

Blind Source Separation Using Modified Contrast Function in Fast ICA Algorithm

Dr. Alka Mahajan

SIES Graduate School of Technology
Nerul, Navi Mumbai
India – 400 706

Gajanan Birajdar

SIES Graduate School of Technology
Nerul, Navi Mumbai
India – 400 706

ABSTRACT

A novel contrast function is proposed to be used in fastICA algorithm for Blind Source Separation (BSS). Simulation results show that the proposed nonlinear function used to separate image mixtures, results in faster execution and good quality image separation. Peak Signal to Noise Ratio (PSNR), Improved Signal to Noise Ratio (ISNR), Signal to Noise Ratio (SNR) and Root Mean Square Error (RMSE) are used to evaluate quality of separated images and Amari error is calculated to prove the performance of separation quality.

General Terms

Image Processing, Algorithms.

Keywords

Blind source separation, Independent component analysis, FastICA Algorithm, Nonlinearity function.

1. INTRODUCTION

Blind source separation (BSS) is the method of extracting underlying source signals from a set of observed signal mixtures with little or no information as to the nature of these source signals. Independent component analysis (ICA) is used for finding factors or components from multivariate statistical data and is one of the many solutions to the BSS problem [1]-[3]. ICA looks for the components that are both statistically independent and nongaussian. The various ICA algorithms extract source signals based on the principle of information maximization, mutual information minimization, maximum likelihood estimation and maximizing nongaussianity. ICA is widely used in statistical signal processing, medical image processing, economic analysis and telecommunication applications [4]-[5].

Many popular ICA methods use a nonlinear contrast function to blindly separate the signals. Examples include equivariant adaptive source separation [10], fastICA [1], and efficient FastICA [11]. Adaptive choices of the contrast functions have also been proposed, in which the probability distributions are obtained by considering a maximum likelihood (ML) solution corresponding to some given distributions of the sources and relaxing this assumption afterward [12]-[13]. This method is specially adapted to temporally independent non-Gaussian sources and is based on the use of nonlinear separating functions. Further Tichavsky et al. [9] have proposed two general purpose rational nonlinearities that have similar performance as \tanh , but can be evaluated faster.

In this paper a different nonlinear function is proposed for use in BSS using the fastICA technique. Natural images [7] are used to validate the performance of proposed nonlinear function.

2. PROPOSED NONLINEAR FUNCTION

2.1 Blind Source Separation and ICA

The BSS problem can be stated as,

$$x(t) = As(t) + d(t) \quad (1)$$

where $x(t)$ is N dimensional vector of observed signals at the discrete time instant t , A is an unknown mixing matrix, $s(t)$ is original source signal of $M \times N$ ($M \leq N$) and $d(t)$ is the observed noise vector. The purpose of ICA is to estimate $s(t)$ which is the original source signal, from $x(t)$ which is the mixed signal, i. e. it is equivalent to estimating matrix A . Assuming that there is a matrix W , which is the demixing matrix or separation inverse matrix of A , then the original source signal is obtained by

$$s(t) = Wx(t) \quad (2)$$

ICA algorithm assumes [6] that the mixing matrix A must be of full column rank and all the independent components $s(t)$, with the possible exception of one component, must be non-Gaussian. Further, the number of observed linear mixtures m must be at least as large as the number of independent components n ($m \geq n$).

2.2 Proposed Nonlinearity

Practically any nonquadratic function Γ which is smooth and robust against outliers, may be used to perform independent component analysis. The function Γ divides the space of probability distributions into two half spaces, depending on whether the nonpolynomial moment in the theorem is negative or positive. While choosing the nonlinearity functions the important considerations are [4]:

1. The practical estimation $E G(x)$ should be easy to calculate, and should not be sensitive to outliers.
2. The function $G(x)$ should not grow too fast (not faster than quadratically) as $|x|$ increases.
3. The G must capture aspects of the distribution of x that are pertinent in the computation of entropy.

It is desired that the estimator is robust against outliers and this robustness depends on the behavior of the contrast function $g(u)$. Also the asymptotic variance of the estimator can be minimized if the primitive function is optimized in the form

$$G_{opt}(u) = B_1 \log f(u) + B_2 u^2 + B_3 \quad (3)$$

where B_1 , B_2 and B_3 are arbitrary constants, f is the density function of $s(i)$ and $G_{opt}(u)$ is the optimal contrast function. Keeping in line with the above, the proposed nonlinearity and its primitive function is given as,

$$g(u) = \frac{u}{(1 + \frac{u^2}{2})} \quad (4)$$

$$G(u) = \log(1 + \frac{u^2}{2}) \quad (5)$$

where $g(u)$ is the derivative of the primitive function $G(u)$. The proposed function is compared with the commonly used *tanh* and *Gaussian* function and the comparative behavior is shown in figure 1.

