http://ansinet.com/itj

ITJ

ISSN 1812-5638

# INFORMATION TECHNOLOGY JOURNAL

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Information Technology Journal 11 (4): 456-460, 2012
ISSN 1812-5638 / DOI: 10.3923/itj.2012.456.460


# Research and Simulation on Blind Separation Algorithm in Mixing Image Based on ICA

Xinling Wen and Manxiang Miao
Zhengzhou Institute of Aeronautical Industry Management,
450015 Zhengzhou, China

**Abstract:** This paper studied the Independent Component Analysis (ICA) algorithm, mainly introduces the ICA model and design an algorithm based on the Fast ICA method, which is the most widely used in blind source separation at present. A kind of Fast ICA algorithm is researched in this paper and through simulation of three mixed images, we prove this algorithm's accuracy in separating blind sources signal, as well as obtain precise result in image blind separation. From the simulation diagram results we can see, that Fast ICA algorithm used in images blind separation can get better separation effect.

**Key words:** Blind signal separation, independent component analysis, performance index, matlab, fast ICA

## INTRODUCTION

In the field of image processing, the various causes lead to images' aliasing such as image acquisition, transmission process, etc. In order to obtain the original image, we often need to separate these mixed image each other. Independent Component Analysis (ICA) had developed a kind of new signal processing technology in 20 century 90's. In recent years, ICA theory and application has gotten domestic and foreign scholars' concern. ICA theory and application has won the considerable development, which has emerged a variety of different algorithms (Hyvarinen, 1999).

ICA is the meaning of having no prior knowledge, according to statistics way observed signals the principle of independent through optimization algorithm is divided into several independent component (Comon, 1994). ICA algorithm has been applicated in communication, image feature extraction, biological signal analysis and other areas widely (Cuiru *et al*., 2006; Yuen and Lai, 2001; Guo and Sun, 2004). Fixed point ICA algorithm (Amari *et al*., 1997) has faster convergence speed, separation effect is more reliable, so, in BSS field, it has widely application. In this study, we present an improved Fast ICA algorithm and applied to image blind separation fields, from the simulation we can see, the fast ICA algorithm in image BSS has obtained better results.

## INTRODUCTION OF ICA

**ICA mathematics model:** ICA aims to carry out blind separation to unknown mixture of observation signals that produced by independent source signal, thus reproduce the original independent source (Yang and Hong, 2006) and its application mainly includes two aspects, one is the blind source separation and the other is feature extraction. Figure 1 expresses independent component analysis problem with the structure diagram (Yang, 2007).

If $x(t) = [x_1(t), x_2(t),..., x_n(t)]^T$ is n dimension random observation mixed signal, there is m numbers of source signal $s(t) = [s_1(t), s_2(t),..., s_m(t)]^T$, each observation signal value of $x_i(t)$ is a sampling of the random variable, which has general character, a mixture of general stochastic variable and independent sources have zero mean. When defining the ICA model in 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 Eq. 1:

$$X = AS = \sum_{j=1}^{m} a_i s_i(t),\ i =,12,...,m \tag{1}$$

where, $s_i(t)$ is independent component, $A = a_1, a_2,..., a_m)^T$ is m×n full rank mixed matrix, $a_i$ is base vector matrix of mixed matrix. From Eq. 1, each observation data $x_i(t)$ is gotten by different linear weighted of $a_{ij}$ by independent

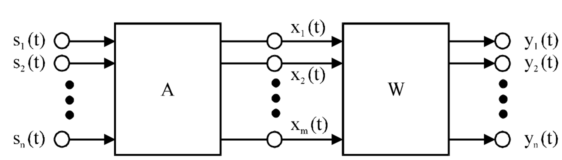


Fig. 1: ICA model frame diagram

**Corresponding Author:** Xinling Wen, Zhengzhou Institute of Aeronautical Industry Management, 450015, China

source $s_i$ (t). Independent source $s_i$ (t) is implied variables, mixing matrix A is also unknown matrix, the information that can be adopted only the observation of random vector X. Without restriction conditions, only X estimates S and A, there are countless equation solution. 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 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 X in linear and get n output vector Y.

$$Y = WX = WAS \tag{2}$$

**The basic hypothesis:** Because in the ICA model, the prior knowledge of source signal and mixed matrix is unknown, only the information of observation signal can be exploited. If without any premise condition, blind separation problems will have a lot of solution, so, it needs to add some basic hypothesis and constraint conditions to source signal and mixed matrix.

- m = n, for convenience, take m = n, that is to say that A is a mixed matrix of full rank
- Each components of source signal S is statistically independent each other
- Each component of the source signal at most only allow one gaussian distribution, this is due to many numbers of gauss signals linear-mixing signal still obey gaussian distribution, which can not to be separated

**Pretreatment:** Fast ICA pretreatment includes two parts of mean and albino processing. In most of the proposed BSS algorithm, we assume that source signal of each component is zero mean and random vector. So, in order to make the actual blind source separation to the proposed mathematical model, we must carry out the mean processing before the signal separation. In order to make X to become a mean for zero of variables, we can put each random variable X minus the mean. From observations signal X minus the mean vector of signal, which making observation signals to become zero mean variable, so, we can simplify the ICA algorithm. If we have N numbers of image blind mixed, the part of standardization program is shown as bellow:

```
    MixedS_mean = zeros (N,1);
  for i=1:N
    MixedS_mean (i) = mean (MixedS (i,:));
end
for i=1:N
    for j = 1: Size (MixedS,2)
    MixedS (I, j) = MixedS (i,j)-MixedS_mean (i);
  end
end
```

Albino as separation algorithm is a common preprocessing method, in great degree, albino process can reduce the complexity of the problem. The so-called albino is carry out linear transformation to suppressive data, make transformation data is not related and is unit variance. If it is albino, the covariance matrix after random vector after suppressive must be a unit matrix. Albino is a commonly used methods of ICA in the pretreatment. It presents each component is not correlation. Part of albino program is shown as bellow:

```
MixedS_cov = cov (MixedS');
(E, D) = eig (MixedS_cov);
Q = inv (sqrt (D))* (E)';
MixedS_white = Q*MixedS;
IsI = cov (MixedS_white');
```

**Uncertainty of the ICA model:** The above definition ICA model has two uncertainty: one is the range, the separated signal amplitude and the source in the signal has a certain relation; Second, in order, separated signal and the order of the source signal may not consistent. But these uncertainty did not affect the algorithm to solve practical problems.

**Fast ICA algorithm:** Fast ICA algorithm (Hyvarinen *et al*., 2001) 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 Eq. 4. 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 bleaching process.

**ICA criterion:** 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 Eq. 3 (Tian *et al*., 2009).

$$J(x) = H(x_{gauss}) - H(x) \tag{3}$$

Among them, $J(x) = \int f(x) \log f(x)\, dx$, $x_{gauss}$ 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 Eq. 4.

$$J(x) \propto [E\{G_i(x)\} - E\{v\}]^2 \qquad (4)$$

Fast ICA algorithm is essentially a minimize weight estimates of the mutual information neural network method, using the maximum entropy principle to approximate negative entropy and through a suitable nonlinear function g achieve the optimal (Hyvarinen and Karhunen, 2007). G (·) takes the quadratic function such as:

$$g_1(u) = \frac{1}{a_1}\log\cos a_1 u$$

$1 \leq a_1 \leq 2$, $g_2(u) = u\exp(-a_2 u^2/2)$, $a_2 \approx 1$ and $g_3(u)\ u^3$, etc. 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. In algorithm simulation, we select $g_3(u)\ u^3$.

**Algorithm workflow**

**Algorithm analysis:** The specific procedure is shown as follows:

**Step 1:** Randomly selecting chosen initialized weights vector $w_0$ and $k = 0$

**Step 2:** Using equation to update weights vector $w_{k+1}$

**Step 3:** Normalized $w_{k+1}$ and $w_{k+1}/||w_{k+1}||$

**Step 4:** 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 Fast ICA 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 independent components needed to be extracted. We can use the formula to realize removing extracted independent component.

**Algorithm simulation:** If we set 3 pictures of image, it is that s (t) has three images of signal sources, it is shown as Fig. 2.

Here, we select three images of 800*600 penoy in Luoyang flower show. And we suggest mixing matrix A is

Fig. 2: Source images

a 3×3 full rank random matrix mixed images of x (t), which are shown as Fig. 3.

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. 4. From Fig. 4 we can see, separation images after using the fast ICA BSS algorithm have better effect. Because

Fig. 3: Mixing images

Fig. 4: Separation images

separation result is uncertainty, so, the order of separation image are different from the source image.

## CONCLUSIONS

This paper detailed study Fast ICA algorithm and through the simulation, successfully realize three images mixed effectively BSS. This algorithm in image processing area has broad prospect of application. In future, we will research and simulation mixed images with noise, which is the development direction.

## ACKNOWLEDGMENTS

This paper is supported by the Aeronautical Science Foundation in China (No.2010ZD55006).

## REFERENCES

Amari, S., T. Chen and A. Cichocki, 1997. Non-Holonomic Constraints in Learning Blind Source Separation. In: Progress in Connectionist-Based Information Systems, Kasabov, N. (Eds.)., Springer, New Zealand, pp: 633-636.

Comon, P., 1994. Independent component analysis a new concept. Signal. Process., 36: 287-314.

Cuiru, W., D. Juan and Z. Xinlin, 2006. Research of image separation based on improved independent component analysis. Chin. J. Sci. Instrum., 27: 785-786.

Guo, S. and Y.L. Sun, 2004. Fast independent component analysis algorithm based on blind source separation. Electron. Meas. Technol., 27: 12-14.

Hyvarinen, A., 1999. Survey on independent component analysis. Neural Comput. Surveys, 2: 94-128.

Hyvarinen, A., J. Karhunen and E. Oja, 2001. Independent Component Analysis. Wiley Interscience, UK.

Hyvarinen, A. and J. Karhunen, 2007. ICA, Zhou Zongtan. Electronic Industry Press, Beijing, China, pp: 191-192.

Tian, Q., W. Zheng and D. Sun, 2009. Study on speech separation and image separation based on fast ICA. Program. Skills Maintenance, 16: 90-91.

Yang, B., 2007. Research and application of BBS algorithm based on fast ICA. Xian University of Science and Technology, China.

Yang, F. and B. Hong, 2006. The Principle and Application of ICA. Tsinghua University Press, Beijing, China.

Yuen, P.C. and J.H. Lai, 2001. Face representation using independent component analysis. Pattern Recognit., 34: 545-553.