Journal of ELECTRICAL ENGINEERING, VOL 72(2021), NO3, 208–212

COMMUNICATIONS

Blind audio source separation based on a new system model and the Savitzky-Golay filter

Pengfei Xu[1], Yinjie Jia[1,2*], Mingxin Jiang[2]

Blind source separation (BSS) is a research hotspot in the field of signal processing. This scheme is widely applied to separate a group of source signals from a given set of observations or mixed signals. In the present study, the Savitzky-Golay filter is applied to smooth the mixed signals, adopt a simplified cost function based on the signal to noise ratio (SNR) and obtain the demixing matrix accordingly. To this end, the generalized eigenvalue problem is solved without conventional iterative methods. It is founded that the proposed algorithm has a simple structure and can be easily implemented in diverse problems. The obtained results demonstrate the good performance of the proposed model for separating audio signals in cases with high signal to noise ratios.

K e y w o r d s: blind source separation, moving average, Savitzky-Golay smoothing, cost function, audio signal

1 Introduction

A pioneering work on the BSS was conducted by Jutten and Herault [1] in 1985. They established the BSS as a multi-dimensional signal processing method for recovering each component of the source signal. Since then, the established BSS and the research results have been widely applied in diverse fields, including the speech recognition, fault diagnosis, array signal processing [2], signal de-noising [3], image processing [4], fetal electrocardiogram [5], biomedicine and so on.

Currently, several BSS algorithms have been established to calculate a demixing matrix so that the source signal can only be separated (estimated) by the observed or mixed signals. Compared with other algorithms, the demixing matrix based on the maximum signal to noise ratio (SNR) can be obtained with no iterative operation [6,7]. The main advantage of this algorithm over the BSS is simplicity. Accordingly, the SNR algorithm is employed in the present study to calculate the demixing matrix. It has also been adopted to calculate the demixing matrix for the peak detection in the spectrum sensing [8]. Studies show that when the complex-valued maximum pseudo-signal-to-noise ratio is considered as the target function, Doppler human gesture signals caused by different motions can be separated for optimizing the demixing matrix [9].

However, the original algorithm [6] directly uses the moving average algorithm to predict the source signal, and the prediction accuracy can be further improved. Based on such a consideration, the present study intends to propose a model based on the Savitzky-Golay smoothing filter and simplified cost function for audio signals in noisy environments. It is expected to improve the separation performance.

2 Methodology

2.1 *Blind separation of noisy mixtures*

The linear model of the basic instantaneous BSS can be expressed as

$$x_i(t) = \sum_{1}^{n} a_{ij}\big(s_i(t) + v(t)\big), \tag{1}$$

where a_{ij} is a mixed coefficient. Equation (1) can be rewritten in the vector form

$$x(t) = A\big(s(t) + v(t)\big), \tag{2}$$

where $x(t) = [x_1(t), \dots, x_n(t)]^\top$ is a vector of mixed or observed signals. Meanwhile, A is an $n \times n$ mixing matrix, $s(t) = [s_1(t), \dots, s_n(t)]^\top$ is a vector of source signals, and $v(t)$ denotes the additive white Gaussian noise (AWGN) [10]. In the BSS problem, only the statistical independence of the mixed signal from the source signal is known. Moreover, each source can be recovered through its probability distribution. Assuming that W is an $n\times n$ demixing matrix or separating matrix, the general solution to the problem is

$$y(t) = Wx(t)\,, \tag{3}$$

where $y(t) = [y_1(t), \dots, y_n(t)]^\top$ is a vector of separated signals. Generally, the BSS consists of two steps. The first step is to create a cost function $F(W)$ with respect to W. When W maximizes $F(W)$ function, then the corresponding W can be used as a demixing matrix. In the second step, an effective iterative algorithm is required for solving the equation $\partial F/\partial W = 0$. The cost function in the present article is the function of the signal-noise ratio so that the demixing matrix can be obtained by resolving the generalized eigenvalue problem with no iterative operation during the optimization of the cost function.

