Blind Source Separation for Remote Sensing Images based on the Improved ICA Algorithm

Di SHEN, Chengfan LI*, Jingyuan YIN, Junjuan ZHAO, Dan XUE

School of Computer Engineering and Science, Shanghai University, Shanghai, 200444, China

*corresponding author email: david-0904@163.com

Abstract. In consideration of some problems including the independence and invariance of components, no noise assumption and the uncertainty of the final solution as well as the inconsistency of the features of remote sensing data in traditional independent component analysis (ICA) model, we put forward a blind source separation algorithm for remote sensing images using variational Bayesian ICA. In the proposed method, the Bayesian network is introduced into the ICA model, the Bayesian inference is used to complete the study of unknown hidden variables, and the computation is optimized by combination with the variational approximation method. Finally, the proposed method is validated by simulation and real remote sensing image tests. The result shows that the variational Bayesian ICA algorithm has both good stability and separation effect, and it overcomes the deficiency of the traditional ICA method in remote sensing application.

Introduction

Remote sensing technology has the advantages of wide coverage, large amount of information and short revisit cycle. It is able to obtain the change information of the surface of the earth timely and accurately, and widely applied in the resource investigation, environmental monitoring, etc. [1, 2] Compared with the traditional digital image, the remote sensing image contains more diverse terrain types, also the distribution is more complex. As a new signal processing technology, ICA has close association with the problem of blind source separation, and it has the ability of separating source signals (terrain components) from the mixed ones without the need of prior knowledge [3-5]. At present, the application field of the ICA has expanded from solution of mixed blind signals to signal detection, image processing and pattern recognition, etc. But the ICA model has problems containing the independence and invariance of components, no noise assumption as well as the uncertainty of the final solution [6]. These assumptions of the ICA make it not completely consistent with the actual state of remote sensing image.

This paper puts forward blind source separation for remote sensing images based on improved ICA algorithm, i.e. variational Bayesian ICA method, the algorithm is based on the analysis of the traditional ICA model and of the characteristics of remote sensing image. In the proposed method, the Bayesian network is introduced into the ICA model, we used Bayesian inference to complete the study of unknown hidden variables (independent components), and we optimized it by combination with the variational approximation algorithm to make the isolated independent components as consistent as the real situation of the surface. This paper analyzed the variational Bayesian ICA method from the aspects of both simulation images and real remote sensing image respectively, and the results of experiments verified the effectiveness of the proposed algorithm.

Variational Bayesian ICA Algorithm

Add noise information in the standard ICA model, making it the linear mixed ICA model with noise [7, 8], and the formula is shown below:

$$x(t) = As(t) + \varepsilon(t) \tag{1}$$

Where the $x(t)$ refers to a mixed signal with M, $s(t)$ refers to a source signal (hidden variable) with L,

mixed matrix A is $M \times L$, $\varepsilon(t)$ refers to Gaussian noise, which usually is a diagonal matrix with the inverse variance for Λ and the mean for zero.

In linear mixed ICA model containing noises, the probability calculation formula for mixed signal $x(t)$ is:

$$p(x \mid s, A, \Lambda) = \mid \det(\frac{1}{2\pi}\Lambda) \mid^{\frac{1}{2}} \exp[-E_D] \tag{2}$$

Where $E_D = \frac{1}{2}(x - As)^T \Lambda(x - As)$, $\det(\cdot)$ refers to the determinant value. In the model we assume the source signals are independent of each other, so the probability distribution formula for s can be expressed as $p(s) = \prod\limits_{i=1}^{L} p(s_i)$.

In Bayesian network model, all the parameters and the hidden variables of the model are represented by W, $W = \{\Lambda，A，S，q，\theta\}$. The maximum negative free energy $F[x|\theta]$ is the objective function of the model, the formula of which is shown by the following:

$$F[x \mid \theta] = \langle \log p(x, W \mid M) \rangle_{p'(W|\theta)} + H[p'(W)] \tag{3}$$

In variational Bayesian ICA model, according to the conditional independence principle of the Bayesian network, we use the approximate form of real posteriori probability distribution $p'(W|\theta)$ to calculate the partial derivatives of various parameters, and the maximum value, respectively, so:

$$p'(W_i \mid \theta) = \frac{1}{Z_i} \exp[\langle \log(x, W \mid M) \rangle_{\prod\limits_{i \neq j} p'(W_{j \neq i}|\theta_{j \neq i})}] \tag{4}$$

Where Z_i refers to specification factors, whose value varies with the variable.

