Image Blind Signal Separation Algorithm based on Fast ICA

Ajit B.Pande
Electronics and Telecomm. Engg Department,
MGM'S College Of Engg. And Technology,
Mumbai University, India.

Shashi Prabha
Electronics and Telecomm. Engg Department,
MGM'S College Of Engg.And Technology,
Mumbai University, India

ABSTRACT
The independent component analysis (ICA) algorithm and ICA basic model were studied in detail in this paper. Here mainly discussed the mathematical theory of ICA and FastICA algorithm of the most widely used at home and abroad. Then , carried out simulation in blind sources process of the mixing images. Through simulation of three color images (640*480) mixed by a 3*3 random matrix, and using the Fast ICA separation algorithm, realized the mixed blind source separation. The simulation experiment results show that the FastICA algorithm can gain a good approach effect, and the separation image is basically consistent with the original image.

Keywords
Blind source separation, ICA, random matrix, mixing images, Fast ICA.

1. INTRODUCTION
At present, the blind signal separation (BSS) [1] in signal processing field has got more and more attention because in the wireless communication, medical analysis, speech recognition, image processing field, it has broad prospect of application. In the past ten years. The relevant theories and research get the rapid development. BSS problem is to separate independent signal source from linear mixing signals in the premise without prior knowledge based on the statistical independent theory, which can also called the independent component analysis. [2] Since J.H erault and Jutten [3] first carried out research of BSS, and a lot of scholars BSS problem carried on a series of research, at present BSS method has put forward a series of algorithm. These methods of the signal processing have been widely applied on many fields, but in the image blind source separation, the discussed algorithm has little research.[4].

2. INTRODUCTION OF ICA
2.1 ICA mathematics model
Independent Component Analysis (ICA) was originally used to solve the 'cocktail party problem', as in condition of many persons' voices of mutual aliasing, required to let the speech separated alone. ICA is to point to the source signal only using source signals' observation (mixed) signals to restore the each independent component of source signal. Figure 1 expresses independent component analysis problem with the structure diagram.[5]

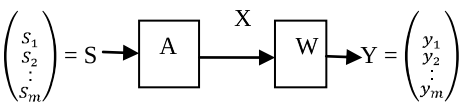

Fig 1.ICA model frame diagram

If we suggest $x(t) = [x(t), x(t),...x_n(t)]^T$ is n dimension random observation mixed signal, now, there is m numbers of source signal $s(t) = [s(t), s(t),...,s_m(t)]^T$, each observation value $x_i(t)$ is a sampling of the random variable, which has general character, a mixture of general stochastic variable and independent sources have zero mean. When we define the ICA model in the matrix form.

$X = (x_1, x_2, ...x_n)^T$ is n random observation vector. $S = (s_1, s_2, ...s_m)^T$ is m dimension unknown source signal, then the ICA linear model can be expressed as formula (1).

$$X = AS = \sum_{j=1}^{m} a_i s_i(t), i = 1,2, \ldots\ldots, m \qquad (1)$$

Among formula (1), $s_i(t)$ is independent component, $A = (a_1, a_2, \ldots a_m)$ is full rank mixed matrix, a_i is base vector matrix of mixed matrix [6]. From formula (1) we can see, each observation data $x_i(t)$ is gotten by different linear weighted of a_{ij} by independent source $s_i(t)$. Independent source $s_i(t)$ is implied variables, and they do not directly measured, mixing matrix $\boldsymbol{A}$ is also unknown matrix, the information that can be adopted only the observation of random vector $\boldsymbol{X}$. Without restriction conditions, only $\boldsymbol{X}$ estimate $\boldsymbol{S}$ and $\boldsymbol{A}$, there are countless equation solution. And in some limited conditions of ICA, according to the statistic characteristics of $\boldsymbol{X}$, it given the only solution, and realize the equation of independent component of the extraction. An important basic assumption of ICA is the requirements of independence character to unknown source signals. According to specific model, unknown source signal is independent, that is to meet formula (2).

$$P(s) = \prod_{i=1}^{N} P_i(S_i) \qquad (2)$$

In ICA model, the source signals need independent, also must satisfy the non-Gaussian distribution characteristics, in addition, in order to simplify the mathematical model, we assume the unknown mixture matrix $\boldsymbol{A}$ is a square formation, which is m=n. So, that is the purpose of the ICA would need to find a transformation matrix, transform $\boldsymbol{X}$ in linear, and get n output vector $\boldsymbol{Y}$.

