Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 154 (2019) 524–530

Procedia Computer Science

www.elsevier.com/locate/procedia

8th International Congress of Information and Communication Technology, ICICT 2019

Blind Source Separation Using an Adaptive Coupling of Four Separation Systems Based on Variable Momentum Term and Variable Step Size

Yuan Yuan Tian [a], Shi Feng Ou [a,*], Ying Gao [a], Qian Qian Jiang [b]

[a]School of Opto-Electronic Information Science and Technology Yantai University, Yantai 264005, Shandong Province, China
[b]School of Foreign Languages Yantai University, Yantai 264005, Shandong Province, China

Abstract

In order to improve the performance of the blind source separation (BSS) algorithm, this paper proposes a four-system adaptive coupling BSS algorithm based on variable momentum term and variable step size. The algorithm improves the convergence rate and steady-state error by adaptively coupling two medium systems. The first medium system is composed of two subsystems with different step sizes and the same large momentum term, and the convergence rate is faster. The second medium system is composed of two subsystems with different step sizes and the same small momentum term, and the steady-state error is smaller. Simulation results show that the convergence rate of the algorithm is faster and the steady-state error is smaller.

Keywords: Momentum term; step size; four separation systems; adaptive; blind source separation

1. Introduction

In recent years, many new technologies have emerged in the signal processing field, including BSS technology. The essence of BSS is to estimate the source signal from the observed mixed signal according to the independent statistical characteristics of the source signal without any prior knowledge. BSS has become one of the hot

* Corresponding author. *E-mail address:*ousfeng@126.com

10.1016/j.procs.2019.06.082

techniques , which has been widely applied in wireless communication, seismic survey, ocean sonar detection, biomedical engineering, digital image processing, speech processing and so on [1,2].

Scholars had proposed many adaptive blind source separation algorithms to solve the problem between the convergence rate and steady-state error. In the early literature, a dual-system adaptive coupling algorithm was proposed, which adaptively selected the optimal separation system in each separation stage. Although the separation performance of the algorithm was effectively improved, the algorithm had the problem of convergence in stages [3]. In order to improve the rate of convergence, some literature added the momentum term to the algorithm, and uses the gradient descent method and convex combination theory to adaptively adjust the momentum term. However, the technique has drawbacks: it did not improve the impact of step size on the algorithm [4]. Another effective solution to solve the principal contradiction is to combine the momentum term with the step size, and use the estimated value of the index of separation performance PI to adaptively adjust them, but the steady-state error was not reduced [5]. To a certain extent, these algorithms have alleviated the basic problem, but the error caused by the step size and momentum term in the operation of the algorithm will inevitably have some impact on the separation performance of the system [6].

According to these adaptive and optimal parameters of step size and momentum term, this paper proposes a new four-system adaptive coupling BSS algorithm based on variable momentum term and variable step size. The algorithm of this paper has a big separation system which is combined by two medium systems. Each medium system is composed of two different subsystems. The first medium system is composed of two subsystems with different step sizes and the same large momentum term, and the convergence rate is faster. The second medium system is composed of two subsystems with different step sizes and the same small momentum term, and the steady-state error is smaller. In order to optimize the algorithm, this paper uses the sliding factor to adaptively adjust two medium systems, and at each iteration, the best medium system is selected to make the algorithm have the best separation performance.

2. Blind Source Separation Algorithm with Momentum Terms

In an ideal blind source separation system, the linear instantaneous mixing system can be presented as: $\mathbf{x}(t)=\mathbf{As}(t)$, where A is an $m\times n$ row-full-rank mixed matrix. The source signals $s(t)=[s_1(t),s_2(t),s_3(t),\cdots s_n(t)]^T$ is the n-dimensional vector at the time t, the mixed signal $x(t)=[x_1(t),x_2(t),x_3(t),\cdots x_m(t)]^T$ is the m-dimensional vector of the observed mixed signals by m sensors.

The observed mixed signal is output by the separation system to obtain an output signal, the relationship is :

$$\mathbf{y}(t)=\mathbf{Wx}(t)=\mathbf{WAs}(t)=\left[y_1(t),y_2(t),y_3(t),\cdots \qquad\qquad ^T\right. \tag{1}$$

where **W** represents the separating matrix. $y(t)$ is the estimation of the source signal s(t).

