Science Development
2021; 2(1): 1-6
http://www.sciencepublishinggroup.com/j/scidev
doi: 10.11648/j.scidev.20210201.11

Speech Signals Separation Using Optimized Independent Component Analysis and Mutual Information

Hussein Mohammed Salman

College of Material Engineering, University of Babylon, Babylon, Iraq

Email address:
Hus14ms@uobabylon.edu.iq

To cite this article:
Hussein Mohammed Salman. Speech Signals Separation Using Optimized Independent Component Analysis and Mutual Information. *Science Development*. Vol. 2, No. 1, 2021, pp. 1-6. doi: 10.11648/j.scidev.20210201.11

Received: November 30, 2020; **Accepted**: December 14, 2020; **Published**: January 4, 2021

Abstract: One of the still problems in the Digital Signals Processing is the Blind Signal (Source) Separation (BSS). The BSS mean how to recover the original (source) signals from mixed (observed) signals via many sensors. There are many methods are used in the Blind Signal (Source) Separation problems specifically Cocktail Party problem, such as Independent Component Analysis (ICA), which has become most commonly used. Also, In more cases of the BSS problems especially the Cocktail-Party case there are number of challenges as number of mixed signals and the mixture type. In order, to enhance the performance of the ICA there are many studies for this purpose that depend on the optimization mechanisms as genetic algorithm and Particle Swarm Optimization (PSO). The advantages of a Quantum Particle Swarm Optimization (QPSO) are employed to improve the efficiency of the ICA approach using mutual information function as modern technique, which is used in de-mixing of the speech signals. In this work, a new technique is introduced, is QPSO-based ICA by using *Mutual Information* function as an objective function for the optimizing process. The presented method has been implemented on the real three different speech signals, with 8 KHz frequency. The results was high accuracy in the signals and more efficient in the computations requirements as the time and space which are measured by the evaluation metrics as the signal plotting, SNR, SDR, and Computation Time.

Keywords: BSS, Mutual Information, ICA, QPSO, Cocktail-Party Problem

1. Introduction

A Blind Signal Separation (BSS) problems are a set of the Digital Signal Processing (DSP), that aim to conclude a some of latent speech signals by using a some of predefined statistical attributes of these latent. The BSS emerged at late of 1980s and then extended swiftly. There are many main resources illustrate the BSS in details [1-3].

In BSS, many sources are observed by a class of microphones (sensors), that are ordered to rebuild the original sources. It suppose that the received data is concluded by interactions among latent factors.

Independent Component Analysis (ICA) is most commonly approach used to analyze multivariate data. The statistical method (ICA) is used for recovering a multivariate observation into additive components assumption that the mutual statistical non-dependence of non-Gaussian distribution of a source signals. ICA approach uses one of two aspects: sample dependence or non-Gaussianity [1, 2].

The non-dependence presumption satisfies in more cases, also, the ICA-based blind separation of an observation signals gives more great attributes. They use statistical advantages of the sources in the non-dependent latent by reducing the statistical dependence of the evaluated signal components (factors). Subsequently, Non-Gaussianity property is exploited for measuring the non-dependence of the factor, by the kurtosis metric or the Negentropy metric [4, 5].

The ICA approach and its methods assume the observations signals have produced by an i.i.d. random process as an invertible filter driven [2]. So that, the ICA approach adapting same pre-processes are the centering and whitening. There some traditional methods of the Linear ICA are: Non-linear PCA [6], SOBI [6], JADE [7], EASI [8, 9], INFOMAX (Bell-Sejnowski) [9], FastICA [9], and RADICAL [10].

This paper introduces a method to solve the cocktail party problem by Quantum Particle Swarm Optimization (QPSO) and employing the mutual information function. The QPSO method is one of the probabilistic algorithms. It requires an easier and fewer parameters for an implementation than PSO algorithm. One of the advantages of the QPSO is used to solve more optimization problems [11, 12].