[1] Hohai University, College of Computer and Information, Nanjing, China, * Corresponding author: jiayinjie@hhu.edu.cn, [2] Huaiyin Institute of Technology, Faculty of Electronic Information Engineering, Huaian, China

https://doi.org/10.2478/jee-2021-0029, Print (till 2015) ISSN 1335-3632, On-line ISSN 1339-309X

Table 1. The proposed algorithm

Input: The mixed signals X. Output: The demixing matrix W and the separated signal Y.
1: XS = smoothdata(X,'sgolay'); % Smooth X using a Savitzky-Golay filter.
2: (W, d)=eig(cov($X - XS$),cov(X)); % Demixing matrix W is obtained from (4).
3: Y =($X * W$)'; % Separated signal Y.

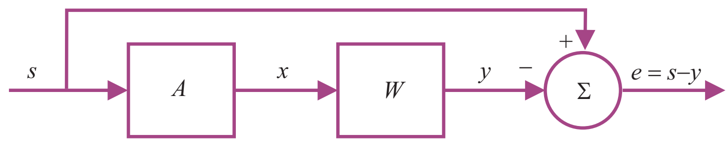

Fig. 1. Block diagram of the simplified cost function

2.2 Inversion problem

As shown in Fig. 1, the simplified cost function can be defined as

$$F_{SNR} = \frac{s\,s^\top}{e\,e^\top} = \frac{s\,s^\top}{(s-y)(s-y)^\top} = \frac{W x\,x^\top W^\top}{W(\overline{x}-x)(\overline{x}-x)^\top W^\top}, \tag{4}$$

where, $\overline{x}(n) = \frac{1}{L}\sum_{j=0}^{L} x_j(n-j)$, $0, 1, \ldots, L-1$, represents the moving average of the mixed signal $x(n)$. By calculating the partial differential equation $\partial F_{SNR}(W)/\partial W = 0$ with respect to W [7], the demixing matrix can be obtained by resolving the generalized eigenvalue problem with no iterative operation [11].

When the proposed scheme is applied to mixed signals containing white Gaussian noise described in equation (2), the moving average is replaced with Savitzky-Golay (S-G) smoothing filter [12, 13]. It is worth noting that the main purpose of applying the S-G smoothing filter is to obtain a smooth signal. Accordingly, the S-G filter is applied to a series of digital data points to improve the signal-to-noise ratio without deforming the signal. Then the linear least square method is applied to fit a subset of consecutive data points with a low order polynomial and obtain the convolution of polynomials [14]. Moreover, when the data points are arranged at fixed and uniform intervals along the selected abscissa, this filter can be applied to any consecutive data. It should be indicated that curves formed by drawing the points must be continuous and almost smooth [15].

Now, $\overline{x}(n)$ in (4) is the smoothed data generated by the Savitzky-Golay smoothing filter. In order to show the superiority of the Savitzky-Golay method over the moving average smoothing, the correlation coefficient is utilized to express the similarity between the original signal and the two smoothed signals for different window lengths and its calculation expression is shown in (6). This is presented in Fig. 2.

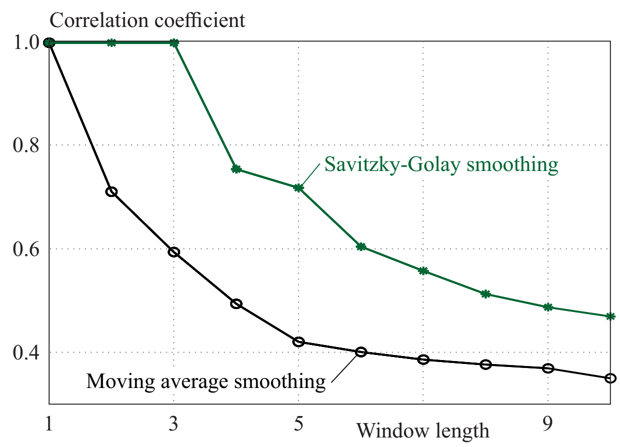

Fig. 2. Correlation coefficients between the original signal and the two smooth estimated signals for different window lengths

In Fig. 2, the abscissa and the ordinate represent the window length ranging from 1 to 10 and the correlation coefficients of the original signal and the smooth data, respectively. The lines represent the similarity (correlation coefficient) between the original signal and the Savitzky-Golay smooth signal and the correlation coefficient between the original signal and the smooth signal by moving average, respectively. Figure 2 indicates that:

(1) For different window lengths, the lower the window length, the better the similarity.

(2) When the window length increases, the similarity deteriorates.

(3) The similarity of the Savitzky-Golay method is always higher than that of the moving average method with the same window length.