In this paper, the natural gradient algorithm is taken as the main research object, according to the minimum mutual information criterion and KL divergence [7], the expression of the cost function is assumed to be:

$$J(t)=-\left|\mathbf{H}(x(t))\right|-\log\left|\det(\mathbf{W}(t))\right|-\sum_{i=1}^{n}E\left\{\log\left[q_i(y_i(t))\right]\right\} \tag{2}$$

From the stochastic gradient algorithm [8, 9], we can get the updated equation of the natural gradient algorithm[10, 11], which can be written as:

$$\mathbf{W}(t+1)=\mathbf{W}(t)-\mu\frac{\partial J(t)}{\partial\mathbf{W}(t)}=\mathbf{W}(t)+\mu\left\{\mathbf{I}-f\left[\mathbf{y}(t)\right]\mathbf{y}^T(t)\right\}\mathbf{W}(t) \tag{3}$$

where μ is a constant, which stands for a step size factor . f(•) is a non-linear activation function, and I represents the identity matrix.

Adaptive blind source separation algorithms are mostly classified as gradient algorithms of LMS type, which still have limitations in the process of blind source separation. In order to alleviate this problem, momentum term technology[12, 13] is incorporated into formula (3) to accelerate the convergence speed of the algorithm, and a natural gradient algorithm with momentum term is proposed [14]. The updating rule can be expressed as:

$$\mathbf{W}(t+1)=\mathbf{W}(t)+\mu\left\{\mathbf{I}-\mathrm{f}\left[\mathbf{y}(t)\right]\mathbf{y}^{T}(t)\right\}\mathbf{W}(t)+\alpha\left[\mathbf{W}(t)-\mathbf{W}(t-1)\right] \tag{4}$$

where α is a momentum factor, and $\alpha[\mathbf{W}(t)-\mathbf{W}(t-1)]$ represents the momentum term.

3. Blind Source Separation Using An Adaptive Coupling Of Four Separation Systems Based On Variable Step Size And Variable Momentum Term

The purpose of this paper is to obtain a large system by adaptively combining the two medium systems. The large system selects the medium system with the best separation performance at each iteration, so the convergence rate and steady-state error are more optimized. The first medium system $\mathbf{W}_a(t)$ is obtained by adaptive coupling of $\mathbf{W}_{a1}(t)$ and $\mathbf{W}_{a2}(t)$, and the convergence rate of this medium system is faster. The second medium system $\mathbf{W}_b(t)$ is composed of $\mathbf{W}_{b1}(t)$ and $\mathbf{W}_{b2}(t)$, and its steady-state error is small. The large system $\mathbf{W}(t)$ takes into account the advantages of the two medium systems. It has a greater convergence rate at the beginning of the algorithm separation or when the environment changes, and has a smaller steady-state error in the stable stage of the system. The algorithm structure presented in this paper is shown in Fig.1.

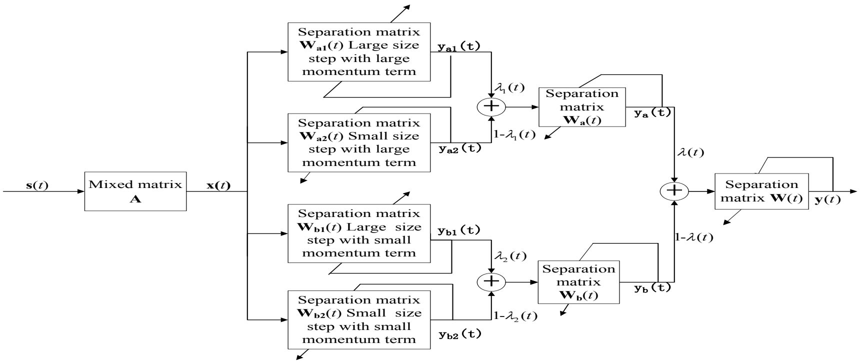


Fig.1. The frame diagram of the proposed algorithm

The separation systems of four different momentum term factors and step size factors are:

$$\mathbf{W}_{ai}(t+1)=\mathbf{W}_{ai}(t)+\mu_i\left\{\mathbf{I}-f\left[\mathbf{y}_{ai}(t)\right]\mathbf{y}_{ai}^{T}(t)\right\}\mathbf{W}_{ai}(t)+\alpha_1\left[\mathbf{W}_{ai}(t)-\mathbf{W}_{ai}(t-1)\right] \tag{5}$$

$$\mathbf{W}_{bi}(t+1)=\mathbf{W}_{bi}(t)+\mu_i\left\{\mathbf{I}-f\left[\mathbf{y}_{bi}(t)\right]\mathbf{y}_{bi}^{T}(t)\right\}\mathbf{W}_{bi}(t)+\alpha_2\left[\mathbf{W}_{bi}(t)-\mathbf{W}_{bi}(t-1)\right] \tag{6}$$

where μ_1 and μ_2 are the step size scalar factors, and $\mu_1 > \mu_2$; α_1 and α_2 respectively represent the momentum factors of the separation system, and $\alpha_1 > \alpha_2$; $\mathbf{W}_{a,bi}(t)$ represent the separation matrix of the four separation system; $y_{a,bi}(t)$ is the output signals of the four separation systems, namely:

$$\mathbf{y}_{a,bi}(t)=\mathbf{W}_{a,bi}(t)\mathbf{x}(t) \tag{7}$$

where i=1,2.