2. Background Theories

A. The Independent Component Analysis (ICA)

It is statistically approach based on some attributes of the observation signals to recover the original signals. This approach assume there no previous knowledge about the mixed (received) signals. The mathematical formula of the mixing process can be formed as:

$$x(t)=A*s(t) \quad (1)$$

$s(t)$ represent original signals, $x(t)$ denotes the mixed signals, A denote the mixing matrix of latent mixing parameters, and t denote to time factor. The ICA is to discover the inverse of the matrix A, *(W)* that results y, the de-mixing transformation model is as:

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

Note that y(t) is the estimation of source signals, and W represents the matrix of the de-mixing treatment. The ICA method base on two axioms are the objective function and the optimization method. The objective function used to maximize the independence among the components and the optimization method used to optimize the maximization of the objective function. The ICA approach includes main two stages are the whitening process, where the second order statistic arithmetic used to perform decorrelation process. The first stage produces an orthogonal matrix indeed necessary in the second stage of the ICA [1, 2, 5].

- Pre-Processes in the ICA

1. Centering: in this process the first order statistic (mean) of the observation signal vectors is computed then subtracted from same mixed signals, this operation computed as in the equation (3).

$$v'=v-E[v] \quad (3)$$

where v is the observation source and $E[v]$ represents the mean of the observation signal (Expectation). Adding the expectation vector to the produced source signals:

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

where s' is the estimated source vector and A^{-1} represents the inverse of the matrix A (mixing matrix), so $A^{-1}E[x]$ is the mean vector.

2. Whitening: This process most pre-processing important in overall ICA approach. It is model of linear transformation of the centered source vectors; include producing unit variance and uncorrelated mixed signals. The whitening transformation can be formed as in equation (5)

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

Where Λ and D two matrices. The Λ' columns are the eigenvectors of $E[xx^T]$, and the diagonal of D are the eigenvalues of $E[xx^T]$. Main advantage of the whitening is produce orthogonal mixing matrix which used to recovering the original signals [2].

Objective Functions in ICA

3. Mutual Information: one of the most information theory measures is used as an objective function to evaluate and/or solve several continuous and discrete problems. The measures of information theory are Kurtosis, Negentropy, Mutual Information, and Maximum Likelihood. The Mutual Information amongst variables, x_i, i=1, 2,.., m, defined based on the differential entropy, as illustrated in the equation (6):

$$MI(x_1, x_2, \ldots, x_m) = \sum_{i=1}^{m} H(x_i) - H(x) \quad (6)$$

Mutual Information (MI) is a measurement of the variables dependency. It equivalent to the Kullback Leibler concept. MI measurement indicates that the x_i are independent when it near to zero [2].

MI can be used as estimation metric of the ICA because it measures the theoretical information of the non-dependency of the observation signal vectors. This manner can be used instead of the standard ICA model (s_i=Wx_i), by estimating the matrix W through minimizing the MI factor among mixed signal vectors [2, 9, 13, 14].

The problem with MI is that it is difficult to estimate. Therefore, to solve this problem, an approximation of mutual information proposed based on polynomial density expansions by using higher-order cumulants (skews and kurtosis) as shown in equation (7) [15]

$$MI(x) \approx C+\frac{1}{48}\sum_{i=1}^{m}\left[4k_3(x_i)^2+k_4(x_i)^2+7k_4(x_i)^4-6k_3(x_i)^2 k_4(x_i)\right] \quad (7)$$

Where C is constant used to adjust the approximation.

B. Optimization Methods

- Particle Swarm Optimization (PSO)

By emulating of the animal life behaviors (especially the bird flocking and fishes), the scientists Kennedy and Eberhart introduce new optimization method used to enhance the performance many optimization methods. That method called "Particle Swarm Optimization" [16].

