ICSP2012 Proceedings

A Novel Evolutionary Approach for Blind Source Separation Based on Stone's Method

Abdullah Ahmed, Zhang Chao Zhu, Cang Yan
College of Information and Communication Engineering, Harbin Engineering University, Harbin, China
Ahmed_Albakri1977@yahoo.com, zhangchaozhu@hrbeu.edu.cn, cangyan@hrbeu.edu.cn

Abstract— **In recent years, much research emerged to modify Stone's BSS method for solving a blind source separation problem; Stone's method used to recover original signals from the mixture. In this work, new direction has been opened to use an intelligent soft computing technique (Fast Genetic Algorithm) with the temporal predictability of signals based on Stone's BSS method. Proposed algorithm has been compared with well-known BSS algorithms (JADE, FICA, and Stone's BSS) over super-Gaussian, sub-Gaussian, and Gaussian signals with linear mixture combination. Then eight voices have mixed randomly; and the proposed approach has successfully recovered the voices with high efficiency. Interpretation based on the responses of two different linear scalar filters to the same set of signals, which indicate to Short-term and Long-term linear predictors with tuned Half-life values (h_L, h_S) genetically is a powerful new technique for solving BSS problem. In addition to the benefits of the Stone's method, the proposed algorithm overcomes the local minima problem by successfully jump out of the potential local minimum. Usually recovery methods depend on the difference between signals and mixture proprieties; generally, there are three famous properties for any signal: (1) Gaussian probability density function based on the central limit theorem (2) Degree of statistical independence (3) Temporal predictability. So (1&2) proprieties, have previously been used as a base for separation but in this work only 3rd property has been used. In order to check the effectiveness of the proposed algorithm two performance indexes are used: Interference to signal ratio (ISR), and Integral square error (ISE).**

Keywords- blind signal separation; signal temporal predictability; fast genetic algorithms; Stone's BSS.

I. INTRODUCTION

Blind source separation is a significant and formal area of research in signal and image processing [1]. It refers to the problem of restoring original sources or data from the mixture without prior knowledge of mixing process and source signals [2]. Various methods have been expanded in statistical, neural computing, and signal processing to find the solution of BSS problem and to get an appropriate linear representation of random variable [3]. The most significant method is Independent Component Analysis (ICA) as a statistical method for finding independent components of the signal, which have the most statistical independence in multidimensional statistical data. ICA is based on random and natural gradient [4, 5]. The famous methods of estimating ICA mode by: maximizing non-Gaussianity [2], minimization of mutual information [3], maximum Likelihood Estimation [6], JADE algorithms [7]. With a view to minimize appearance of probability density functions of source signals, a class of second-order statistic method is called temporal predictability (TP) method or Stone's method is proposed by Stone [8]. TP method established on conjecture that TP of any mixture is less than (or equal to) that of any of its components. However, this conjecture according to [9] is incorrect and modified. It can be considered the improved conjecture as a theoretical basis for BSS problem. Another view is talk over to measure TP, which depends on difference measure [10]. Modified algorithm using signal temporal predictability is proposed in [11] depends on the fact that temporal predictability (TP) of signals is predominantly different. In recent years, genetic algorithm (GA) has been successfully applied to solve different problems in business, engineering, and science [13]. GA has been utilized to solve the BSS problem. Two evolutionary algorithms, namely, particle swarm optimization (PSO) and genetic algorithms (Continuous and Binary) with a novel fitness function based on the fusion of two criteria, mutual information and kurtosis are proposed in [5] to solve BSS. Evolutionary algorithm is used for nonlinear blind source separation; Niche Genetic Algorithm is used in combination with nonlinear BSS to solve global optimization of parameters with high accuracy, fast convergence, and robustness against local minima [12]. A method for blind separation of digital signals based on Elitist Genetic Algorithms is proposed in [12]. In this work, a fast genetic algorithm (FGA) is used to generate and tune Half-life (h_L, h_S) parameters, which used by Stone's method, because the proposed algorithm is based on the responses of two different linear scalar filters to the same set of signals, linear mixing model is considered here . The schematic diagram for mixing and separation process in blind source separation techniques is shown in Fig.1.