In variational Bayesian ICA model, through the conditional independence assumption of Bayesian inference and source signal generation model, we can choose the approximate form of real posteriori probability distribution $p'(W|\theta)$:

$$p'(W \mid \theta) = p'(\Lambda)p'(A)p'(s \mid q)p'(q)p'(\theta) \tag{5}$$

Where $p'(\theta) = p'(\pi)p'(\mu)p'(\beta)$, $p'(s \mid q)$ can be obtained in GOM model.

According to the source signal generation model of variational Bayesian ICA, the following formula can be launched:

$$p(x \mid W) = p(x \mid A, \Lambda, s)p(s \mid q, \mu, \beta)p(q \mid \pi)p(\theta)p(A)p(\Lambda) \tag{6}$$

Substitute the approximate form of real posteriori probability distribution $p'(W|\theta)$ into maximum negative free energy F. Combining with the definition of prior distribution of the parameters in the source signal model, the best approximation true posterior probability distribution parameters of the source signal s can be deduced:

$$p(S \mid q) = \prod\limits_{i=1}^{L}\prod\limits_{t=1}^{T} N(S_i^t; \hat{\mu}_{i,q_i}^{(t)}; \hat{\beta}_{i,q_i}^{(t)}) \tag{7}$$

Similarly, we can also continuously deduce the real posterior probability distribution of other parameters and hidden variables in the variational Bayesian ICA model, and obtain the approximation value through iteratively approaching true value until convergence.

Experiments

The simulation experiments

In this paper, we randomly use vibration signal function model to construct the source signal $s_1 = (1+\sin(30\pi t))\cos(100\pi t)$ and $s_2 = (1+\sin(20\pi t))\cos(200\pi t)$, respectively. The frequency of s_1 and s_2 is 1000Hz, the sampling length is 1024 points. The source signal obtained after calculation is shown in Figure 1.
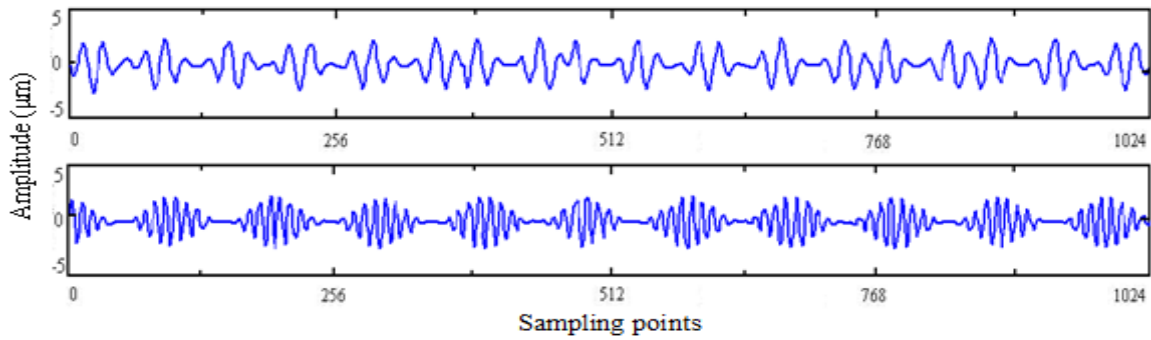
Fig.1. The simulated images of source signal s_1 (upper) and s_2 (lower)