The PSO algorithm begins with selecting any set of particles randomly. These particles represent initial state of the solution of the problem. After that, it search about better state than initial state and reset the swarm for the optimal state based on its research state. The method reset of its factors based on two parameters: the position and the velocity of each particle in n-dimension in the space state for all particles. It repeats the looking and resetting for n iteration as:

$$v_i(t+1)=wv_i(t)+c_1 r_1(t)(pbest_i(t)-x_i(t))+c_2r_2(t)(gbest_i(t)-x_i(t)) \quad (8)$$

$$x_i(t+1)=x_i(t)+v_i(t+1) \quad (9)$$

The parameter v represents the velocity of the current particle, x denotes the general of the particle. The local position of the current particle denoted by *pbest.* So, the global position of the current position denoted by *gbest.* Also, there is important parameter used for the convergence speed of the algorithm called inertia weight denoted by "w". The parameters c_1 and c_2 are the two acceleration constants. The two parameters r_1 *and* r_2 are valued in the range [0 to 1] randomly, [11, 16].

- *Quantum Particle Swarm Optimization (QPSO)*

At late 2004, *Jun Sun* and others mentioned new version of the PSO algorithm named as Quantum Particle Swarm Optimization (QPSO). There are some differences between PSO and QPSO as the velocity where the QPSO not require it, so the QPSO easier in the implementation because it have few parameters. It can improve the performance of many algorithms used in many discrete and continuous optimization issues [11, 12, 17]. The description of the QPSO method is:

Assume that each particle searches in each dimension with a δ potential on a particular dimension, nearly the point p_{ij}. In simple state, by considering the particle in a selected dimensional space with position of p of the potential. By solving the *Schrödinger* function of a selected dimensional δ potential good, it can get on the distribution formula *F,* and the probability density function *Q* as:

$$Q(X_{i,j}(t+1))=\frac{1}{L_{ij}(t)}e^{-2|P_{i,j}(t)-X_{i,j}(t+1)|/L_{i,j}(t)} \quad (10)$$

$$F(X_{i,j}(t+1))=e^{-2|P_{i,j}(t)-X_{i,j}(t+1)|/L_{i,j}(t)} \quad (11)$$

Note that $L_{i,j}(t)$ denotes the standard deviation, $X_{i,j}$ denotes the *i*th, *j*th position of the particle P_{ij}. By applying the Monte Carlo model as in (14), the position of the particle is:

$$X_{i,j}(t+1)=P_{i,j}(t)\pm\frac{L_{ij}(t)}{2}\ln(1/u),\ u=\text{rand}\ (0,1) \quad (12)$$

when the $L_{i,j}(t)$ is evaluated, the *mean best position* (denoted by *m*), represent the global best point of the population, is calculated. The *m* is defined as the mean of all the *pbest* positions of all particles.

$$m(t)=(m_1(t),m_2(t),,m_n(t)= \left(\frac{1}{M}\sum_{i=1}^{M}P_{i,1}(t),\frac{1}{M}\sum_{i=1}^{M}P_{i,2}(t),\frac{1}{M},\dots,\frac{1}{M}\sum_{i=1}^{M}P_{i,n}(t),\right) \quad (13)$$

Note that *M* represent the size of the population and P_i is the *pbest* position of *i*th particle. The values of $L_{i,j}(t)$ is:

$$L_{i,j}(t)=2ß\cdot|m_j(t)-X_{i,j}(t)| \quad (14)$$

and thus the position is:

$$X_{i,j}(t+1)=P_{i,j}(t)\pm ß\cdot|m_j(t)-X_{i,j}(t)|\cdot\ln(1/u) \quad (15)$$

Inform that ß is particular parameter used to control with the algorithm convergence called contraction–expansion factor. As result, the equation (15) represent the QPSO formula.

C. Evaluation Measurements

To measure the performance of the introduced method, there are many subjective measurements used as the signals plotting; and many objective measurements used as SNR, SDR, MSE, and Computation Time.

The measurement metrics have been applied on the source signals and the recovered signals. The rane scale of SNR measurement is between 0 to 1, best scale of SNR is nearest to 0; whilst in the SDR measurement, the best scale is higher value [3, 18, 19].

1. Signal-to-Noise-Ratio (SNR): the general form of the SNR measurement is:

$$SNR=10\log_{10}\frac{\sum_{n=-\infty}^{\infty}s^2(n)}{\sum_{n=-\infty}^{\infty}(s(n)-\hat{s}(n))^2}(dB) \quad (16)$$

Where *s(n)* represent the source signal vector, and $\hat{s}(n)$ represent the recovered signal vector.