$$Y = WX = WAS \qquad (3)$$

When allow the premise of un-qualitative in order and proportion, $\boldsymbol{Y}$ is an estimate of independent component s_i . From above BSS view, we can expound the ICA model. The following are the ICA model in multi-dimension signal linear description view. Set n dimension observation data

$X = (x_1, x_2, ...x_n)^T$, the purpose of the ICA is looking for a coordinate system $\{\xi_1, \xi_2, \ldots \xi_n\}$, which making each component $x_1, x_2, ...x_n$ in the coordinate projection is formula (4).

$$x_i = s_1\xi_1 + s_2\xi_2 + \ldots + s_n\xi_n, i = 1,2, \ldots, n \quad (4)$$

The projection coefficient $s_1, s_2, ...s_n$ is independent. If make $\boldsymbol{Y}$=$\boldsymbol{WX}$, in ICA algorithm, the goal is to find a optimal matrix $\boldsymbol{W}$, which making output y_i statistical independence, namely each other mutual information of $\boldsymbol{Y}$ for zero. At this time, $W^{-1} = [\xi_1, \xi_2, \ldots \xi_n]$ is the ICA coordinate system of the linear description model.

2.2 ICA criterion

Multi-independent random variables and the probability density distribution obey Gaussian distribution asymptotically according to center limit theorem. So, we can act the non-Gaussian character of each component as the criterion of independence. Each component closes to the independent more, the non-Gaussian character is stronger. The key of independent component analysis model estimate is the non-Gaussian character, and the size of non-Gaussian usually use negative entropy and kurtosis to measure.

By the theory of information theory, entropy value is related to the information of the observation data, in all having the random variable variance, the Gaussian character is stronger, and the Gaussian distribution information entropy is smaller. Usually, this means that using entropy can measure the Gaussian character. Negative entropy is kind of differential entropy, it is the amount of information theory in normalizing difference entropy, and the definition of negative entropy is shown as follows formula (5). [7]

$$J(X) = H\left(X_{guass}\right) - H(X) \qquad (5)$$

Among them, $H(X) = \int f(X) \log f(x) dx$. x_{guass} is the Gaussian random variables, which have the same covariance with *x*. It remains the same to *x* any linear transformation. This is an important characteristic of negative entropy. Usually, negative entropy is always non-negative and just zero when *x* is Gaussian distribution. Usually, in order to simplify the calculation in real application, the negative entropy approximate taken value as formula (6).

$$J(X) \propto [\, E\{G_i(X)\} - E\{v\}]^2 \qquad (6)$$

Among them, *v* is a Gaussian random vector with zero mean and unit variance, the mean of *x* is zero, and the variance is unit variance. Among them, *G* (●) takes the quadratic function such as formula(7).

$$g_1(u) = \frac{1}{a_1} \log cos a_1 u \ \ (1 \le a_1 \le 2)$$
$$g_2(u) = u \exp(-a_2 u^2/2) \ \ (a_2 \approx 1)$$
$$g_3(u) = u^3 \qquad (7)$$

In the Gaussian character measure, negative entropy can be got. They are the good compromise between negative entropy and classical kurtosis. The approximate characteristic is calculations quickly, concept simple and good robustness. [8]

3. Fast ICA ALGORITHM

Fast ICA algorithm is based on the maximum principle of non-Gaussian character, uses fixed- point iterative theory to look for non-Gussian character maximum of $W^{T}x$, this algorithm adopts Newton iterative algorithm, and carries out batch to amount of sampling points of observed variables *x*, isolates a independent component from observation signal every times. The Gaussian character measure function of this algorithm is shown as formula (6). In order to reduce the estimate parameters of the algorithm, and simplify the calculation of algorithm, before running Fast ICA algorithm, we need carry out data pretreatment, that is removing mean value and whitening process.

Fast ICA algorithm acts as one of the most popular algorithm of independent component analysis algorithm, it can start from observation signal and estimate source signals with little known information, and get the approximation of original signal independent each other. Fast ICA algorithm is a kind of fixed point iteration method, which in order to find out the maximum of Gaussian character, and use formula $J(y) = [E\{G_i(y) - E\{G(v)\}]^2$to measure its independence, and can also approximately derive it by Newton iterative method.