Therefore, the Savitzky-Golay filter is applied in the proposed algorithm to smoothen the noisy signal $x(n)$. As a result, the estimated signal is closer to the original signal and a more reasonable source signal can be separated. Table 1 presents the proposed algorithm with three lines of the MATLAB code.

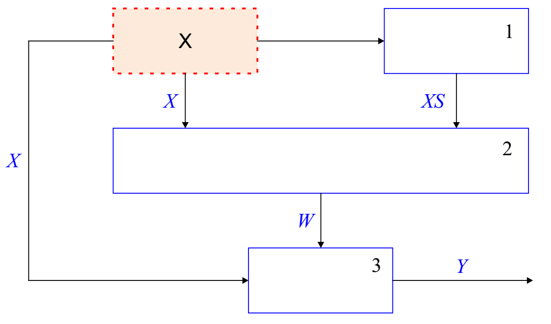

Fig. 3. Block diagram of the proposed BSS model

Figure 3 presents the block diagram of the proposed BSS model based on the foregoing algorithm. The dashed box indicate represents the mixed signals X. Note that the sequence numbers, 1, 2 and 3 denote the three steps accordingly, Tab. 1.

By constructing the cost function and estimating the demixed matrix, the closed-form solution can be found directly with no need for an iterative optimization process by the generalized eigenvalue decomposition. In other words, the original signal S can be estimated or separated only through mixed signals X without any prior knowledge. The proposed algorithm is a matrix eigenvalue decomposition method. Moreover, it has a simple structure and it is an appropriate scheme for being implemented in the FPGA hardware and real-time processing.

3 Simulations and results

$$C(s,y)=\frac{\operatorname{cov}(x,y)}{\sqrt{\operatorname{cov}(s,s)}\sqrt{\operatorname{cov}(y,y)}}\,, \tag{5}$$

where s and y denote the source signal and the separated signal, respectively. $C(s,y)=0$ indicates that s is independent of y, while s and y are fully correlated when $C(s,y)=1$. The higher the value of C is, the better the separation performance of the proposed algorithm is and vice versa.

In addition to the correlation coefficient, it is intended to utilize another common metric in this simulation, named the signal-distortion-ratio (SDR), to evaluate the separation effect.

$$SDR=10\,\log_{10}\frac{\|s_t\|^2}{\|e_i+e_n+e_a\|^2}\,, \tag{6}$$

where, s_t is a modified version of s [16], e_i, e_n and e_a denote the interferences, noise and artifacts error terms, respectively. It should be indicated that SDR is evaluated by computing energy ratios expressed in decibels (dB). The higher the value of SDR is, the better the separation performance of the proposed algorithm is and vice versa.

In the first simulation, the sources are three music tunes, including guitar.wav (Source signal-1), piano.wav (Source signal-2) and trumpet.wav (Source signal-3). For each source signal, the number of samples is $N=400000$. The AWGN channel model is the most basic noise and interference model, which is used in the signal transmission channel. Moreover, the amplitude distribution is Gaussian and the power spectral density distribution is uniform. It is worth noting that the SNR in the AWGN channel is set to 50 dB. The mixing matrix $\boldsymbol{A}$ is randomly generated. Figure 4 shows the separation results obtained by the proposed algorithm.

Figure 4 shows that the proposed algorithm can effectively separate the source signals from the three mixed signals.