2. Signal to Distortion Ratio (SDR): the general formula of the SDR measurement is:

$$\text{SDR}=10\log_{10}\ \frac{\sum_t s(t)^2}{\sum_t(\tilde{s}(t)+\text{interf}+\text{artif})^2}\ (\text{dB}) \quad (17)$$

Where $s(t)$ represent the source signal vector, and $\tilde{s}(t)$ represent the recovered signal vector.

3. The Computation Time: this measurement used to measure the consumed time to separate the mixed signals under the proposed method (ICA-QPSO) and the standard method (ICA-PSO) in seconds. By using MATLAB R2017b under Windows7 (64-bit) and the Intel Core i7-5500U with 2.40 GHz.

3. Research Methodology

A. Initializing the Mixed Sounds

In this phase, initializing the required mixed files under some considerations: as mono speeches under the frequencies 8KHz, clean (noiseless), and wav format; also achieve the identical, independent distribution (i.i.d.) as possible.

A number of speech files recorded for many cases found in an international database are used in the proposed system, as the database of the ITU (International Telecommunication Union) [20], and the database of the scientist Loizou from the University of Dallas [21].

After that, randomly determine the mixed matrix that can achieve best mixed-case under the well-condition number of the mixed matrix. For each mixed case, the mixing matrix was created separately in normal distribution with certain intervals between -1 to 1 randomly. These intervals were determined manually and relative to achieve well-condition number of the mixing matrix. The mixing matrix initialized by A=a+(b-a)*randn (2, 2).

Table 1 describes the mixed (observation) speech signals, and the mixture matrix and its condition number for two mixed (observation) cases of three source speech signals for each mixed case. The manipulated sounds was taken from the database of ITU organization and the University of Dallas database. The speech signals are mixed with linear instantaneous mixture based on the ICA model as defined in equation (1). The columns of the *kurtosis* of the source signals and mixed signals show that the selected sources achieved the i.i.d. condition.

Additionally, the determined mixing matrix achieved the well-condition, where the sources are super-gaussian.

Table 1. *The Speech Signals and Mixing Matrix.*

Mixed Case No.	Source Signals Files Names	Kurtosis of Source Signals	Kurtosis of Recovered Signals	Length (samples)	a	b	Mixing Matrix
	Julia	7.3709	7.3503				-2.9349, 2.8909, -0.2048
1	Lauren	7.2118	7.2267	61038	-1	1	2.5983 -1.1030, -2.3122
	Ray	7.4982	7.4981				-2.6364,-1.7024, -2.3940
	Julia	7.3709	7.5246				2.5981, 1.6145, 1.7437
2	Source11	4.2686	4.2678	50000	0	1	-1.5430, 2.9281, 1.7149
	Source22	6.1309	6.1296				1.5557, 2.2052, -2.4348

B. Separation Process (Proposed Method)

Two main stages in the presented method are, first stage, is the mixing process and whitening. Second stage is enhancement results of the previous stage by implementing a specified optimization method. In the proposed method the QPSO method is used as an optimization method. The proposed method is performed in the ICA approach to apply the cocktail-party idea, which concentrate how separate (recover) many mixed signal received by number of sensors. Both two stage will illustrated in detail as follow:

First stage includes the following steps:

1. Set three noiseless speech signals. These signals must be under same frequency (mono speech 8KHz), and same length (samples) and with the i.i.d. condition.
2. Perform the mixing process among the received speech signals under the equation (1). The equation (1) requires mixing matrix under the *well-condition* criteria.
3. Centering and whitening (main two pre-processes) are implemented on results of the mixing process.
4. Implementing the *Mutual Information* as the contrast function in ICA as shown in equation (7), which is used to separate the mixed signals.