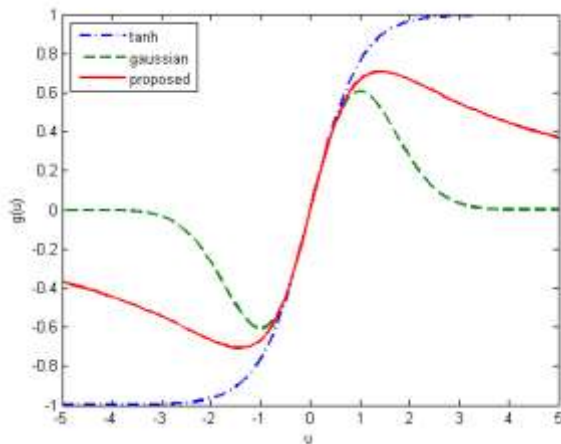


Figure 1: Comparative behavior of *tanh*, *Gaussian* and *proposed nonlinearity*

The proposed nonlinear function has a very similar behavior as *tanh* and *Gaussian* in the neighborhood of zero. As compared to the *tanh* nonlinearity, the derivative of nonlinear function g , i.e. g' of the proposed nonlinearity has a wider spread. Using Matlab it was found that, the evaluation of *tanh* takes 0.12 seconds whereas the proposed nonlinearity requires 0.05 seconds. Thus showing that, the proposed function is also executed faster.

It is known that robustness of an estimator depends on the behavior of the function $g(u)$. The slower $g(u)$ grows, the more robust the estimator i.e. single highly erroneous observation will not influence the estimator. Hence as is the case with the proposed nonlinearity, the contrast function is chosen such that it does not grow very fast when $|u|$ grows.

When the independent components are estimated one-by one, the order in which they appear is influenced by the choice of contrast function [15]. Practically it is difficult to determine certainty of this order, but a suitable choice of the contrast function means that independent components with certain distributions tend to be found first.

3. FASTICA ALGORITHM WITH PROPOSED NONLINEARITY

FastICA is the most popular algorithm used in various applications as it is simple, fast convergent and computationally less complex. It is a fixed point iteration scheme for finding a maximum of nongaussianity or negentropy of ' $w^T x$ ' i.e. one of the independent component ($S_i = w_i^T x$). FastICA learning rule finds the direction of a unit vector ' w ' such that the projection of $w^T x$ maximizes nongaussianity. Nongaussianity is measured by negentropy $J(w^T x)$ i.e.

$$J(y) = [E G(y) - E G(v)]^2 \quad (6)$$

where ϖ is Gaussian variable of zero mean and unit variance and Γ is some non-quadratic function. The variable ψ is assumed to be zero mean and unit variance. The fastICA algorithm was applied to a set of natural images.

Data pertaining to the selected natural images was first subjected to preprocessing before applying to ICA algorithm. Preprocessing is necessary to simplify and reduce the complexity of the problem. The steps applied in the fastICA are as listed below:

1. The data x is centered by subtracting its mean vector $m = E x$ so as to make x a zero mean variable.

2. The observed data x is transformed linearly so that a components of a new vector \tilde{x} are uncorrelated and their variance equals unity i.e. $E \begin{Bmatrix} \tilde{x} \\ \tilde{x}^T \end{Bmatrix} = I$. This whitening is done by

using eigenvalue decomposition (EVD) of the covariance matrix. The new mixing matrix obtained using whitening is given by,

$$\tilde{x} = ED^{-1/2} E^T x = ED^{-1/2} E^T A s = \tilde{A} x \quad (7)$$

3. An initial random mixing matrix w is chosen.

4. Let,

$$w^+ = E xg(w^T x) - E g'(w^T x) w \quad (8)$$

where g is the nonlinearity function.

5. The matrix w is then orthogonalized.

6. Let,

$$w = \frac{w^+}{\|w^+\|} \quad (9)$$

7. Step 4 is repeated till convergence is obtained. Convergence means that the old and new values of w point in the same direction i.e. their dot product equal to 1. To estimate several independent components, we need to run one unit FastICA algorithm using several units with weight vectors w_1, \dots, w_n . When we have estimated p independent components, or p vectors w_1, \dots, w_p we run the one unit fixed point algorithm for w_{p+1} , and after every iteration step subtract from w_{p+1} the projections $w_{p+1}^T w_j w_j$, $j = 1, \dots, p$ of the previously estimated p vectors, and then renormalize w_{p+1}