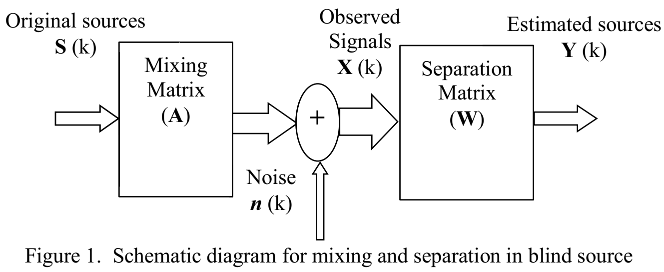

Figure 1. Schematic diagram for mixing and separation in blind source separation (BSS)

978-1-4673-2197-6/12/$31.00 ©2012 IEEE

Typical linear mixing model of blind source separation with n sensors and n sources are taken. $\boldsymbol{S}(k) = [s_1(k), . s_n(k)]^T$ refers sources; superscript T refers transpose operator; $\mathbf{X}(k) = [x_1(k), \dots x_n(k)]^T$ refers observation, $\mathbf{A} \in \mathbf{R}^{n\times n}$ refers mixing matrix. Symbol (k) is time or sample index.

The mixing model used in this work is:

$$\mathbf{X}(k) = \mathbf{A}\,\mathbf{S}\,(k) \tag{1}$$

And the separating model is:

$$\mathbf{E}\,\mathbf{Y}(k) = \mathbf{W}\,\mathbf{X}\,(k) \tag{2}$$

That $\mathbf{E}$ is permutation and scaling matrix and $\mathbf{Y}(k) = [y_1(k), \dots y_n(k)]^T$ denotes recovered sources. BSS problem is to estimate the best separating matrix $\mathbf{W} \in \mathbf{R}^{n\times n}$.

The remnant of the paper is organized as follows: specific measures for the temporal predictability are explained in the next section. The proposed algorithm is presented in section III, and then simulation results are shown in section IV. Finally; the conclusions are given in section V.

II. Signals Temporal Predictability Measure

Nearly every signal within a physical system is a factually mixture of statistically independent source signals. However, because source signals are usually generated by motion of mass (e.g., Membrane), form of physically possible source signals is underwritten by laws that govern how masses can move over time [8]. This indicates that most parsimonious clarification for complexity of giving an observed source signal is that combine of mixture. Stone proposes a method depend on the temporal predictability of signal and then several measures invented by researchers.

A. Stone's Temporal Predictability Measures.

Stone has been drawing many researchers' attention for BSS by using short- term and long-term predictors. Attempts for a weight vector which pack up orthogonal projection of signals such that each extracted signal is maximally predictable. It's a batch approach with low complexity. Stone's BSS conjecture refers that temporal predictability (TP) of any mixture is less than (or equal to) that of any of its components. Stone's measure of TP for N-sampled signal has been defined as follows [8, 14]:

$$F(y) = log\frac{V_y}{U_y} = log\frac{\sum_{t=1}^{N}(y_{long}(k)-y(k))^2}{\sum_{t=1}^{N}(y_{short}(k)-y(k))^2} \tag{3}$$

Where the symbol (k) is a time index, y *(k)* is the signal value of at time *k*. Term U_y contemplates the extent to which *y (k)* is predicted by short-term moving average (y_{short}). In contrast, the term V_y is a measure of the overall variability in y that is measured by the extent to which *y (k)* is predicted by long-term moving average (y_{long}). Predicted values $y_{\text{short}}(k)$, $y_{\text{long}}(k)$ of *y (k)* are both exponentially weighted sums of signal values measured up to time (k – 1), such that recent values have a larger weighting than that indistinct past.