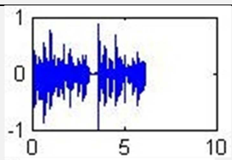
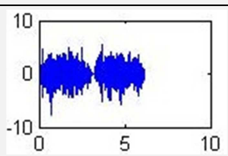
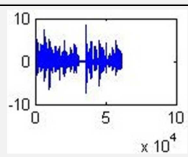
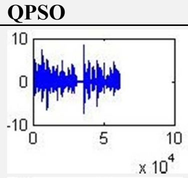
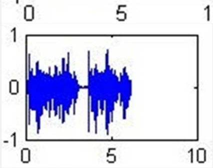
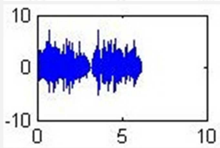
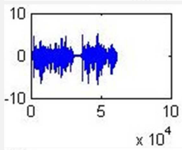
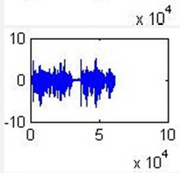
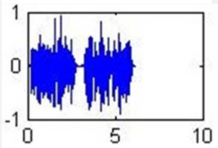
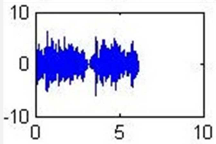
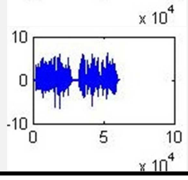
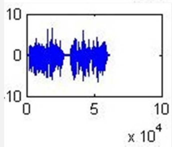
Second stage includes some steps are:

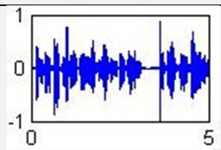
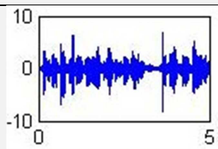
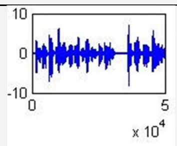
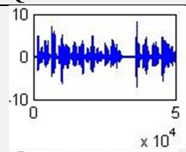
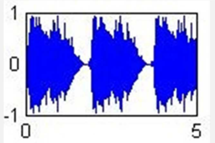
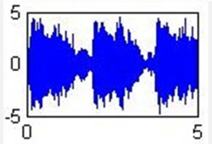
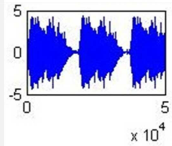
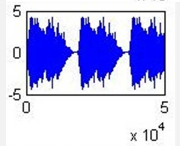
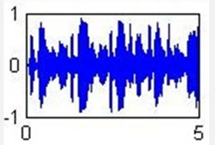
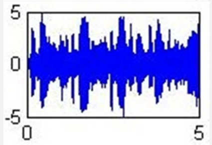
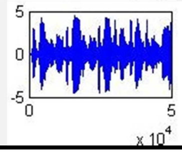
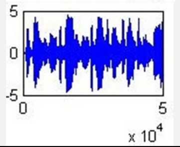
1. The initial coefficients of the QPSO method are maximum iteration (maxiter), population, and CE (Contrast-Expansion) coefficient (abbreviated with α). Where set maxiter=80 and population=20. So that, in this presented method, the CE coefficient will take a value 0.75 that gave good results in the optimization stage.
2. Set the cost value and cost function of the optimization algorithm. Here, the *Mutual Information* function is used as the cost function.
3. In each loop of the optimization algorithm, the centering and whitening are performed.
4. Set the "*mean best*" coefficient for each particle according to the equation (13).
5. In all iterations, the method updates the cost value based on the results of the cost function, and determines the best values.
6. Final step in the second stage is the evaluation process. In this measurements, applying some subjective measurements (as sketching the source and recovered signals) and objective measurements (as SNR, SDR, and Computation Time).

4. Results and Discussion

To perform the validity of the introduced method, two triples mono-speech signals are tested, these speeches are clean (noiseless) and in different conditions spoken. By simulating the cocktailing with triple different mono-speech signals, and input them to the system. To evaluate the proposed method, the ICA based PSO method is applied for this purpose. Based on the subjective metric, the sketching of the (source, mixed, separated) signals are showing in Table 2.

