Comparative Study of Performance of Particle Swarm Optimization and Fast Independent Component Analysis method in Cocktail Party Problem

Hawraa S. Hamza
University of Babylon/ College of Information Technology/ Department of Software
hawraa_shareef@yahoo.com

ARTICLE INFO
Submission date: 21/ 7/ 2019
Acceptance date: 3 /10/ 2019
Publication date: 31/ 3 / 2020

Abstract
There are many methods used for solving the Blind Source Separation problem, such as Independent Component Analysis which became the most commonly used method. ICA methods depend on one of two properties: sample dependency or non-Gaussianity. In our study, the cocktail-party problem processed using ICA method.

In this work, we studied the performance of two techniques with the independent component analysis is standard FastICA, and PSO; and compare the results of each algorithm with others according to some evaluation metrics (objective such as SNR and SDR) and (subjective such as signals plotting and playing). The implement of these algorithms was to be made with two source signals and three source signals. As in the evaluation process, the PSO gives more accurate results than FastICA.

Many input speech signals of 8 KHz sampling frequency, that achieve i.i.d. condition and well-condition were tested for different speeches for men and/or women, also music.

Keywords: BSS, Fast ICA, ICA, PSO.

1. Introduction
Blind Source (Signal) Separation (BSS) is a popular signal treatment technique that suggested and used in the 1990s. As an output of reaction among the neural networks, statistics processes, and the information theory. After that, it became an excellent topic for many types of research and used in a number of applications, as in the medical and medical sciences, signal telecommunications processing, sounds and speech processing, images processing, etc. [1]-[3].

The BSS, which is one of the challenges in the digital signal processing, requires recovering the original signals from multivariate data. The term “Blind” implies that the original signals are unknown, also the properties of the system are hidden or there is some of principles information about of the sources (such as non-Gaussian distribution, and statistically independence). In many applications, the observation signals can be noticed as a mixed signal of original signals, where the observation signals can represent as a sequence the outputs of sensors, each sensor receives multi-combinations of the original signals. The key function of BSS is to separate the mixed signals and recovers the original signals which desired from the observation signals data. Independent Component Analysis (ICA) is the most popular mechanisms for analyzing latent data and it is a statistical process for splitting a multivariate signal into additive subcomponents supposing the shared statistical independence of the non-gaussian signals of the sources. Only one of these properties (sample dependence or non-Gaussianity) are used by the ICA method [1],[2],[4].

The non-linear and linear ICA depending on the function of the mixing process. There are many methods in the linear ICA, such as Non-linear PCA [5], SOBI [5], JADE [6], EASI [7],[8], INFOMAX [8], FastICA [8], and RADICAL [9]. In all these methods, the source signals generated by using invertible filter driven by i.i.d. assuming to be a random process [2]. So, the similar pre-processing techniques such as whitening, centering, and de-noising processes followed by all ICA methods.

In this paper, we studied the performance of Particle Swarm Optimization (PSO) and compare them with the standard FastICA depending on some objective evaluation metrics (as SDR (signal-distortion ratio) and SNR (signal-to-noise ratio)) and subjective evaluation metrics (as signals plotting and playing). The implement of these algorithms was to be made with two source signals and three source signals.

We implement cocktail-party problem processed using ICA method with negentropy and kurtosis as fitness function, with the mentioned algorithms. Also, the study appears that the PSO gives more accuracy than other algorithms in signal separation. The speeches signals that tested were 8 KHz sample frequency for different speakers men and/or women and/or music.

The remainder of this paper ordered as follows: section 2 shows the related work , section 3 shows the related theories which include the basics of ICA, PSO, FastICA, and evaluation criteria (SNR, SDR). section 4 presents the comparative study. The results of experiments and evaluation measurements described in section 5 and finally section 6 shows the conclusion.

2. Related Works

The hybrid two algorithms, genetic algorithm and PSO were proposed for solving maximum likelihood ICA problem [16]. In other side, to reduce the shortcoming in Blind Source Separation through improved PSO via updating the dynamic inertia weight in PSO [17]. The advantages of the PSO algorithm was used to enhance the nonparametric ICA [18]. Quantum Particle Swarm Optimization (QPSO) is a new approach to improve the performance of ICA through Negentropy as a fitness function [19]. The application of scrambling based on ICA and PSO was presented in [20]ز

3. Background Theories

3.1: Independent Component Analysis (ICA)

Let the observation signal $x(t)=[x_1,x_2,\dots,x_n]^T$ represent $n\times 1$ mixed signals vector. Where n denote the number of sensors, t time coefficient, and T means the transposed of the vector x. Each variable in the vector x represents the mixed signal has been initialized in Multiple-Input/Multiple-Output model (MIMO). This mixture system as given in equation (1).