$$y_{short}(k) = \beta_S\, y_{short}(k-1) + (1-\beta_S)y(k-1) \tag{4}$$

$$y_{long}(k) = \beta_L\, y_{long}(k-1) + (1-\beta_L)y(k-1) \tag{5}$$

According to Stone's method $\beta_s, \beta_L \in [0,1]$ are two different parameters, $y_{long}(1) = y_{short}(1) = y(1)$. Half-life h_L of β_L is much longer (typically 100 times longer) than corresponding half-life h_S of β_S , and the relation is:

$$\beta = 2^{-1/h} \tag{6}$$

Stone has offered TP of y_i for ith extracted signal by separator vector $\boldsymbol{w_i}$ as Rayleigh's quotient as follows:

$$F(y_i) = \log\frac{\mathbf{w_i}\mathbf{C}_{\mathbf{xx}}^{\text{long}}\mathbf{w_i^T}}{\mathbf{w_i}\mathbf{C}_{\mathbf{xx}}^{\text{short}}\mathbf{w_i^T}} \tag{7}$$

Where $\mathbf{C}_{\mathbf{xx}}^{\text{long}}$ and $\mathbf{C}_{\mathbf{xx}}^{\text{short}}$ are covariance matrices of error signals of predictions of mixed by long-term and short-term predictors, respectively. Stone BSS aims to maximize Rayleigh's quotient to yield un-mixing vectors. Later, generalized eigenvectors of $\mathbf{C}_{\mathbf{xx}}^{\text{long}}\,[\mathbf{C}_{\mathbf{xx}}^{\text{short}}]^{-1}$ are considered as un-mixing vectors in Stone's BSS [8, 14].

B. Mao's Temporal Predictability Measures

In the same strategy of Stone's measure as mentioned above, Mao's measure (M-Measure) for TP signal is [10]:

$$F(y) = \frac{1}{N}\sum_{k=1}^{N}(y_{long}(k) - y(k))^2 - \frac{1}{N}\sum_{k=1}^{N}(y_{short}(k) - y(k))^2 \tag{8}$$

For signal series y (k) with zero-mean, the covariance C difference is defined as:

$$\boldsymbol{R}_y = C(f_y(k), f_y(k)) - C\left(g_y(k), g_y(k)\right) \tag{9}$$

Where $f_y(k) = y(k) - y_{long}(k)$, $g_y(k) = y(k) - y_{short}(k)$

R_y implies the difference measure in the mean value sense. In this measure, BSS problem is changed to standard symmetric eigenproblem, and separation matrix is orthogonal.

III. Proposed Algorithm

Proposed approach depends on the integration between Stone's measure and Fast Genetic Algorithm evolutionary technique, which is called (S'M-FGA) approach. FGA has lots of improvements about population, selection, crossover and mutation in comparison with simple GA [15]. Reference [14] was interpreted how Stone's BSS deploys generalized eigenvalue decomposition to obtain the un-mixing matrix; proposed theory is based on the responses of two different linear scalar filters to the same set of signals. Indeed, the response of linear filter to signals is a comprehensive case which includes short-term and long-term linear predictors used by Stone's BSS too. Linear filters are assumed as scalar filters rather than matrix filters unless stated otherwise. Fig.2 shows a schematic diagram for the theoretical foundation of proposed approach.

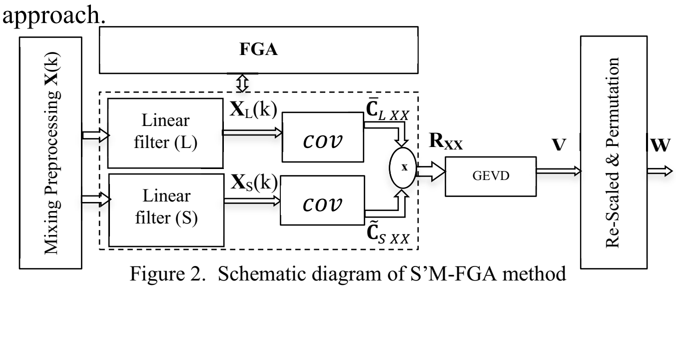

Figure 2. Schematic diagram of S'M-FGA method

Where that:
$\mathbf{X}$ (k) = Mixture observation signals
$\mathbf{X}_L$ (k) =Filter Response (L)
$\mathbf{X}_S$ (k) =Filter Response (S)
$\overline{\boldsymbol{C}}_{L\,XX}$ = Long-term covariance matrix
$\widetilde{\boldsymbol{C}}_{S\,XX}$ = Short-term covariance matrix
$\mathbf{R}_{XX} = \overline{\boldsymbol{C}}_{L\,XX}\widetilde{\boldsymbol{C}}_{S\,XX}$
V= Eigenvector matrix $\mathbf{R}_{XX}\mathbf{V}=\mathbf{VD}$
W=Un-mixing matrix