First, we notice that the maximum value of approximation negative entropy of $w^{T}x$ is gotten through $E\{G(w^{T} x)\}$ optimized. According to the conditions of Kuhn-Tucker, the most optimized point of $E\{G(w^{T} x)\}$ is obtained in meeting formula (8) under the restriction of
$E\{\, G(w^T X^2)\} = ||w||^2 = 1$.

$$\beta = E\{\, w_o^T X g(w_o^T X)\} \qquad (8)$$

In formula (8), β is a constant value and can easy to get through . And w_0 is the initial boundary value of *w* . If we assume we had solve the equations by the Newton method, and express the left of function formula (8) by *F*, then we can get the Jacobian matrix *JF*(*w*) as bellow formula (9).

$$JF(w) = E\{\, xx^T g'(w^T x\,)\} - \beta I \qquad (9)$$

In order to simplify the calculation of transposed matrix, we take the approximation value of first item and the reasonable estimate is shown as formula (10).

$$E\{xx^T g'(w^T x)\} \approx E\{xx^T\} E\{g(w^T x)\}$$
$$= E\{g'(w^T x)\} I \qquad (10)$$

Jacobian matrix is the Lord diagonal matrix and singular, and easy to deferring relatively, the approximate iteration formula is shown as formula (11).

$$w_{k+1} = w_k - \frac{E\{xg(w_k^T x)\} - \beta w_k}{E\{g'(w_k^T x)\} - \beta} \qquad (11)$$

Among formula (11), $\beta = E\{xg(w_K^T x)\}$

In order to enhance the stability of the iterative algorithm, after $w_{k+1} = w_{k+1} \,/\, ||w_{k+1}||$ iteration, we use formula to normalization *w*, and in formula (11) multiplied $\beta - E\{\, g'(w_K^T x)\}$ on both sides, then, the algorithm will be further simplified, which will get protection iteration algorithm formula is shown as formula (12).

$$w_{k+1} = E\{xg(w_K^T x)\} - E\{xg'(w_K^T x)\} w_k \qquad (12)$$

If not convergence, then this process will be repeated, until convergence, so far, we can estimate one independent component. If there are *n* numbers of source signal, we should estimate the *n* numbers of independent component. Every times a component extracted, the independent component will be subtracted form the observation signal, repeat until all the components are extracted from observation signal. The method of removing extracted independent component is shown as following formula (if suggest we have estimated the *k* numbers component).

$$w_{k+1} = w_{k+1} - \sum_{j=1}^{k} {w_{k+1}}^T w^j w^j$$

$$w_{k+1} = w_{k+1} \,/\, \sqrt{w_{k+1}^T w_{k+1}} \qquad (13)$$

Among them, w_{k+1} is the *k+1* times Newton Iterated results.

4. IMAGE BSS APPLICATION BASED ON Fast ICA

4.1 Algorithm analysis

The realization of Fast ICA algorithm includes three points, the first, removing mean value of the observation signal, the second, whitening processing to the observation signal after mean processing, the third, the process of independent components extraction algorithm and realization. Algorithm realization flow is shown as Figure 2. The first two steps can be regarded as the pretreatment to the observation signal. Through mean and whitening processing, ICA algorithm process can be simplified.[9]

The specific procedure is shown as follows:

(1) Inputing mixing signal X;

(2) Meaning the obversation data X and making the mean value is 0;

(3) Whitening the data and getting unit variance signal Z;

(4) Choosing the need to assess the number of components, set Iteration times p←1;

(5) Randomly selecting chosen initialized weights vector w_o and k = 0;

(6)Using formula (11) to update weights vector w_{k+1}

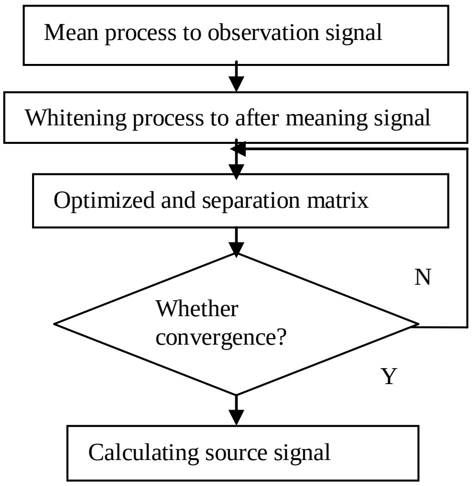

Fig 2. Flow chart

(7) Normalized w_{K+1} and

$$w_{k+1} = {w_{k+1}}/{||w_{k+1}||} \ ;$$

(8) If $|w_{k+1} - w_k| > \varepsilon$, then the algorithm is not convergence, return to step (2), or Fast ICA algorithm estimate a independent component, and the algorithm is over. For many numbers independent component extraction, we can repeat to use above basic form of FastICA algorithm. To be sure each time's different components extracted, we only need to remove independent component having be extracted from the observation signal in every times, and repeat this process until all the dependent components needed to be extracted. We can use the formula (13) to realize removing extracted independent component.

4.2 Algorithm Simulation

If we set *m*=3, *s*(*t*) has three images of signal sources, it is shown as Fig3. Here, we select three images of 640*480 in Silver Hole Gorge of Baoji, China. And we suggest mixing matrix A is 3×3 full rank random matrix, mixed images of *x*(*t*) are shown as Fig 4.

Fig 3 Source Images.

Fig.4 Mixed Images

From the Fig 4 we can see, three images are hard to identify their clear initial colony through mixing by random matrix. The results by using Fast ICA separation algorithm to separate source images are shown as Fig 5.

Fig .5 Separated Images.

Separation images after using the Fast ICA BSS algorithm have better effect.In this separation experiment, images order are the same range, and images have a slightly increase. Through the simulation we can see, this paper writing the Fast ICA algorithm program can gain a good approach effect, separate the original image well.

5. CONCLUSIONS

This paper detailed study Fast ICA algorithm, and through the simulation, successfully realize three images mixed effectively BSS. But in nature, there are many polluted images with unknown noise, which is hard to deal with. So, this study is just in the basic model of ICA, the next step of work is to realize images containing noises or other signals' better separation and process.

6. REFERENCES

[1] Guan weihua, Li lifu, Lin yongman, "Underdetermined Blind Separation Algorithm Base d on Subtractive Clustering", AISS: Advances in Information Sciences and Service Sciences, Vol. 3, No. 7, pp. 75- 81, 2011.

[2] Chengfan LI, Jingyuan YIN, Junjuan ZHAO, Feiyue YE, "Detection and Compensation of Shadows based on ICA Algorithm in Remote No. 7, pp. 46-54, 2011. Sensing Image", IJACT : International Journal of Advancements in Computing Technology, Vol. 3

[3] Wang Lei, "Research of BSS Method Based on ICA", Lanzhou university of technology, 2010.

[4] Bi Yang, "Research and Application of BSS Algorithm Based on Fast ICA", Xian university of science and technology, 2007.

[5] Huang Liyan, Gao Qiang, Kang Haiyan, Zhao Zhenbing, Xu Yixi, "Improved Fast ICA algorithm", Journal of North China Electric Power Univenity, Vol. 5, pp.59-60, 2006.

[6] Yang F. S., Hong B., " Theory and application of independent component analysis", Beijing: tsinghua university press, (2006).

[7] Cao Huirong, Zhang Baolei, Ma Li, "The Separation of Mixed Images Based on Fast ICA Algorithm", Computer Study, Vol.1, pp. 45, 2004.

[8] Hyvarinen A., et al. "Independent component analysis", John Wiley and Sons, (2001).

[9] ARIE Y, "Blind sour ce separation via the second characteristic function", Signal Processing, Vol. 80, pp. 897-902, 2000.

[10] Hyvarinen A., Oja E., " Independent component analysis: algorithm and application", Neural Network, 13(4-5), 411-430, (2000).

[11] Comon P, "Independent Component Analysis a New Concept", Signal Processing, Vol. 36, pp. 287-314, 1994.

[12] Jutten C, Herault J, "Space or time adaptive signal processing by neural network models", Intern. Conf . on Neural Network for Computing, Snowbird (Utah ,USA), pp. 206-211, 1986.