Table 2. *Plot Original, Mixed and Separated Speech Signals.*

Speech Signal Names	Original Speech	Mixed Speech	Separated Speech PSO	Separated Speech QPSO
Julia				
Lauren				
Ray				

Speech Signal Names	Original Speech	Mixed Speech	Separated Speech	
			PSO	QPSO
Julia				
Source11				
Source22				

In addition to the signals plotting, this paper used objective measurements as SNR, SDR, MSE, and computation time. These measurements applied with both methods ICA-PSO and ICA-QPSO, as shown in table 3.

Table 3. *Results of evaluation measurements.*

		SNR (*dB*)	SDR (*dB*)	Time Computation (seconds)
1st separating case	ICA-PSO	0.1798	18.2063	98.9799
	ICA-QPSO (proposed method)	0.0886	18.6450	97.6865
2nd separating case	ICA-PSO	0.1839	14.3748	81.5166
	ICA-QPSO (proposed method)	0.1130	15.4167	80.1542

Table 3 describe the evaluation metrics between the source signals and the separated signals. SNR, SDR, and Computation Time are used in the evaluation process. The best range scales of SNR measurement is between 0 to 1, the best results are closest to 0. In the first separated case of the proposed method is note that SNR=0.0886 *dB*, less than PSO, also in the second separated case note that SNR=0.1130 *dB*, less than PSO. This prove that the proposed method gave good separating results. Additionally, in the SDR measurement, a higher value is good and desired. As shown in the table 3. The proposed method ICA-QPSO in the first case, the SDR=18.6450 greater than PSO, also in the second, the SDR=15.4167 greater than ICA-PSO. These results mean that the separated signals are more accuracy and the error is more slightly. So, the computation time measurement illustrates that the separation process in the proposed method ICA-QPOS consume time less than ICA-PSO method. The results in the table 3 are illustrated in the figure 1.

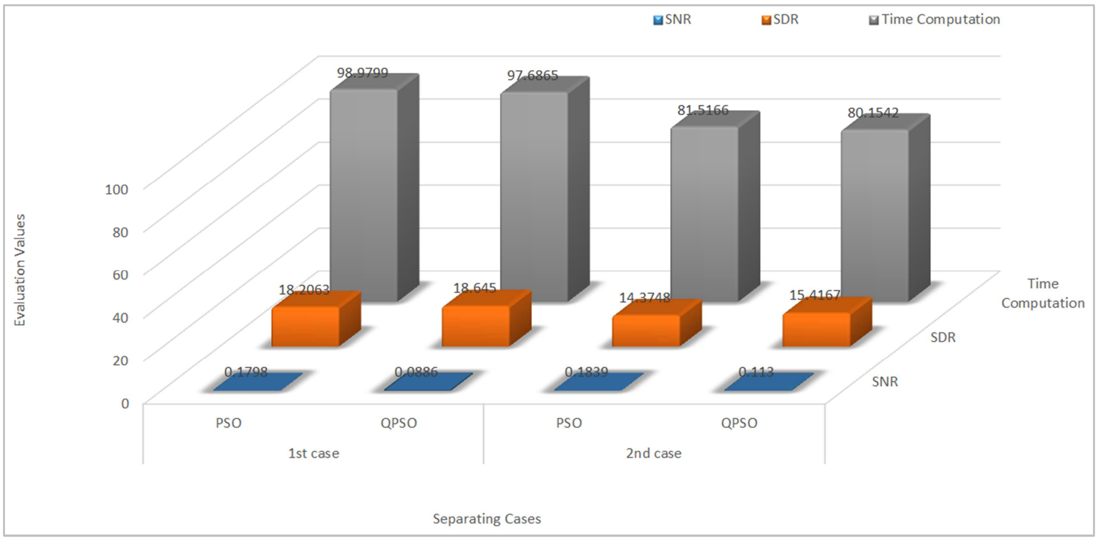


Figure 1. *Evaluation Measurements of Separating Cases.*

5. Conclusions

To solve the BSS problem, there are number of methods as the Independent Component Analysis (ICA) method. Also, In more cases of the BSS problems especially the Cocktail-

Party case there are number of challenges as number of mixed signals and the mixture type. The ICA depends on objective function and optimization method. In this paper, we simulate cocktail party with mixing three signals in instantaneous linear mixture. The *Mutual Information* transformation is used as an objective function in the ICA method, so the QPSO optimization method is used as an optimization method to enhance the performance of the ICA method. The signals that be used in this paper are mono speech signals with 8KHz frequency.