The preprocessing of BSS consists of (Centering and Whitening), from the most basic and necessary preprocessing is to center the received observation mixture of signals x, i.e. subtracts its mean vector $m = E\,[x]$ so as to make x zero-mean. The centering process should be implemented because kurtosis basically obtained as [5]:

$$Kurt(x) = \frac{E\{x^4\} - 3(E\{x^2\})^2}{3(E\{x^2\})^2} + \frac{12(E\{x\})^2E\{x^2\}}{(E\{x^2\})^2} - \frac{4E\{x\}E\{x^3\} + 6(E\{x\})^4}{(E\{x^2\})^2} \quad (10)$$

Assumption of mixed data centering, makes easy to find kurtosis. So, the kurtosis can be computed as:

$$Kurt(x) = \frac{E\{x^4\}}{(E\{x^2\})^2} - 3 \quad (11)$$

Another useful preprocessing strategy called whitening; the observed vector x linearly transforms to obtain a new vector $\widetilde{x}$ which is white, i.e. its components are uncorrelated with unity variances or covariance matrix of $\widetilde{x}$ equals the identity matrix [6]:

$$E[\tilde{x}\tilde{x}^T] = \mathrm{I} \quad (12)$$

Eigenvalue decomposition (EVD) of the covariance matrix is a popular method for whitening. From "(1)", vector of n mixture $\mathbf{X}(k) = [x_1(k), \ldots x_n(k)]^T$ has been received by n sensors. BSS considers the best estimate of un-mixing matrix $W_{n\times n}$ in order to estimate n unknown sources as mentioned in "(2)", $\mathbf{Y}(k) = [y_1(k), y_2(k), \ldots, y_n(k)]^T$. X_L (k) and X_S (k) are respectively responses of two different linear filters L and S to estimated (recovered) the signals by **W**. From [14] some plausible assumptions and properties refer to estimate **W** such as:

- ***Assumption 1:*** Mixing matrix is full column rank.
- ***Assumption 2***: Sources are mutually uncorrelated and autocorrelation functions are not equals.
- ***Assumption 3***: Responses of sources to first filter (L) are not the same from their responses to the second filter (S).
- ***Assumption 4***: Unmixing matrix is orthogonal separating.
- ***Property 1:*** If sources signals are mutually uncorrelated then response of linear filter are also mutually uncorrelated and covariance matrix of the response signals is a diagonal matrix
- ***Property 2:*** If $\bar{y}(k)$ and $\bar{x}(k)$ are respectively responses of a linear filter to $y(k)$ and $x(k)$.

$$\bar{y}(k) = \mathbf{W}\bar{x} \quad (13)$$

$$\boldsymbol{C}_{\bar{y}\bar{y}} = \mathbf{W}\boldsymbol{C}_{\bar{X}\bar{X}}\mathbf{W}^{\mathrm{T}} \quad (14)$$

The covariance matrices of X_L(k) and X_S(k) are diagonal matrices because the source signals S (k) are mutually uncorrelated (Assumption 2)

$$\overline{\boldsymbol{C}}_{L\,XX} = E[\bar{X}\bar{X}^T] = diag(\mathrm{E}[\bar{X}_1\bar{X}_1], \mathrm{E}[\bar{X}_2\bar{X}_2], ..\,\mathrm{E}\,[\bar{X}_n\bar{X}_n] \quad (15)$$

$$\widetilde{\boldsymbol{C}}_{S\,XX} = E[\tilde{X}\tilde{X}^T] = diag(\mathrm{E}[\tilde{X}_1\tilde{X}_1], \mathrm{E}[\tilde{X}_2\tilde{X}_2], ..\,\mathrm{E}\,[\tilde{X}_n\tilde{X}_n] \quad (16)$$