$$x(t)=As(t) \ldots\ldots\ldots\ldots\ldots \quad (1)$$

where $s(t)=[s_1,s_2,\ldots,s_n]^T$ is $n\times 1$ latent vector having independent and zero-mean non-normality distribution elements s_i (represent the original signals), and A is an unknown $n\times n$ non-singular and full-rank mixing matrix [2,4]. The model in equation (1) represents the general linear model of the BSS and/or ICA method. Mathematically, could not solve the equation (1) to find $s(t)$ and recover the original signals. Therefore, statistical estimation approaches have been employed for this purpose. The main output of the BSS methods expression is as in equation (2)

$$y(t)=Wx(t) \ldots\ldots\ldots\ldots\ldots \quad (2)$$

where $y(t)=[y_1,y_2,\ldots,y_n]^T$ represent $n\times 1$ separated vector and estimation of the original source signal, and W is an $n\times n$ estimated unmixing matrix (called separated matrix). More ICA algorithms can involve two stages; firstly: whitening process, where applying the second-order statistics for decoration. The process of estimating the orthogonal matrix is the objective of the whitening process. The orthogonal matrix is required to independence that accomplished in the next stage [1],[2].

The outcome of the independence hypothesis is the estimated sources signals then converted into the contrast function (optimization problem) that is become minimum when the estimated sources are independent.

3.1.1. Pre-processing of ICA

• **Centering**: this concept strongly relative to the central moment, the most important preprocessing for the ICA methods. It concerns with calculating the mean of the mixed signal vectors and this mean is subtracted from the mixed vector itself:

$$x' = x - E\{x\} \quad \ldots\ldots\ldots\ldots\ldots \quad (3)$$

A result vector called the mean vector or zero-mean vector. In addition, the independent components are made zero mean as well, as given in equation (4):

$$E\{s\} = A^{-1}E\{x'\} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (4)$$

After this process, the mixing matrix remains the same after this preprocessing, also this process can be done without affecting the estimation of the mixing matrix

• **Whitening**: the second most pre-processing is so-called whitening process. Also, it called the sphering process. After this process, obtaining uncorrelated mixed signals x and having the unit variance of the observed data. By applying the model in equation (5) can achieve the whitening process:

$$\tilde{x} = \Lambda D^{-1/2}\Lambda^T x \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (5)$$

Columns of Λ representing the eigenvectors of $E[xx^T]$ and the diagonal D represent the eigenvalues of $E[xx^T]$. The main function of this process (whitening) is to make the mixing matrix orthogonal [2].

3.1.2. Objective Functions in ICA:

• **Negentropy(Negative Entropy)**:

It depends on the amount of the theoretical information of the entropy. The entropy measures the randomness of the variable. The entropy is given in equation (6) for a discrete variable:

$$H(X)=-\sum P(X=a_i)\log P(X=a_i) \quad \ldots\ldots\ldots\ldots\ldots \quad (6)$$

The H denotes to the entropy of the observation signals and an estimation of the original signals, p is the probability of x, and x represent possible values of x. The Gaussian distribution variables have greater entropy than other variables. To measure the Gaussianity of the components, the Negentropy concept defined, as given in equation (7)

$$J(y) = H(y_g) - H(y) \quad \ldots\ldots\ldots\ldots\ldots \quad (7)$$

where y_g represent a Gaussian variable. That $J(y)$ is nonnegative, and for the Gaussian variable equal to zero. Negentropy is non-parametric possibly, statistically robust, and exhaustive computationally. Moreover, the estimation of it is too hard, hence the approximations used for this determination: [2]

$$J(y) \propto [E\{G(y)\} - E\{G(v)\}]^2 \quad \ldots\ldots\ldots\ldots\ldots \quad (8)$$

Where v and y signify Gaussian vectors without means (zero mean), and G signifies the no quadratic function.

• **Kurtosis:** The kurtosis is most popularly used to measure the Gaussianity, it represents the 4th order cumulant, and defined in equation (9).

$$k_4 = E\{x^4\} - 3[E\{x^2\}]^2 \quad \ldots\ldots\ldots\ldots\ldots\ldots.. \quad (9)$$

For the unit variance of x, the equation of the *kurtosis* can be rewritten as $E\{x^4\}-3$. This proves that the kurtosis is a normalized model of the 4th order moment $E\{x^4\}$, which is equal to $3[E\{x^2\}]^2$ for a Gaussian distribution. And therefore, the kurtosis value is zero. The kurtosis value of the most non-Gaussian distribution is nonzero.
Kurtosis could be defined in sign (-,+,0): (negative) sub-Gaussian, (positive) super-Gaussian and (zero) Gaussian, respectively.