The proposed method compared with standard PSO method and evaluate with some metrics as SNR, SDR, and the Computation Time. The proposed method gave good results according these evaluation measurements and compared with PSO method.

References

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

[2] A. Hyvarinen, Juha Karhunen, and Erkki Oja, 2001, "Independent Component Analysis", john wily & son.

[3] A. Tharwat, "Independent Component Analysis: An Introduction", 2018, *Applied Computing and Informatics*, Elsevier, King Saud University.

[4] H. M. Salman, 2019, "Mono Speech Signal Separation Using Optimize Independent Component Analysis Algorithm", Ph. D. Dissertation, University of Babylon, Iraq.

[5] A. Nidaa Abbas, M. Hussein Salman, 2020, "Enhancing Linear Independent Component Analysis: Comparison of Various Metaheuristic Methods", *Iraqi Journal for Electrical and Electronic Engineering*, vol. 16, Issue 1, 113-122, https://doi.org/10.37917/ijeee.16.1.14.

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

[7] K. Zhang, G. Tian, and L. Tian, 2015," Blind source separation based on JADE algorithm and application ", 3rd ICMRA-2015.

[8] A. Hyvarinen, 2016, "Independent Component Analysis: recent advances", royal society publishing.

[9] R. Mutihac, and M. M. Van Hulle, 2009, "A comparative survey on adaptive neural network algorithms for independent component analysis ", Faculty of physics, university of Bucharest, 76900 Romania.

[10] V. Krishnaveni, S. Jayaraman, P. M. Manoj Kumar, K. Shivakumar, K. Ramadoss, 2005, "comparison of independent component analysis algorithms for removal of ocular artifacts from electroencephalogram", Measurement Science Review, vol-5, sec-2, 67-78.

[11] Sun J, Lai C., and Wu X., 2012, "particle swarm optimization classical and quantum perspectives", CRC-press.

[12] J. Sun, B. Feng, and W. Xu, 2004, "particle swarm optimization with particles having quantum behavior", IEEE, DOI: 10.1109/CEC.2004.1330875.

[13] B. Paprocki, A. Pregowska, and J. Szczepanski, 2020, "Optimizing information processing in brain-inspired neural networks", bulletin of the polish academy of sciences technical sciences, Vol. 68, No. 2.

[14] D. Fantinato, L. Duarte, Y. Deville, R. Attux, C. Jutten, and A. Neves, 2019, "A second-order statistics method for blind source separation in post-nonlinear mixtures", *Signal Processing*, Elsevier, 155 (2019) 63-72.

[15] A. Hyvarinen, 1999, "*Survey on Independent Component Analysis*", Neural Computing Surveys 2, 94-128, 1999.

[16] J. Kennedy and R. Eberhart, 1995, "particle swarm optimization ", proceeding of the IEEE international conference on neural networks, Australia.

[17] M. Xi, J. Sun, and W. Xu, 2008, "An improved quantum-behaved particle swarm optimization with weighted mean best position", journal of applied mathematics and computation, Elsevier, vol. 205 (2008), pp 751–759, DOI: 10.1016/j.amc.2008.05.135.

[18] N. A. Abbas, 2015, "Image encryption based on Independent Component Analysis and Arnold's Cat Map", Egyptian Informatics Journal, dx.doi.org/10.1016/j.eij.2015.10.001.

[19] E. Vincent, R. Gribonval, C. Fevotte, 2006," Performance measurement in blind audio source separation", IEEE, vol. 14, issue 4, pp: 1462-1469 DOI: 10.1109/TSA.2005.858005.

[20] https://github.com/dennisguse/ITU-T_pesq/tree/master/comform, accessed at 2019.

[21] http://www.utdallas.edu/~loizou/speech/noizeus/, accessed at 2019.