As shown $\bar{C}_{L\,XX}$ and $\tilde{C}_{S\,XX}$ are distinct diagonal matrices their multiplication $\boldsymbol{R}_{XX}$ is a diagonal matrix;

$$\boldsymbol{R}_{XX} = \overline{\boldsymbol{C}}_{L\,XX}\widetilde{\boldsymbol{C}}_{S\,XX} = diag(\mathrm{E}[\bar{X}_1\bar{X}_1]\mathrm{E}[\tilde{X}_1\tilde{X}_1], \mathrm{E}[\bar{X}_2\bar{X}_2]\mathrm{E}[\tilde{X}_2\tilde{X}_2], ..\,\mathrm{E}\,[\bar{X}_n\bar{X}_n]\mathrm{E}\,[\tilde{X}_n\tilde{X}_n] \quad (17)$$

Or can be said:

$$\boldsymbol{R}_{yy} = \overline{\boldsymbol{C}}_{L\,yy}\widetilde{\boldsymbol{C}}_{S\,yy} = diag(\mathrm{E}[\bar{y}_1\bar{y}_1]\mathrm{E}[\tilde{y}_1\tilde{y}_1], \mathrm{E}[\bar{y}_2\bar{y}_2]\mathrm{E}[\tilde{y}_2\tilde{y}_2], \ldots..\,\mathrm{E}\,[\bar{y}_n\bar{y}_n]\mathrm{E}\,[\tilde{y}_n\tilde{y}_n\,] \quad (18)$$

From "(13)" and "(14)"

$$\boldsymbol{R}_{yy} = [\boldsymbol{W}\overline{\boldsymbol{C}}_{L\,XX}\boldsymbol{W}^{\boldsymbol{T}}][\boldsymbol{W}\widetilde{\boldsymbol{C}}_{S\,XX}\boldsymbol{W}^{\boldsymbol{T}}] = \boldsymbol{W}\overline{\boldsymbol{C}}_{L\,XX}[\boldsymbol{W}^{\boldsymbol{T}}\boldsymbol{W}]\widetilde{\boldsymbol{C}}_{S\,XX}\boldsymbol{W}^{\boldsymbol{T}} \quad (19)$$

Now, can be representing the problem as generalized eigenvalue decomposition [8] and from the assumption "(4)" [14].

$$\mathbf{W}^{\mathbf{T}}\mathbf{W} = \mathbf{I} \quad (20)$$

$$\overline{\boldsymbol{C}}_{L\,XX}\widetilde{\boldsymbol{C}}_{S\,XX} = \mathbf{W}^{-1}\mathbf{R}_{yy}\mathbf{W} \quad (21)$$

Then, the un-mixing matrix W is organized of the eigenvector matrix of "(19)", also they are orthogonal. The FGA evolutionary technique used here for generating and tuning accurate optimum values of Half-life parameters (h_L , h_S) instead of fixed as referees in Stone's measure (S-Measure), subsequently these values are being affected on the responses of two linear filters . Since two employed linear filters are error terms of short-length and long-length prediction. According to FGA, initial population can be adapted into new population that independence among its components is maximized using the suitable fitness function. FGA intrinsically has advantages to jump out of potential local minimum and fast algorithm with operators:

- Maximum number of Generation (MaxGen) = 40.
- Population Size (Pop) = 80.
- Length of Chromosome (number of genes) Len= 2.
- Probability of Crossover P_C= 0.95.
- Probability of Mutation P_M= 0.05.
- The fitness function Fit:

$$Fit(y) = -\sum_{i=1}^{n}[|E\{y_i^4\} - 3E^2\{y_i^2\}| + H(y_i) - H(y)] \quad (22)$$