3.2 Standard FastICA Algorithm:

The FastICA learning rule discoveries a direction unit vector w such that the projection $w^T x$ maximizes the value of nongaussianity. The approximation of negentropy is used to measure nongaussianity [1, 15]. The FastICA is established on a fixed-point iteration scheme for finding a maximum of the nongaussianity of $w^T x$. It can be also resulting in an approximate Newton iteration.
The basic formula of the FastICA can be summarized as follows:

1. Randomly set the initial weight vector W.
2. Repeat until convergence :
   2.1: Suppose $w^+ = E\{xg(w^T x)\} - E\{g'(w^T x)\}w$
   2.2: Suppose $w = w^+ / \|w^+\|$

3.3 The Particle Swarm Optimization algorithm (PSO)

PSO algorithm is presented by Kennedy and Eberhart in 1995. It's one of the search approaches that use the heuristical population search. In this method, each search term called Particle, have two main parameters position and velocity. Each particle search on best position in the local search space called local best position, even it stores all its positions in its memory. These positions defined as the experience of the current particle in the current dimension. Through the search procedure, the particle swarm discovers new positions in another dimension, the new positions called new experience, and so on.

The final and main job of the particle swarm is to find a global best position during the search iteration. The global best position represents the best position among the local best positions in n-dimensions of the current particle. The position and the velocity of each particle calculated in equations (10) and (11) respectively, for n iterations [12].

$$v_i(t+1) = wv_i(t) + c_1 r_1(t)(pbest_i(t) - x_i(t)) + c_2 r_2(t)(gbest_i(t) - x_i(t)) \quad \ldots\ldots\ldots\ldots\ldots \quad (10)$$

$$x_i(t+1) = x_i(t) + v_i(t+1) \quad \ldots\ldots\ldots\ldots\ldots \quad (11)$$

Where *v* denotes a velocity of a particle *P*, *x* represents the location of particle *P*, the best local location of the current *P is denoting by pbest*, and *gbest* denote a best global location for all particle in the search area in *n*-dimension. Where *w* denotes an inertia weight (for the convergence speed). So, c_1 and c_2 are the acceleration factors constants, Likewise, r_1 and r_2 denote two parameters valued randomly between 0 to 1. [10],[12].

3.4 Measurement Criteria:

we used many metrics for evaluating the performance of the proposed system, these metrics include subjective (such as plotting and playing) and objective (such as SDR and SNR).

The measurement metrics (SNR and SDR) that is viewed in the table (3) compare the speech signals evaluation process between the original signals and the separated one. For the SDR metric, a good and desired value is as high as possible, while in the SNR metric the best range is from 0 to 1, and if the value is closest to 0, then it considered as the best results. [3],[13],[14].

- **Signal to Noise Ratio (SNR)**:
  is considered as :

$$SNR = 10log_{10}\frac{\sum_{n=-\infty}^{\infty} s^2(n)}{\sum_{n=-\infty}^{\infty}(s(t)-\hat{s}(t))^2}\,(dB) \qquad (12)$$

before mixing, the original source signal denoting by *s*, and the estimated signal denoting by $\tilde{s}$.

- **Signal to distortion ratio (SDR)**:
  is considered as:

$$SDR = 10log_{10}\frac{\sum_t s(t)^2}{\sum_t (s(t)-\hat{s}(t))^2}\,(dB) \quad \ldots\ldots\ldots\ldots\ldots \quad (13)$$

before mixing, the original source signal denoting by *s*, and the estimated signal denoting by $\tilde{s}$.

4. The Proposed Study

The proposed study includes mainly two parts: the first part include solve the ICA problem using PSO as an optimization method. The second part, include a comparative study among some the ICA methods and the PSO based ICA proposed method.

Firstly, handling the cocktail-party problem by employing the PSO algorithm to optimize the linear ICA for separating various mixing sounds received from various microphones(sensors).

The proposed method is summarized in the following steps:

1. Under the property, i.i.d., Initialize at least two free noise speech signals with the same frequency and length (the system dealing with 8KHz sounds).
2. Implement Eq. (1) that is mixing the speech signals and after that initializing the mixture matrix that gives the best-mixed signal and accomplishes the *well-condition*,
3. Implement the preprocessing of ICA (centering and whitening) on the result of step 2.
4. Implement the ICA contrast functions (*Negentropy* and *Kurtosis*) for separating the mixing signals .