The separating matrices $\mathbf{W}_a(t)$, $\mathbf{W}_b(t)$ and $\mathbf{W}(t)$ can be expressed as:

$$\mathbf{W}_a(t)=\lambda_1(t)\mathbf{W}_{a1}(t)+\left[1-\lambda_1(t)\right]\mathbf{W}_{a2}(t) \tag{8}$$

$$\mathbf{W}_b(t)=\lambda_2(t)\mathbf{W}_{b1}(t)+\left[1-\lambda_2(t)\right]\mathbf{W}_{b2}(t) \tag{9}$$

$$\mathbf{W}(t)=\lambda(t)\mathbf{W}_a(t)+\left[1-\lambda(t)\right]\mathbf{W}_b(t) \tag{10}$$

where $\lambda_1(t)$, $\lambda_2(t)$ and $\lambda(t)$are sliding factors, which are kept in interval [0,1].The output signal of this system is represented by $y(t)$, then:

$$\mathbf{y}(t)=\mathbf{W}(t)\mathbf{x}(t) \tag{11}$$

During the operation of the algorithm, adaptive sliding factors are used to adjust different separation systems at different separation stages to better improve the algorithm. Based on stochastic gradient theory, the update equation for $\lambda_i(t)$ is expressed as :

$$\lambda_i(t+1)=\lambda_i(t)-\rho\nabla_{\lambda_i}J(t)\Big|_{\lambda-\lambda(t)} \tag{12}$$

where, ρ is a small constant. We introduce the inner product matrix to facilitate the algorithm is introduced, the definition of the formula for:

$$\langle\mathbf{C}\bullet\mathbf{D}\rangle=tr\left(\mathbf{C}^T\mathbf{D}\right) \tag{13}$$

where, the matrix inner product is < >. tr() is the track operator, which is the trace of the matrix, and the matrices C and D are the m×n dimensional matrices. Therefore, the equation (14) can be expressed as:

$$\nabla_\lambda J(t)\Big|_{\lambda-\lambda(t)}=\left\langle\frac{\partial J(t)}{\partial\mathbf{W}_{a,bi}(t)},\frac{\partial\mathbf{W}_{a,bi}(t)}{\partial\lambda(t)}\right\rangle=tr\left\{\left[\frac{\partial J(t)}{\partial\mathbf{W}_{a,bi}(t)}\right]^T\times\left[\frac{\partial\mathbf{W}_{a,bi}(t)}{\partial\lambda(t)}\right]\right\} \tag{14}$$

From (3), we can get:

$$\frac{\partial J(t)}{\partial\mathbf{W}_{a,bi}(t)}=-\left\{\mathbf{I}-f\left[\mathbf{y}_{a,bi}(t)\right]\mathbf{y}_{a,bi}{}^T(t)\right\}\mathbf{W}_{a,bi}(t) \tag{15}$$

According to the formula of (3) (8) (9),

$$\frac{\partial\mathbf{W}_{a,bi}(t)}{\partial\lambda_i(t)}=\frac{\partial\left\{\lambda_i(t)\mathbf{W}_{a,bi}(t)+\left[1-\lambda_i(t)\right]\mathbf{W}_{a,bi+1}(t)\right\}}{\partial\lambda_i(t)}=\mathbf{W}_{a,bi}(t)-\mathbf{W}_{a,bi+1}(t) \tag{16}$$

where i=1, 2.

Substitute (15) and (16) into (12):

$$\nabla_{\lambda} J(t)\Big|_{\lambda-\lambda(t)} = -tr\left\{\left\{\left\{\mathbf{I} - f\left[\mathbf{y}_{a,bi}(t)\right]\mathbf{y}_{a,bi}{}^{T}(t)\right\}\mathbf{W}_{a,bi}(t)\right\}\left[\mathbf{W}_{a,bi}(t) - \mathbf{W}_{a,bi+1}(t)\right]\right\} = -tr\left[\Gamma^{T}(t)\mathbf{H}(t)\right] \tag{17}$$

Then, the final adaptive update rule for $\lambda(t)$ is:

$$\lambda_i(t+1) = \lambda_i(t) + \rho tr\left[\mathbf{\Gamma}^{T}(t)\mathbf{H}(t)\right] \tag{18}$$

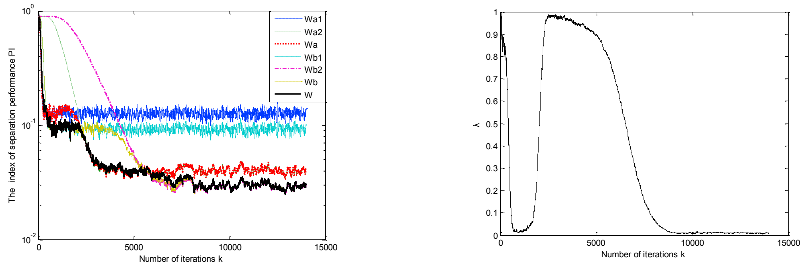
4. Simulation Experiment and Performance Analysis

The performance index PI is used to verify the validity of the algorithm's separation performance [15], which is proposed in this paper:

$$PI(G) = \left[\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{m}\frac{|g_{ij}|}{\max_j |g_{ij}|} - 1\right] + \left[\frac{1}{m}\sum_{i=1}^{n}\sum_{j=1}^{m}\frac{|g_{ij}|}{\max_i |g_{ij}|} - 1\right] \tag{19}$$

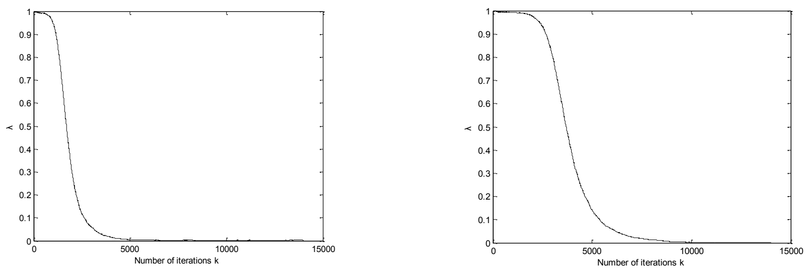
4.1. Stationary environment

First of all, the separation performance of the algorithm is verified in stationary environment.We select two under Gaussian signals with zero mean as the source signals. According to the under Gaussian property of the source signal, the nonlinear activation function is selected as a cubic function; The two step sizes are selected as μ1=0.01and μ2=0.001, and the values of momentum factor are set at α1=0.5 and α2=0.05, and the parameter is 0.05 that which is called P. Fig.4.(a) plots the convergence curve of each adaptive algorithm averaged over 100 Monte Carlo trials.



(a) (b)

Fig.2. (a) Convergence curve of algorithm (b) Convergence curve of sliding factor gg



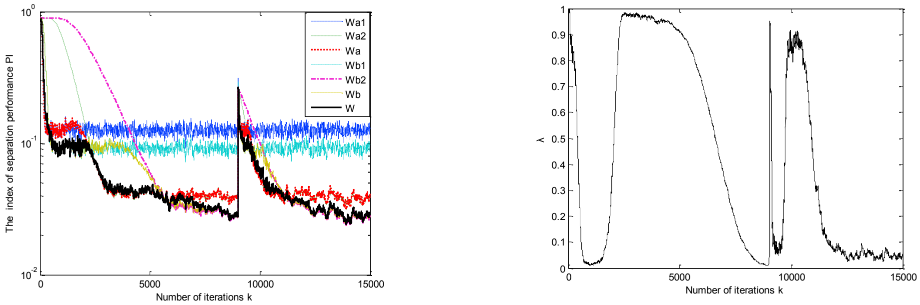
(a) (b)

Fig.3. (a) Convergence curve of sliding factor ga (b) Convergence curve of sliding factor gb

As discussed in Section 3, in the first middle system Wa(t), its separation performance index is close to 10-1 after 1000 points, it has a fast convergence rate . While the latter middle separation systems Wb(t) reaches 10-1 after about 1300 iterations, which is reduced misalignment in steady-state error. The large system W(t) coupled by the two medium systems has faster convergence rate and lower steady-state error. By analyzing Fig 2(a), we can know that the improved algorithm not only effectively avoid the shortcomings of the separation system, but also further optimize the separation effect through multiple coupling.