That H is the entropy of mixed signals. The entropy is always non-negative, and zero if and only if the variables are statistically independent. The fitness function, which proposed here takes the same fashion in [5] based on the fusion of two criteria, kurtosis and mutual information. When Fitness function is maximized the dependence among the estimated signals is minimized, for more details view [5]. It is not necessary to assume that the sources have the same sign of kurtosis; the absolute of fitness function is directly maximized. So Supergaussian and Subgaussian signals can be separated from each other. The algorithm would be imperfect without the orthogonal process, since the estimate of short-term half-life (h_S) and long-term half-life (h_L) using maximization of fitness function to extract original signals is not enough, and the orthogonal separating matrix can be obtained by "(12)". Orthogonalization is applied to FGA before fitting each population, as shown in the structure of the proposed evolutionary algorithm (Fig.3). In order to check the effectiveness of the proposed algorithm two performance indexes used:

-Interference to signal ratio (ISR) in dB as:

$$ISR_i = 10log\frac{E[\,(S_i(k) - y_i(k))^2]}{E[(S_i(k))^2]} \quad (23)$$

- Integral Square Error (ISE):

$$ISE = \sum_{k=0}^{T} e^2 \;=\; \sum_{k=0}^{T} (S_i(k) - y_i(k))^2 \quad (24)$$

That e is the error signal for T samples.

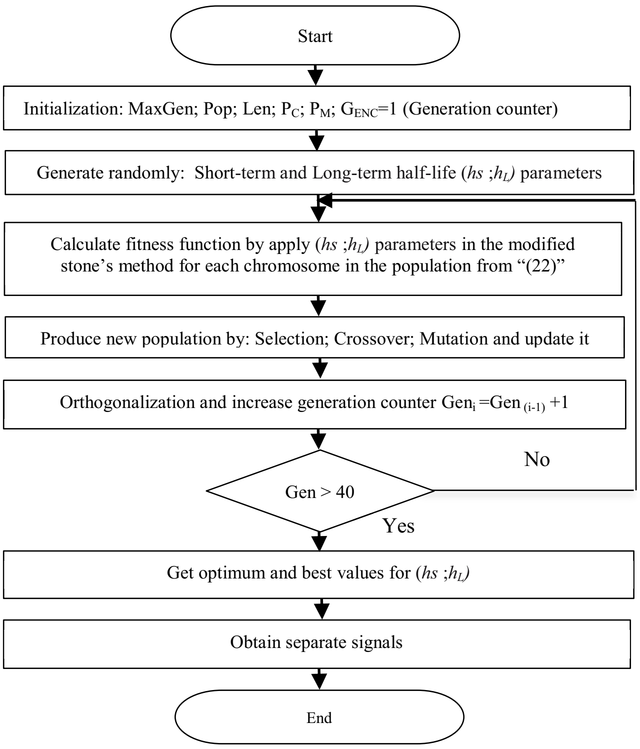

Figure 3. Flowchart of for S'M-FGA Approach

IV. SIMULATION RESULT

In order to illustrate experimental demonstration of validity of S'M-FGA approach, simulation used same experimental data in [8]. Three sources with different pdf (Supergaussian s_1 (k), Subgaussian s_2 (k), and sorted Gaussian noise s_3 (k) are generated randomly, and mixed by random matrix **A**. For more details about these signals, view [8]. The mixture tested by JADE algorithm based on [7], fastICA based on [2], Stone's approach based on [8], and the proposed approach, for having better evaluation, performances and compare the result, as shown in Table (I). Short-term and long-term half-life (*h_S*, *h_L)* parameters are generated and selected from FGA algorithm as shown in the flowchart for the proposed approach. Fig.4, shows three sources and corresponding recovered signals using S'M-FGA approach with shifted vertically for display purposes. Finally proposed algorithm was tested with eight voices mixed by random matrix **A** and proven that the recovered signals are very near to original signals as shown in Fig.5. In this work, it's impossible to recognize a difference between original and recovered signals.