Particle Swarm Optimization includes the steps below:

1. Initialize the parameters of PSO as a population.
2. Set Kurtosis and Negentropy as a fitness function for the optimization method.
3. In each iteration, implement the centering and whitening.
4. After PSO, evaluate the results according to some evaluation metrics : (subjective such as playing and signals depict) and (objective such as SDR and SNR).

Secondly, the comparative part includes studying the behavior of the specified three methods (PSO based ICA, and standard FastICA) and the performance of each method according to many evaluation metrics (SNR and SDR).

5. Simulation Result and Analysis

During the proposed study, many pairs of speeches signals were examined. These speeches token from the database “ecs.utdallas.edu/loizou/speech/noizeus/”, which achieve the required sound properties (as the i.i.d.), and for men and women spoken in different circumstances and without noise. A cocktail party with two different speech signals were simulated and used as system input. Table (1) showed the original signals and the mixing one, and table (2) showed the separated signals.

According to the evaluation metrics (SNR and SDR), the proposed method (PSO based ICA) give , in most cases, give the best separation results, more accuracy in less elapsed time than other methods, as shown in table (3) and in analyzed figure (1) for the SNR measurement and figure (2) for the SDR measurement.

Table (1) : Original three pairs signals and mixsing Signals

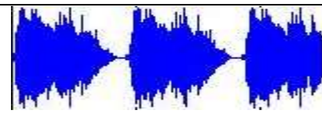
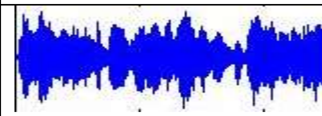
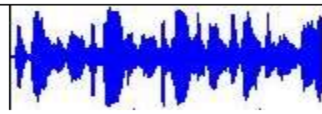
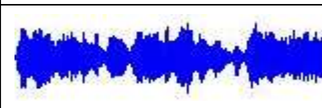
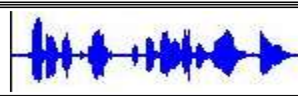
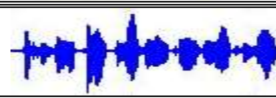
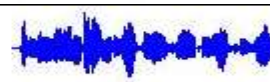
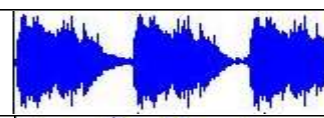
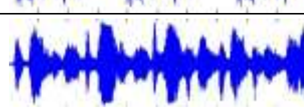
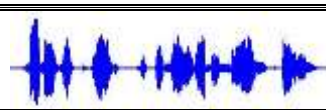
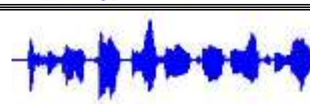
Signals Indications	Length	Source Signals	Mixed Signals
A-1	50000		
A-2	50000		
C-1	23323		
C-2	23323		
D-1	21582		
D-2	21582		

Table (2) : the separated signals depending the selected methods

Signals Indications	Separated Signals	
	FastICA	PSO
A-1		
A-2		
C-1		
C-2		
D-1		
D-2		

Table(3): the evaluation metrics of the selected signals under the specified methods

Speeches Indication	Speeches Length (T)	PSO		FastICA	
		SNR	**SDR**	**SNR**	**SDR**
A-1,A-2	50000	0.1992	**13.9693**	**0.0949**	13.6689
C-1,C-2	23323	0.1031	20.1893	**0.0048**	27.0463
D-1,D-2	21582	**0.0179**	**26.6643**	0.3772	20.0863

6. Conclusion

The most popular and trusty method used for handling the BSS problem is the Independent Component Analysis (ICA) method. It depends on the statistical independence and the non-Gaussianity of the mixed signals. This paper introduced the study for enhancing the performance of the ICA algorithm in speech separation depending on the benefits of PSO and comparing the results of PSO based ICA and the results of other methods. The cocktail-party problem was studied in this work by taking a real different speeches signals for different sentences to different speakers of sampling frequency 8 KHz. According to the evaluation metrics such as signal plotting, SNR and SDR, the results were accurate.

Conflict of Interests.
There are non-conflicts of interest

References

[1] P. Comon, and C. Jutten, “Handbook of Blind Source Separation, independent component analysis and applications”, academic press, Oxford 2010.

[2] A. Hyvarinen, J. Karhunen, and F. Oja, ”independent component analysis”, john wily & son,2007.

[3] S. Makino, T. W. Lee, and H. Sawada, “ Blind Speech Separation”, Springer, 2007.