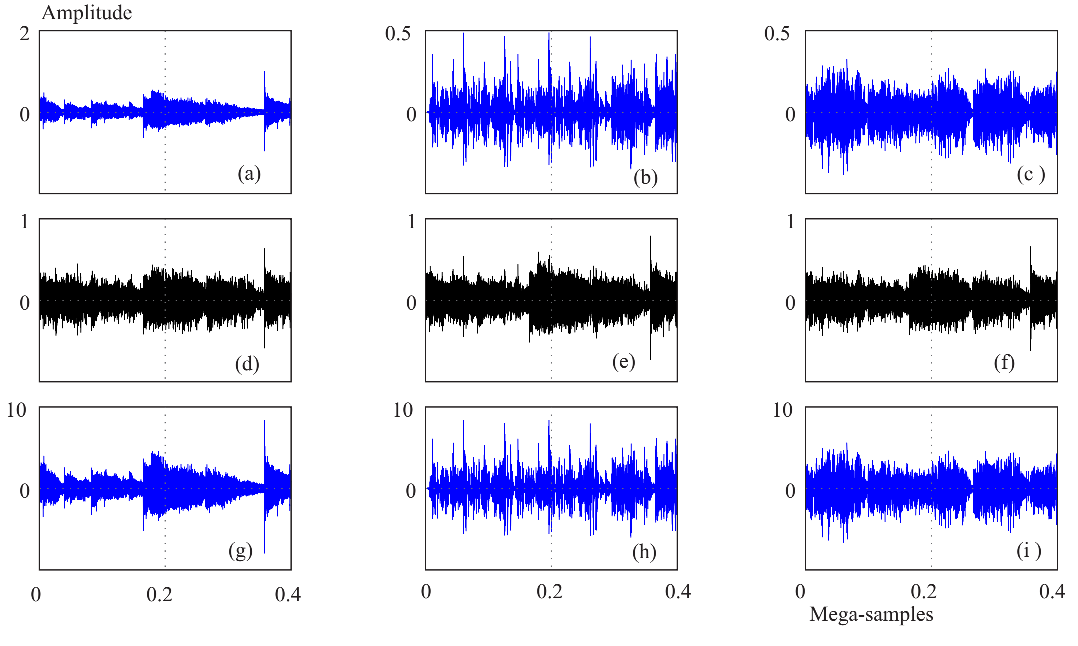

Fig. 4. Signals of the three different sources (1, 2, 3), in left to right columns: first row (a),(b),(c) – source signals, second row (d), (e), (f) - mixed signals, and third row (g), (h), (i) - separated signals

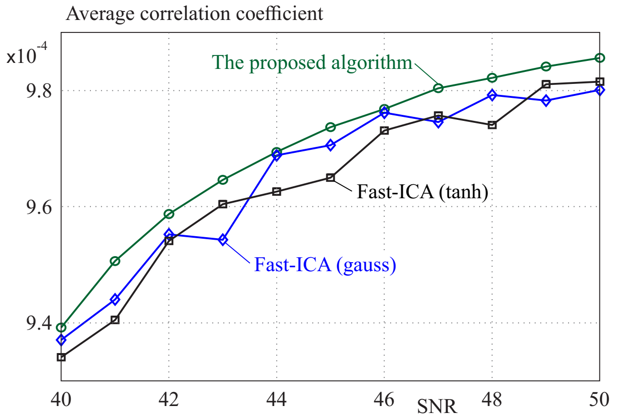

Fig. 5. Distribution of Average correlation coefficient for different algorithms

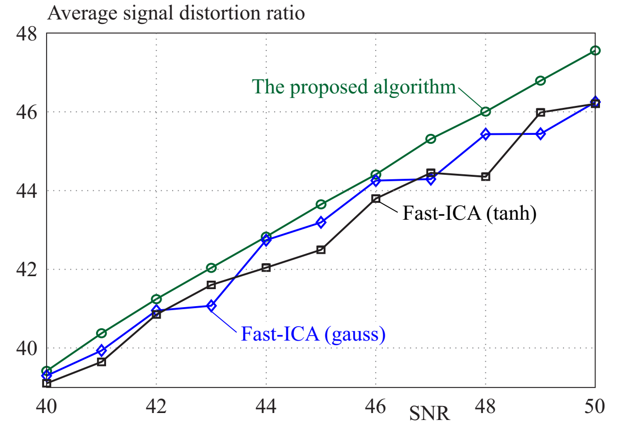

Fig. 6. Distribution of the average signal distortion ratio for different algorithms