TABLE I. PERFORMANCE INDEX COMPRESSION

BSS Algorithms		**Performance Index P.I**	
		ISR	***ISE***
JADE	S_1	-26.2113	7.1778
	S_2	-31.6204	2.0657
	S_3	-22.1245	18.3936
	Mean	**-26.6521**	**9.2124**
FICA	S_1	-54.0022	0.0119
	S_2	-35.7753	0.7936
	S_3	-25.9105	7.6925
	Mean	**-38.5627**	**2.8327**
Ston's-BSS	S_1	-30.394	2.74
	S_2	-61.484	0.0021
	S_3	-44.029	0.1186
	Mean	**-45.302**	**0.95358**
S'M-FGA	S_1	-30.8749	2.4526
	S_2	-79.1398	0.0000
	S_3	-44.3759	0.1095
	Mean	**-51.4635**	**0.8541**

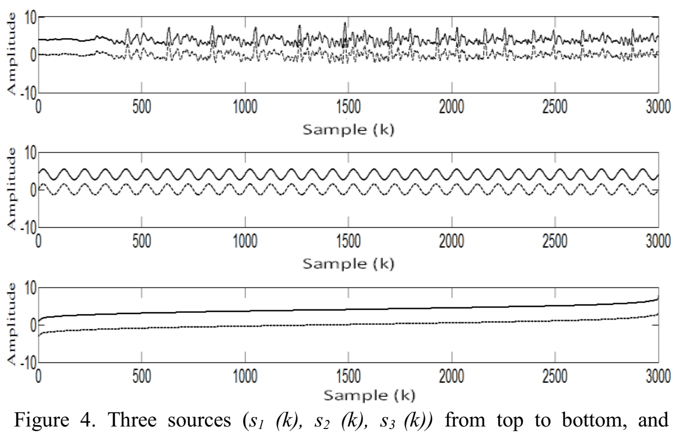

Figure 4. Three sources (*s_1 (k), s_2 (k), s_3 (k))* from top to bottom, and corresponding recovered signal (*y_1 (k), y_2 (k), y_3 (k)*) using proposed approach (S'M-FGA) with shifted vertically for display purposes.

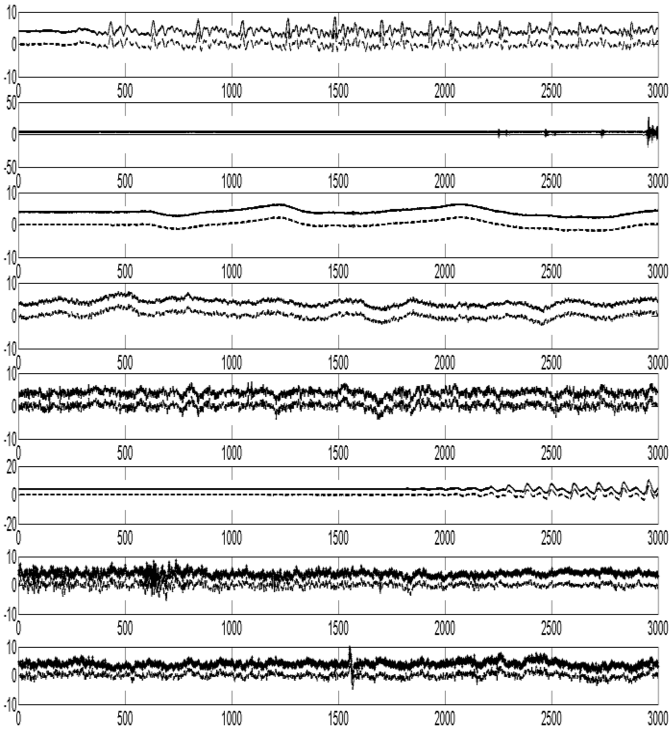

Figure 5. Eight voices and corresponding recovered signals, for 3000 samples, y-axis displays signal amplitude, using proposed approach (S'M-FGA) with shifted vertically for display purposes.