Then we select one random 2×2 matrix A, conduct mixed processing to two source signals, and add 5db Gaussian noise to the mixed signal, respectively. The result is shown in Figure 2.
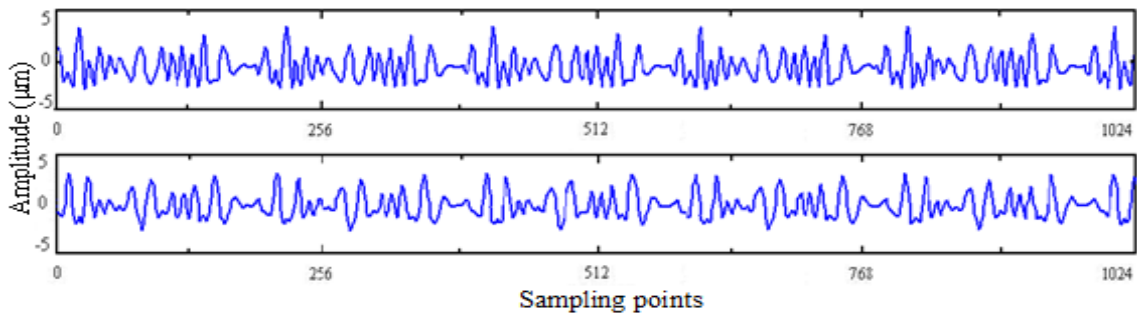

Fig.2. The mixed signals x_1(upper) and x_2 (lower) containing 5db Gaussian noise mixed by source signal s_1 and s_2

Finally, we conduct separation processing to the mixed observation signal x_1 and x_2. To contrast the performance of separating the mixed signals by variational Bayesian ICA method, this article separately conduct separation processing and make contrasts with other methods based on the characteristic matrix joint approximate diagonalization (JADE) and standard ICA. The separation results for mixed signals by three methods are shown in Figure 3.
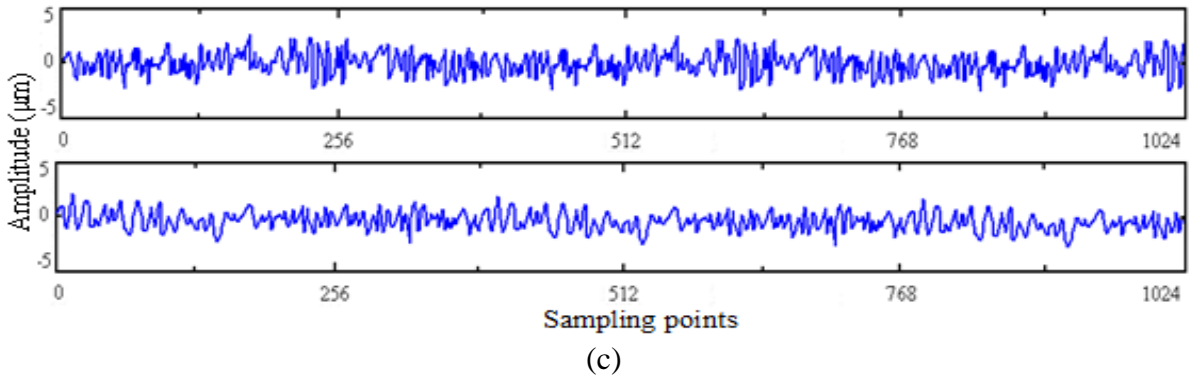

(a)

(b)