Let

$$w_{p+1} = w_{p+1} - \sum_{j=1}^p w_{p+1}^T w_j w_j \quad (10)$$

$$w_{p+1} = w_{p+1} / \sqrt{w_{p+1}^T w_{p+1}}$$

8. The various performance factors are then calculated using the following equations.

(i) Root mean square error (RMSE):

$$RMSE = \sqrt{\frac{1}{MN} \sum_{i=1}^N \sum_{j=1}^M (I_{ij} - \tilde{I}_{ij})^2} \quad (11)$$

where $N \times M$ = Image size, I_{ij} = original source image, \tilde{I}_{ij} =separated image

(ii) Peak signal to noise ratio (PSNR):

$$PSNR = 20 \log_{10} \left(\frac{255}{RMSE} \right) \quad (12)$$

(iii) Signal to noise ratio (SNR):

$$SNR = 10 \log_{10} \frac{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y)^2}{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [f(x, y) - \tilde{f}(x, y)]^2} \quad (13)$$

where $f(x,y)$ = original image and $\tilde{f}(x, y)$ =separated image

(iv) Improved signal to noise ratio (ISNR):

$$ISNR = 10 \log_{10} \left[\frac{s(m, n) - x(m, n)}{s(m, n) - y(m, n)} \right] \quad (14)$$

where $s(m,n)$ =source image, $x(m,n)$ =mixed image, $y(m, n)$ = separated image

(v) Amari error is then used for evaluating separation quality of unmixing matrix w and to compare various ICA algorithms [8]. Before calculating Amari error, normalization is carried out on the mixing matrix and the estimated separation matrix. This normalization is done in the range of [0, 1] where 0 corresponds to perfect separation. It is given by the equation,

$$d(W, A) = \frac{1}{2m(m-1)} \times \left\{ \sum_{i=1}^m \left[\frac{\sum_{j=1}^m |a_{ij}|}{\max_j |a_{ij}|} - 1 \right] + \sum_{j=1}^m \left[\frac{\sum_{i=1}^m |a_{ij}|}{\max_i |a_{ij}|} - 1 \right] \right\} \quad (15)$$

where $a_{ij} = (W^T A)_{ij}$, and W and A are $n \times m$ matrices.

The proposed nonlinear function was used to perform ICA, which resulted in faster convergence and minimum distortion in the separated source images.

4. SIMULATION RESULTS

FastICA algorithm was applied to three natural images [7] Lenna, plane and pool of size 128×128 which are shown in Figure 2. These images were mixed using 3×3 random mixing matrix. The mixed images so obtained are shown in Figure 3

which were then separated by applying fastICA algorithm with the proposed nonlinearity. Simulations were performed using Matlab[®] R 7.9 on a Pentium 4 1.6 GHz PC. Figure 4 shows the separated images using the proposed nonlinearity function. Image separation quality was obtained by evaluating RMSE, PSNR, SNR and ISNR

FastICA algorithm was also applied using *tanh* and *Gaussian* nonlinearity function. Figure 5 and 6 shows the simulation results for the same. Output images are scaled between [0, 255].

Performance parameters obtained using the *tanh*, *Gaussian* and the *proposed* nonlinear function are tabulated in table I - III. Similarly, Amari error and execution time was calculated and tabulated in table IV.

It is observed from Table I - III that, the blind source separation using the proposed nonlinearity function results in better source separation with minimum distortion in terms of SNR, PSNR, RMSE and ISNR as compared to *tanh* and *Gaussian* nonlinear function. Table IV shows better separation quality in terms of Amari error and also confirms faster convergence.



Figure 2: Original source images



Figure 3: Mixed images



Figure 4: Extracted images using proposed nonlinearity



Figure 5: Extracted images using tanh nonlinearity

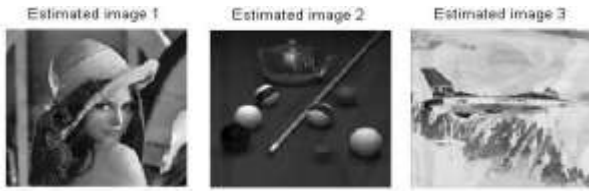


Figure 6: Extracted images using Gaussian nonlinearity

Table 1. Performance parameters for image 1 - Lenna

| Parameter | tanh | Gaussian | Proposed |
|-----------|---------|----------|----------|
| RMSE | 14.9520 | 26.5953 | 15.0824 |
| PSNR | 24.6339 | 19.6347 | 24.5614 |
| SNR | 16.8265 | 11.8272 | 16.7539 |
| ISNR | 26.4679 | 21.4687 | 26.3954 |