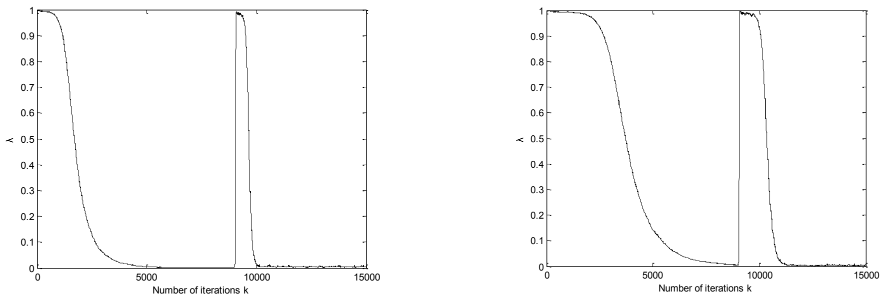
4.2. Non-stationary environment

This experiment was conducted in a non-stationary environment to verify that the algorithm had the ability to react quickly in the external environment. When the sample number parameters are 9000 points, the channel was changed abruptly in order to alter the algorithm environment. Other parameters of the algorithm are same as those of the stationary environment. After 100 Monte Carlo trials, the experimental results are shown in Fig 4(a), which again verifies the advantage of the proposed algorithm after the abrupt change in the channel.



(a) (b)

Fig.4. (a) Convergence curve of algorithm (b) Convergence curve of sliding factor gg



(a) (b)

Fig.5. (a) Convergence curve of sliding factor ga (b) Convergence curve of sliding factor gb

By analysing the variation curve and separation performance curve, it can be seen that even if the external environment changes, the algorithm can react quickly and continue to complete the convergence process. It can be concluded that the algorithm can better complete the signal separation even in the case of external interference, and adjust the proportion of each separation system in the algorithm update process by adjusting its own size.

5. Conclusions

A new algorithm called BSS for four separation systems based on variable momentum term and variable step size was proposed, which can effectively improve the separation performance of the algorithm. The good separation performance of the algorithm was verified in stationary and non-stationary conditions by simulation analysis.

6. Acknowledgement

This work is supported by the Natural Science Foundation of Shandong Province (No.ZR2017MF008).

7. References

1. Cichocki A and Amari S. Adaptive Blind Signal and Image Processing: Learning Algorithms and Application. New York: Wiley Press;2002.
2. Peng S.Y., Yang s., Wang Z.H., Zhu Y.Q. Present Situation and Development of Blind Source Separation. SHIOBOARD ELECTRONIC COUNTERMEASURE 2016;39: 54-57.
3. OU S.F., Zhao X.H., Gao Y.. Blind source separation using an adaptive combination of two separation systems. Journal of Jilin University(Engineering and Technology Edition) 2010;40:1366-1370.
4. OU S.F., Geng C., Gao Y.. Momentum Term Based Blind Source Separation Algorithm and Its Performance Modified Strategies. ACTA ELECTRONICA SINICA 2014;42:42-48.
5. Zhang T.Q., Ma B.Z., Qing X.Z., Quan S. R.. Variable-step blind source separation method with adaptive momentum factor. Journal on Communications 2017;38:16-24.
6. OU S.F., Zhao X.H., Gao Y.. Variable Step-Size Blind Source Separation Algorithm with an Auxiliary Separation System. ACTA ELECTRONICA SINICA 2009;37: 1588-1593.
7. Jimin Ye, Haihong Jin, Shuntian Lou, Kejun You. An optimized EASI algorithm. Signal Processing 2009; 89: 333-338.
8. Aapo Hyvarine, Juha Karhunen, Erkki Oja; Translated by Zhou Z. T., etc. Independent Component Analysis. Beijing: Publishing House of Electronics Industry; 2014.
9. Martin. T Hagan, Howard B. Demuth, Mark H. Beale; Translated by Dai K. etc. Neural Network Design. China Machine Press;2002.
10. Amari S. Natural Gradient Learning for Over and Under Complete Bases in ICA. Neural Computation 1999; 1: 1875-l883.
11. Cichocki A , Amari S, Adachi M, et al. Self-adaptive Neural Networks for Blind Separation of Sources.1996 IEEE International Symposium on Circuits and Systems[C],Piscataway NJ 1996;2:157-160
12. Yang X. J., Zheng J. l.. Artificial Neural Network and Blind Signal Processing. Beijing: Tsinghua University Press; 2002.
13. Gao Y., Li Y., Yang B.J.. Overview on the variable step size techniques for on-line blind source separation. Computer engineering and Applications 2007;43:75-79.
14. Lin Y.M., Lin T.S.. Improved Blind Separation Algorithm by Adding Momentum term. Journal of South China University of Technology (Natural Science Edition) 2006;34:6-9.
15. Yuan L. X., Wang W. W., Chambers J A. Variable step-size sign natural gradient algorithm for sequential blind source separation. IEEE Signal Process Letters 2005;12:589-592.