(c)
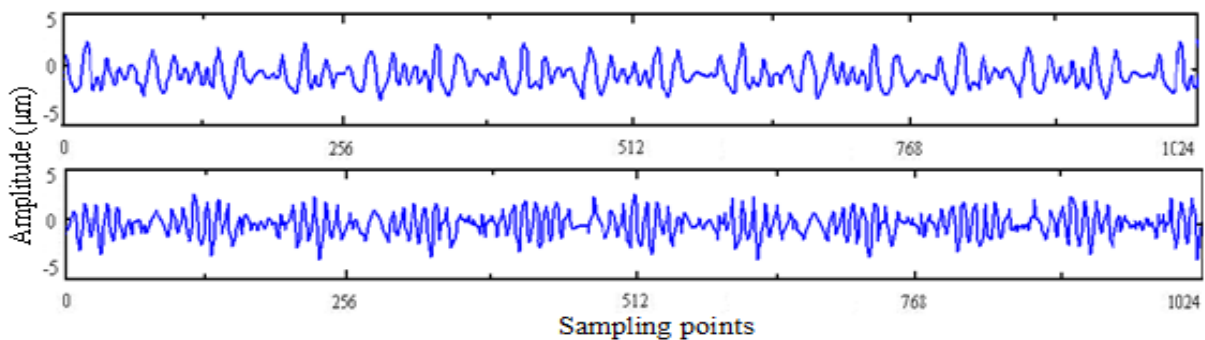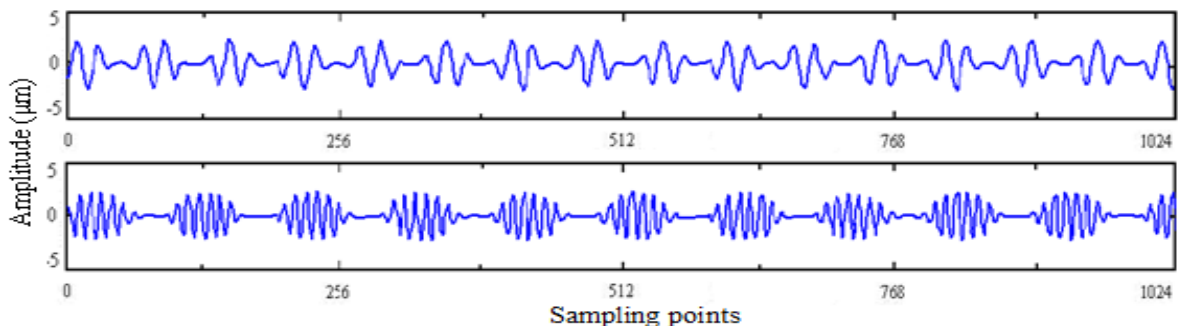
Fig.3. The isolated source signals obtained by JADE(a), variational Bayesian ICA(b) and ICA(c)

It can be seen from the Figure 3 that in the noise environment the shape of isolated signal by variational Bayesian ICA method (Figure 3b) is relatively close to that of the source signal, there is only a certain scale on the amplitude. And the effect of the isolated signal conducted by JADE method and ICA method is poorer, not only the amplitude of the separated signal scales, but also its waveform changes greatly, thus has clear difference with the source signal. In addition, this article also uses the error (ERR) indicators to make further comparison for the three above methods.

The ERR calculation formula is shown below:

$$ERR = \frac{\| \hat{s}(n) - s(n) \|_2}{\| s(n) \|_2} \times 100\% \tag{8}$$

Where $\hat{s}(n)$ refers to the isolated signal, $s(n)$ refers to the source signal.

Make calculations for the three above methods by means of ERR formula respectively, the results are shown in Table 1.

Table.1. The ERR comparison of different separation methods

ERR (%)	JADE method	Standard ICA method	Variational bayesian ICA method
Source signal s_1	70.16	76.58	7.86
Source signal s_2	62.83	75.48	5.39

Seen from the Table 1, the standard ICA method has the largest ERR values, which are 76.58% and 75.48% respectively, it indicates that the separation effect is the worst, the next is JADE method, the ERR values of which are 70.16% and 62.83% respectively, its separation effect is also poor. Relative to the JADE method and standard ICA method, the variational Bayesian ICA method has ERR values of 7.86% and 7.86%, which are far less than the former two separation methods, and the separation effect of which has obvious advantages.

Analysis of TM remote sensing image

The imaging time for the TM remote sensing image used in this experiment is 1988, the size of study area is 2 000 × 2 000 pixels. TM remote sensing image has 7 bands, in this experiment we mainly used band 1, 2, 3, 4, 5, 7, etc. The image preprocessing for TM remote sensing images mainly includes geometric correction and resampling. The TM remote sensing image after preprocessing is shown in Figure 4.
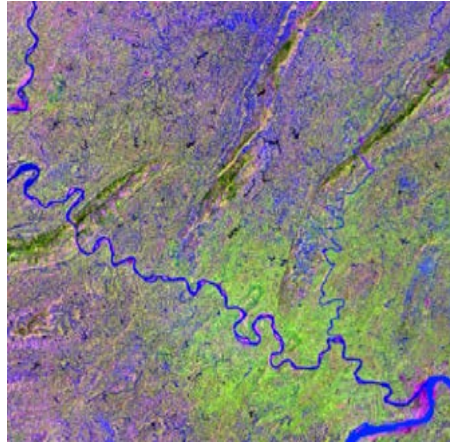
Fig.4. TM remote sensing image