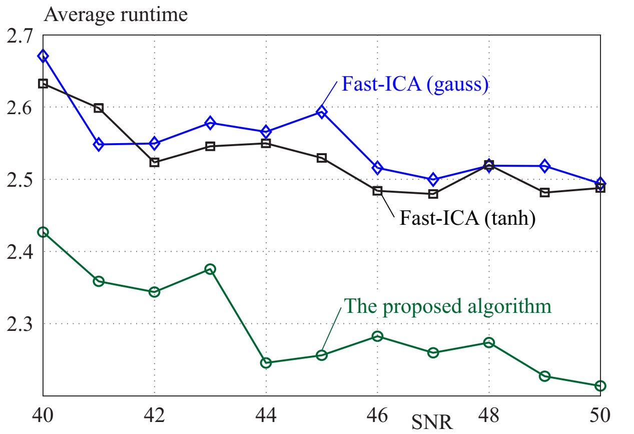

Fig. 7. Distribution of the Average runtime for different algorithms

In the second simulation, in order to investigate the advantages of the proposed algorithm, the widely adopted Fast-ICA (gauss and tanh) algorithm is selected to compare with the proposed algorithm under different signal-noise ratios, from 40 to 50 dB. Since the source signals are super-Gaussian signals (kurtosis (S1)=4.9107, kurtosis (S2)=6.5883, kurtosis (S3)=4.5102), so, $g(u) = ue^{-}u^2/2$ and $g(u) = \tanh(2t)$ are selected as two non-linear functions in the Fast-ICA algorithm, respectively.

The simulation based on the same mixed signal can make the comparison between the three algorithms as fair as possible. It is worth noting that the SNR is increased from 40 dB to 50 dB on the basis of the abovementioned simulation, while other conditions remain unchanged.

It should be indicated that repeated tests can reduce the randomness and improve the reliability of results. Therefore, the simulation runs repeatedly to check its stability and obtain the average correlation coefficients, average signal distortion ratio and average runtime. Total number of iterations in this simulation is set to 100. Figure 5 shows a comparison between the results of the average correlation coefficient.

The comparison results of the average signal distortion ratio are shown in Fig. 6. Moreover, Fig. 7 illustrates obtained results of the average runtime from different algorithms.

Figures 5-7 demonstrate that the proposed algorithm outperforms the conventional schemes from the separation performance (separation accuracy) and computational speed viewpoints.

4 Conclusions

In the present study, a novel model is proposed to calculate the demixed matrix from the mixed signals without any prior condition. The proposed algorithm is assisted by the Savitzky-Golay smoothing filter and simplified cost function. The simulation results demonstrate the reasonable separation performance of the new scheme for audio signals. It should be indicated that due to the simple and clear principle of the proposed algorithm, the BSS model based on it may be applied in other applications of digital signal processing.

Acknowledgements

This work was partially supported by the project of industrial-academic-research cooperation of Jiangsu province (No BY2020241), qing lan project of Jiangsu province and opening project of Henan engineering laboratory of photoelectric sensor and intelligent measurement and control (No HELPSIMC-2020-002).

References

[1] J. Herault, C. Jutten, and B. Ans, "Détection de grandeurs primitives dans un message composite par une architecture de calcul neuromimetique en apprentissage non supervise", *Proceedings of the 10 th Symposium on Signal and Image Processing*, Nice, France, pp. 1017-1022 (1985).

[2] P. Xu, Y. Jia, Z. Wang, and M. Jiang, "Underdetermined Blind Source Separation for Sparse Signals Based on the Law of Large Numbers and Minimum Intersection Angle Rule", *Circuits Systems and Signal Processing*, 39, 5, pp. 2442-2458 (2020).

[3] Y. Jia and P. Xu, "Convolutive Blind Source Separation for Communication Signals Based on the Sliding Z-Transform", *IEEE Access*, 8, pp. 41213-41219 (2020).

[4] P. Xu and Y. Jia, "Blind source separation based on source number estimation and fast-ICA with a novel non-linear function", *Proceedings of the romanian academy series a-mathematics physics technical sciences information science*, 21, 2, 93-194 (2020).

[5] D. D. Taralunga, I. Gussi, and R. Strungaru, "A new method for fetal electrocardiogram denoising using blind source separation and empirical mode decomposition", *Rev. Roum. Sci. Techn. - lectrotechn. Et nerg.*, 61, 1, pp. 94-98 (2016).

[6] J. Ma and X. Zhang, "Blind Source Separation Algorithm Based on Maximum Signal Noise Ratio", *Proceedings of the 2008 IEEE First International Conference on Intelligent Networks and Intelligent Systems*, Wuhan, China, pp. 625-628 (2008).

[7] Z. Gu, T. Fan, Q. Lv, J. Chen, D. Ye, J. Huangfu, and L. Ran, "Remote Blind Motion Separation Using a Single-Tone SIMO Doppler Radar Sensor", *IEEE Transactions on Geoscience and Remote Sensing*, 57, 1, pp. 462-472 (2019).

[8] Z. Saleem, S. Al-Ghadhban, and T. