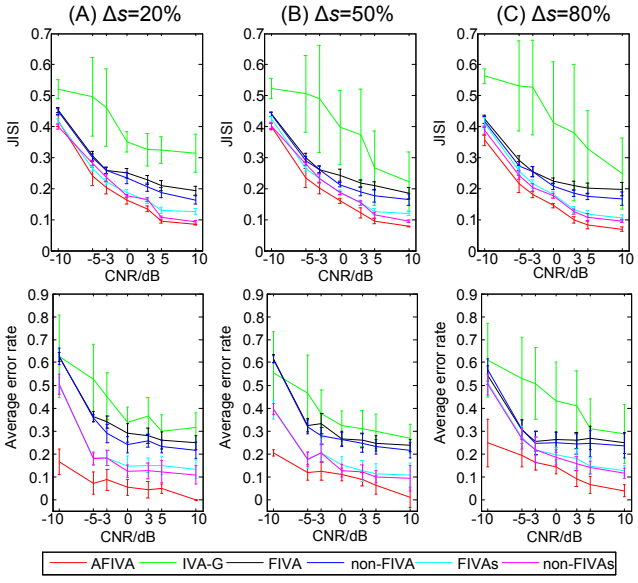

Fig. 2. Comparison of the mean and standard deviation results of JISI (top row) and average error rate (bottom row) of the six algorithms from all cases of the simulated fMRI datasets (Δs=20%~80%, CNR=-10dB~10dB).

Observing the top row of Fig. 2, we can see that our proposed method yields the smallest JISI when CNR≥-5dB, and exhibits advantages over the other algorithms at CNR =

-10dB when the SM variability is larger (Δs=80%). This is basically followed by non-FIVAs, FIVAs, non-FIVA, FIVA and IVA-G. As for the average error rate shown at the bottom row of Fig. 2, the proposed algorithm achieves the best results by significantly decreasing the number of wrong components, while non-FIVAs and FIVAs illustrate better results than non-FIVA, FIVA and IVA-G. Generally speaking, non-FIVAs and FIVAs demonstrate improvements over non-FIVA and FIVA due to utilizing the subspace nonlinearity to denoise the data, and non-FIVAs and non-FIVA are better than FIVAs and FIVA by virtue of incorporating the pseudo-covariance matrix to emphasize the noncircularity. IVA-G did not obtain satisfying results. There may be two reasons. One is that IVA-G was not developed for the noisy complex-valued fMRI data as no denoising strategy was used. The other is that IVA-G utilized a fixed multivariate Gaussian distribution [17], this does not match the various distribution of the eight SCVs, as shown in Fig. 3 by the changing shape parameters of MGGD. In contrast, our method benefits from the subspace denoising scheme to deal with the high noise level problem of the complex-valued fMRI data and the good estimation of the shape parameters to match the varying SCV distributions (see Fig. 3).

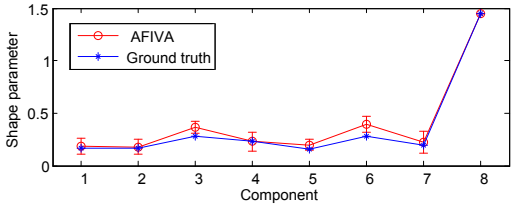

Fig. 3. Comparison of the shape parameters estimated by the proposed algorithm with the ground truth.

4.2. Actual fMRI Data

For the actual fMRI data, we analyzed a task-related component that has SM and TC priors. More precisely, we utilized the GLM reference and the model TC used in [16] as the SM and TC ground truth. To compare among the six algorithms, we calculated the Pearson correlation coefficients between the magnitude of estimates and the ground truth, and computed the error rates as well. Note all of the SM estimates were denoised using the newly proposed phase-based method [16]. The results were also averaged over 20 runs.

Fig. 4 illustrates much similar results as above. Our proposed method provides the best results followed by non-FIVAs, FIVAs, non-FIVA, FIVA and IVA-G, while non-FIVAs and FIVAs obtain similar results for the TC Pearson correlation coefficients and the error rates. Additionally, non-FIVAs and FIVAs show better results than non-FIVA and FIVA, and non-FIVAs and non-FIVA are better than FIVAs and FIVA. Fig. 5 displays a single run of estimated shape parameters for all 40 SCVs, ranging from 0.2887 to 0.4950. This demonstrates that the SCV distributions from the actual fMRI data are also varying. Among others, the estimated shape parameter of the task-related component is 0.4135 (mean) ± 0.0058 (standard deviation).

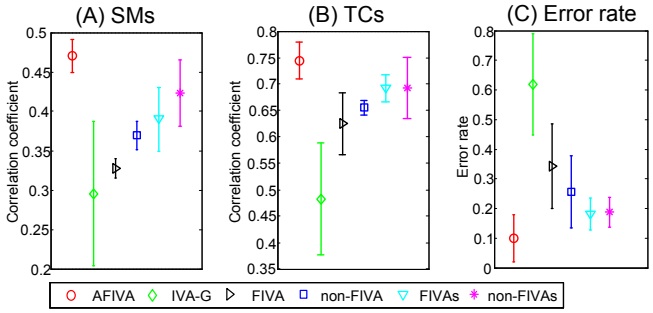

Fig. 4. Results of task-related component from actual fMRI data. (A) Pearson correlation coefficients of SMs with the GLM reference. (B) Pearson correlation coefficients of TCs with the model TC. (C) Error rates.

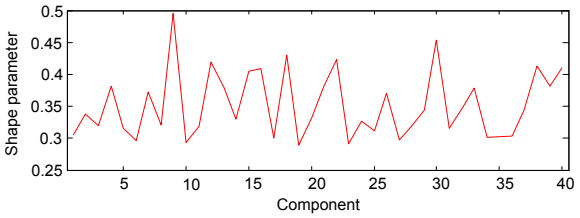

Fig. 5. An example of estimated shape parameters for all 40 SCVs from the actual fMRI data.

5. CONCLUSION

This study proposes an adaptive fixed-point IVA algorithm in an effort to deal with the challenging problems posed by the extremely noisy nature, the varying SCV distribution, and noncircularity of the complex-valued fMRI data. The MGGD-based nonlinearity is exploited to match the SCV distribution and is updated in the dominant subspace of SCV to denoise the data. The pseudo-covariance matrix is explicitly incorporated into the learning rule to emphasize the non-circularity. We test the propose algorithm using both the simulated and actual fMRI data. The experimental results show that our method can well estimate the shape parameters of MGGD, and yield significant improvements over the four circular and noncircular fixed-point IVA algorithms, and the complex-valued IVA-G algorithm.

6. ACKNOWLEDGEMENT

This work was supported by National Natural Science Foundation of China under Grants 61379012, 61105008, 61331019 and 81471742, NSF grants 0840895 and 0715022, NIH grants R01EB005846 and 5P20GM103472, the Fundamental Research Funds for the Central Universities (China, DUT14RC(3)037), and China Scholarship Council.

7. REFERENCES

[1] V.D. Calhoun, T. Adali, J.J. Pekar, and G.D. Pearlson, "A method for making group inferences from functional MRI data using independent component analysis," *Human Brain Mapping*, vol. 14, no. 3, pp. 140-151, June 2001.

[2] J.T. Dea, M. Anderson, E. Allen, V.D. Calhoun, and T. Adali, "IVA for multi-subject fMRI analysis: A comparative study using a new simulation toolbox," In: *IEEE International Workshop on Machine Learning for Signal Processing (MLSP)*, Santander, pp. 1-6, September 2011.

[3] A.M. Michael, M. Anderson, R.L. Miller, T. Adali, and V.D. Calhoun, "Preserving subject variability in group fMRI analysis: performance evaluation of GICA vs. IVA," *Frontiers in Systems Neuroscience*, vol. 8, pp. 1-8, June 2014.

[4] J. Laney, K. Westlake, S. Ma, E. Woytowicz, V.D. Calhoun, and T. Adali, "Capturing subject variability in fMRI data: a graph-theoretical analysis of GICA vs. IVA," *Journal of Neuroscience Methods*, vol. 247, pp. 32-40, May 2015.

[5] T. Kim, T. Eltoft, and T.W. Lee, "Independent vector analysis: an extension of ICA to multivariate components," In: *Independent Component Analysis and Blind Signal Separation*, ser. *Lecture Notes in Computer Science*, J. Rosca, D. Erdogmus, J. Príncipe, and S. Haykin, Eds. Berlin, Germany: Springer-Verlag, vol. 3889, pp. 165-172, March 2006.

[6] J.H. Lee, T.W. Lee, F.A. Jolesz, and S.S. Yoo, "Multivariate analysis of fMRI group data using independent vector analysis," In: *Independent Component Analysis and Blind Signal Separation*, ser. *Lecture Notes in Computer Science*, M.E. Davies, C.J. James, S.A. Abdallah, M.D. Plumbley, Eds. Berlin, Germany: Springer-Verlag, vol. 4666, pp. 633-640, September 2007.

[7] J.H. Lee, T.W. Lee, F.A. Jolesz, and S.S. Yoo, "Independent vector analysis (IVA): multivariate approach for fMRI group study," *NeuroImage*, vol. 40, no. 1, pp. 86-109, March 2008.

[8] I. Lee, T. Kim, and T.W. Lee, "Fast fixed-point independent vector analysis algorithms for convolutive blind source separation," *Signal Processing*, vol. 87, no. 8, pp. 1859-1871, August 2007.

[9] H. Zhang, L. Li, and W. Li, "Independent vector analysis for convolutive blind noncircular source separation," *Signal Processing*, vol. 92, no. 9, pp. 2275-2283, September 2012.

[10] Y. Liang, J. Harris, S.M. Naqvi, G. Chen, and J.A. Chambers, "Independent vector analysis with a generalized multivariate Gaussian source prior for frequency domain blind source separation," *Signal Processing*, vol. 105, no. 12, pp. 175-184, December 2014.

[11] I.H. Lee, J. Hao, and T.W. Lee, "Adaptive independent vector analysis for the separation of convoluted mixtures using EM algorithm," In: *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Las Vegas, pp. 145-148, April 2008.

[12] M. Anderson, T. Adali, and X.L. Li, "Joint blind source separation with multivariate Gaussian model: algorithms and performance analysis," *IEEE Transactions on Signal Processing*, vol. 60, no. 4, pp. 1672-1683, April 2012.

[13] M. Anderson, G.-S. Fu, R. Phlypo, and T. Adali, "Independent vector analysis, the kotz distribution, and performance bounds," In: *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Vancouver, pp. 3243-3247, May 2013.

[14] Z. Boukouvalas, G.-S. Fu, and T. Adali, "An efficient multivariate generalized Gaussian distribution estimator: application to IVA," *49th Annual Conference on Information Sciences and Systems (CISS)*, Baltimore, pp. 1-4, March 2015.

[15] D.B. Rowe, "Modeling both the magnitude and phase of complex-valued fMRI data," *Neuroimage*, vol. 25, no. 4, pp.1310-1324, 2005.

[16] M.C. Yu, Q.H. Lin, L.D. Kuang, X.F. Gong, F. Cong, and V.D. Calhoun, "ICA of full complex-valued fMRI data using phase information of spatial maps," *Journal of Neuroscience Methods*, vol. 249, pp. 75-91, July 2015.

[17] M. Anderson, X.L. Li, and T. Adali, "Complex-valued independent vector analysis: application to multivariate Gaussian model," *Signal Processing*, vol. 92, no. 8, pp. 1821-1831, August 2012.

[18] P.L. Schreier, T. Adali, and L.L. Scharf, "On ICA of improper and noncircular sources," In: *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Taipei, pp. 3561-3564, April 2009.

[19] P.L. Schreier and L.L. Scharf, *Statistical Signal Processing of Complex Valued Data: The Theory of Improper and Noncircular Signals*, Cambridge University Press, 2010.

[20] Y. Na, J. Yu, and B. Chai, "Independent vector analysis using subband and subspace nonlinearity," *EURASIP Journal on Advances in Signal Processing*, vol. 74, no. 1, pp. 1-16, April 2013.

[21] E. Bingham and A. Hyvärinen, "A fast fixed-point algorithm for independent component analysis of complex valued signals," *International Journal of Neural Systems*, vol. 10, no. 1, pp.1-8. March 2000.

[22] E. Gómez, M.A. Gómez-Villegas, and J.M. Marín, "A multivariate generalization of the power exponential family of distributions," *Communications in Statistics-Theory and Methods*, vol. 27, no. 3, pp. 589-600, January 1998.

[23] F. Pascal, L. Bombrun, J.Y. Tourneret, and Y. Berthoumieu, "Parameter estimation for multivariate generalized Gaussian distributions," *IEEE Transactions on Signal Processing*, vol. 61, no. 23, pp. 5960-5971, February 2013.