After completing the preprocessing for TM images, we used variational Bayesian ICA method for independent component separation experiment. Figure 5 shows the independent component image obtained by variational Bayesian ICA method (ICI).
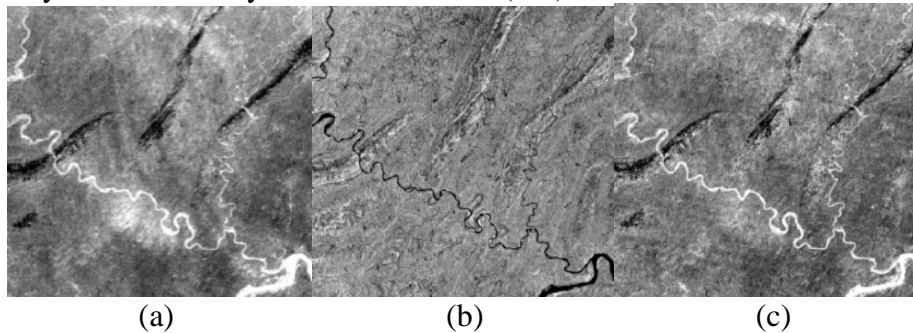

| (a) | (b) | (c) |

Fig.5. The results of TM images obtained by variational Bayesian ICA method, (a) - (c) represents ICI1 - ICI3, respectively

Seen from the Figure 5, the order of the ICI isolated by the variational Bayesian ICA method is random; it does not appear in turn according to the amount of information. As mentioned before, this is mainly due to the random initialization of the projection matrix, but this does not affect its application in practice.

Conclusion

This paper puts forward the variational Bayesian ICA method based on remote sensing images, aiming at the shortcomings of the ICA model and its difference with the characteristics of remote sensing data. We can make effective analysis on the remote sensing image by means of the variational Bayesian ICA algorithm, as well as isolate the different types of terrain. Then we validated the proposed method by simulation and real remote sensing image tests. The results show that the variational Bayesian ICA algorithm has the best stability, maximum separation degree and can accurately isolate different types of terrain information from remote sensing images. The proposed variational Bayesian ICA method is an improvement of the traditional ICA method for remote sensing images; it has certain theoretical and practical significance for application in practice.

Acknowledgement

References

[1] Nascimento, J.M.P. & Dias, J.M.B, Does independent component analysis play a role in unmixing hyperspectral data[J]. IEEE Transactions on Geoscience and Remote Sensing, 2005, 43(1), 175-187.

[2] Li Chengfan, Dai Yangyang, Zhao Junjuan, Yin Jingyuan & Dong Jiangshan, Volcanic ash cloud detection from remote sensing images using principal component analysis[J]. Computers & Electrical Engineering, 2014, 40(8), 204-214.

[3] Hyvarinen, A., Independent component analysis in the presence of Gaussian noise by maximizing joint likelihood[J]. Neurocomputing, 1998, 22(1-3), 49-67.

[4] Li Zhinong, He Kuang & Zhang Fen. Blind source separation method based on principal component analysis and variational Bayesian independent component analysis[J]. Journal of Nanchang University (Natural Science), 2014, 38(2), 187-191.

[5] Li Chengfan, Yin Jingyuan & Zhao Junjuan, Using improved ICA method for hyperspectral data classification[J]. Arabian Journal for Science and Engineering, 2014, 39(1), 181-189.

[6] Li Chengfan & Yin Jingyuan, Variational Bayesian independent component analysis-support vector machine for remote sensing classification[J]. Computers & Electrical Engineering, 2013, 39(3), 717-726.

[7] Li Chengfan & Yin Jingyuan, A multispectral remote sensing data spectral unmixing algorithm based on variational Bayesian ICA[J]. Journal of the Indian Society of Remote Sensing, 2013, 41(2), 259-268.

[8] Li Zhinong, Zhou Wei, Liu Weibin & He Kuang, The blind source separation method based on wavelet packet and variational bayes independent component analysis[J]. Journal of Nanchang Hangkong University (Natural Sciences), 2013, 27(1), 61-65.