Table 2. Performance parameters for image 2 - Plane

| Parameter | tanh | Gaussian | Proposed |
|-----------|---------|----------|----------|
| RMSE | 5.6952 | 8.8210 | 6.1717 |
| PSNR | 33.0207 | 29.2205 | 32.3226 |
| SNR | 30.1147 | 26.3145 | 29.4166 |
| ISNR | 36.9301 | 33.1299 | 36.2321 |

Table 3. Performance parameters for image 3 – Pool

| Parameter | tanh | Gaussian | Proposed |
|-----------|---------|----------|----------|
| RMSE | 2.7415 | 5.6172 | 2.6815 |
| PSNR | 39.3710 | 33.1405 | 39.5632 |
| SNR | 26.2770 | 20.0464 | 26.4692 |
| ISNR | 39.7768 | 33.5463 | 39.9691 |

Table 4. Separation quality and execution speed using different nonlinear functions

| Nonlinearity | Amari error | Execution speed (seconds) |
|--------------|-------------|---------------------------|
| Tanh | 0.4689 | 3.109946 |
| Gaussian | 0.3644 | 2.956239 |
| Proposed | 0.3343 | 2.457466 |

5. CONCLUSION

Linear ICA techniques are primarily based on minimizing nonlinear contrast function and most of them use a hyperbolic tangent (*tanh*) as their built in nonlinearity. In this paper, a rational function is proposed to replace the *tanh* that is used for separating supergaussian sources. The advantage of the proposed rational function is many-fold. First, the rational function requires significantly lower computational complexity than *tanh* resulting in faster execution. Further, it is observed that PSNR, ISNR and SNR are improved resulting in better quality image

separation. Fast convergence is achieved with minimum distortion in the separated signals as evident from the Amari error. It has been observed that the order in which the independent components are separated cannot be determined using ICA. This limitation is partially overcome with the proposed contrast function as the probability of identifying supergaussian distributions first, is found to be maximum. It can be shown that the proposed rational function allows to achieve a better performance of blind source separation.

6. REFERENCES

- [1] A. Hyvärinen and Erkki Oja, 1997, "A fast fixed-point algorithm for independent component analysis", *Neural Computation*, 9(7):1483-1492.
- [2] Bell, A. J. and Sejnowski, T. J., , 1995, "An information-maximization approach to blind separation and blind deconvolution", *Neural Computation*, 7(6):1129-59.
- [3] J. F. Cardoso, and A. Souloumiac, , 1993, "Blind beamforming for non-Gaussian signals", *IEEE Proceeding Part F*, Vol.140, No. 6: 362-370.
- [4] A. Hyvärinen, J. Karhunen, E. Oja, , 2001, "Independent Component Analysis", *John Wiley & Sons* .
- [5] Aapo Hyvärinen and Erkki Oja, 2000."Independent Component Analysis: Algorithms and Applications", *Neural Networks*, 13(4-5):411-430.
- [6] A. Hyvärinen, 1999, "Survey on Independent Component Analysis", *Neural Computing Surveys* 2, 94-128.
- [7] A. Cichocki, S. Amari, and K. Siwek, 2002, "ICALAB toolbox for image processing – benchmarks",
- [8] Cichocki, A., Amari, S.-I., 1996, "A new learning algorithm for blind signal separation", *Advances in neural information processing* 8, 757-763.
- [9] P. Tichavsky, Z. Koldovsky and E. Oja, Sept. 2007, "Speed and Accuracy Enhancement of Linear ICA Techniques Using Rational Nonlinear Functions", *Proceedings of 7th International Conference on Independent Component Analysis (ICA2007)*, pp. 285-292.
- [10] Cardoso J.F., Laheld B., ,1996, "Equivariant adaptive source separation" *IEEE transaction Signal Processing*, 45, pp. 434-444.
- [11] Koldovsky Z., Tichavsky P., Oja E., 2006, "Efficient variant of algorithm fastICA for independent component analysis attaining the Cramer-Rao lower bound" *IEEE transactions Neural Networks*, 17, pp. 1265-1277.
- [12] Karvanen, J., Eriksson, J., Koivunen, V., 2000, "Maximum likelihood estimation of ICA model for wide class of source distributions", *Neural Networks for Signal Processing* 1, pp. 445-454.
- [13] Pham, D.T., Garat, P., 1997, "Blind separation of mixture of independent sources through a quasi-maximum likelihood approach", *IEEE Transaction Signal Processing*, 45, pp. 1712-1725.
- [14] Aapo Hyvärinen, 1997, "One unit contrast functions for Independent Component Analysis: A statistical analysis", *IEEE Signal Processing*, pp.388- 397.
- [15] Aapo Hyvärinen, August 1997, "Independent Component Analysis by Minimization of Mutual Information", Helsinki University of Technology, Laboratory of Computer and Information Science, Finland, Report A46.