[4] N. A. Muhsin, ”A comparison among adaptive ICA algorithms for blind speech signals separation: cocktail party problem”, Ph.D. thesis, University of technology, Iraq, 2006.

[5] S. B. Sadkhan, A. J. Alnaji and N.A. Muhsin, ”performance evaluation algorithms based on neural networks”. A 5th International Symposium on Communication Systems, Networks and Digital Signal Processing, CSNDSP, Greece, 2006.

[6] K. Zhang, G. Tian, and L. Tian,” Blind source separation based on JADE algorithm and application“, a 3rd international conference on mechatronics, robotics, and automation, ICMRA, 2015.

[7] A. Hyvarinen, “Independent Component Analysis: recent advances”, royal society publishing, December 2016.

[8] R. Mutihac, and M.M. Van Hulle, “ A comparative survey on adaptive neural network algorithms for independent component analysis“, Faculty of Physics, University of Bucharest, 76900 Romania.
[9] V. Krishnaveni *et al.* ,” comparison of independent component analysis algorithms for removal of ocular artifacts from electroencephalogram”, measurement science review, volume 5 section 2, 2005.
[10] J. Sun, C. Lai and X. Wu, “particle swarm optimization classical and quantum perspectives”, CRC Press, 2012.
[11] J. Sun, B. Feng and W. Xu, “particle swarm optimization with particles having quantum behavior”, IEEE-2004.
[12] J. Kennedy and R. Eberhart, ”particle swarm optimization“, Proceedings of the IEEE international conference on neural networks, Perth, Australia, pp. 1942-1948, 1995.
[13] Nidaa AbdulMohsin Abbas, “Image encryption based on Independent Component Analysis and Arnold’s Cat Map”, Egyptian Informatics Journal, 2015, http://dx.doi.org/10.1016/j.eij.2015.10.001.
[14] E. Vincent *et al*,” Performance measurement in blind audio source separation”, IEEE Transactions on Audio, Speech and Language Processing, Institute of Electrical and Electronics Engineers, 2006, 14 (4), pp.1462-1469.
[15] A. Hyvarinen and E. Oja, “Independent Component Analysis: Algorithms and Applications”, Neural Networks, 1999.
[16] H. Azad and M. Hatam, “Maximum Likelihood Independent Component Analysis using GA and PSO “, 2016 24th Iranian Conference on Electrical Engineering (ICEE), 2016 IEEE.
[17] M. Li et al, “Blind Source Separation based on Improved Particle Swarm Optimization”, International Conference on Artificial Intelligence and Computational Intelligence 2009, IEEE Computer Society 2009.
[18] D. Krusienski and W. Jenkins, “Nonparametric Density Estimation Based Independent Component Analysis via Particle Swarm Optimization”, ICASSP-2005, IEEE 2005.
[19] Nidaa A. Abbas, Hussein Mohammed Salman, “Independent component analysis based on quantum particle swarm optimization”, Egyptian Informatics Journal, Vol. 19, Issue 2, July 2018, pp. 101-105, Publisher Elsevier.

الخلاصة

هنالك الكثير من الطرق التي تستخدم لحل مشكلة فصل المصدر المحجوب، مثل طريقة تحليل المكونات المستقلة والتي اصبحت من اكثر الطرق استخداما. طريقة تحليل المكونات المستقلة تعتمد على واحدة من اثنتين من الخصائص: استقلالية العينة او non-Gaussianity. في هذا البحث استخدمت طريقة فصل المكونات المستقلة لحل مشكلة حفلة الكوكتيل. حيث تمت دراسة انجازية طريقتين: طريقة فصل المكونات السريعة وطريقة تحسين سرب الطيور ومقارنة النتائج بالاعتماد على بعض مقاييس الانجازية مثل (الموضوعي مثل SNR وSDR) و (ذاتي مثل single plotting و playing) . حيث طبقت الخوارزميتين على مصادر ذوي اشارتين وثلاث اشارات. وكنتيجة لعملية التقييم فأن خوارزمية فصل المكونات السريعة اعطت نتائج اكثر دقة من خوارزمية تحسين سرب الطيور. حيث استخدمت اشارات للكلام بتردد 8 كيلو هرتز والتي حققت شروط كل من ال i.i.d و well-condition والتي اختبرت على احاديث مختلفة لرجال ونساء وكذلك الموسيقى.

الكلمات الدالة: فصل المصدر المحجوب، طريقة فصل المكونات السريعة، طريقة فصل المكونات، طريقة تحسين سرب الطيور.