V. CONCLUSION

In this work, novel BSS approach represented by a merging intelligent evolutionary algorithm with signal temporal predictability measure, based on the difference between the temporal predictability of signals, then can be modified Stone's approach, the proposed approach has opened a new field for Stone's approach. Interpretation based on the responses of two different linear scalar filters to the same set of signals, which indicate to short-term and long-term linear predictor is not new but tuning half-life values (h_L, h_S) by fast genetic algorithms is new. Usually half-life h_L value is longer than corresponding half-life h_S, but not necessarily be the same ratio which specified by Stone's method. Changes in half-life values have a significant effect on signals; therefore, adaptation algorithms are powerful to find the best or optimum values. Fast Genetic algorithm (FGA) is really a search optimization algorithm to extract optimum long and short half-life values, as shown in scheduled results. Chromosome representation is effected on the output results, since real coding for FGA is powerful from binary coding. During observation of this work, half-life values (h_L, h_S) must not be equal to each other, since when $h_{L=}$ h_S very bad response for linear filter will be obtained. Contrary of what Stone said about half-life values, proposed approach seems to be very sensitive to short-term and long-term parameters defined in "(6)". Finally; this work is a new direction to use intelligent soft computing techniques with temporal predictability of signals to solve the blind source separation problem. Furthermore, filters representation facilitates for practical cases.

ACKNOWLEDGMENT

This paper is funded by an international exchange program of Harbin Engineering University for innovation oriented talents cultivation.

REFERENCES

[1] A. Cichocki and S.-i. Amari, *Adaptive blind Signal and Image Processing: Learning Algorithms and Applications*. New York: John Wiley and Sons, 2005.

[2] A. Hyv¨arinen, "Fast and robust fixed-point algorithms for Independent Component Analysis," *IEEE Trans. Neural Netw.*,vol. 10, pp. 626-634, Jun.1999.

[3] P. Comon, **"**Independent Component Analysis, a new concept ?,**"** *Signal Processing, Elsevier,* vol.36, pp. 287-314, 1994.

[4] S. Sun, J. Zheng, and D. Wu, "Research on blind source separation based on natural gradient algorithm," *journal of airforceering university (natural science edition)*, vol. 4, pp. 50-54, Jun 2003.

[5] M. Samira and Ayatollah, "Research of Blind Signal separation with Genetic Algorithm and Particle Swarm Optimization Based on Mutual Information," *Spring Journal of Radioelectronic and communication systems,* vol. 54, pp. 315–324, 2011.

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

[7] J. F. Cardoso and A. Souloumiac, "Blind Beamforming for non Gaussian Signals," *Radar and Signal Processing, IEE Proceedings F,* vol. 140, pp. 362–370, 1993.

[8] J. V. Stone, "Blind source separation using temporal predictability," *Neural Computation*, vol. 13, pp. 1559-1574, 2001.

[9] S. L. Xie, Z. S. He, and Y. L. Fu, "A note on the Stone's conjecture of blind separation," *Neural Computation*, vol. 17, pp. 321-330, 2005.

[10] Y. Mao and L. Xue, "An efficient measure of signal temporal predictability for blind source separation," *Springer:Neural Processing Letters,* vol. 26, pp. 57-68, 2007.

[11] Z. Yang*, et al.*, "Blind Source Separation Based on Signal Temporal Predictability," presented at the Innovative Computing Information and Control, 2008. ICICIC '08. 3rd International Conference, Dalian, Liaoning, June 2008.

[12] S. Kai*, et al.*, "Approach to Nonlinear Blind Source Separation Based on Niche Genetic Algorithm," presented at the Proceedings of the Sixth International Conference on Intelligent Systems Design and Applications (ISDA'06), Oct. 2006.

[13] E. A. Gonz´alez*, et al.*, "Elitist genetic algorithm guided by higher order statistic for blind separation of digital signals," presented at the IECON 2010 - 36th Annual Conference on IEEE Industrial Electronics Society, Glendale, AZ, Nov. 2010.

[14] M. Khosravy*, et al.*, "A theoretical discussion on the foundation of Stone's blind source separation," *Springer:Signal, Image and Video Processing,* vol. 5, April 2011.

[15] X. Meng*, et al.*, "Fast Genetic Algorithms Used for PID Parameter Optimization," presented at the Proceedings of the IEEE International Conference on Automation and Logistics, Jinan,China, Aug. 2007.

[16] T. Rasheed, "Constrained blind source Separation of human brain Signals " Ph.D., Computer Engineering Kyung Hee Seoul, Korea February,2010.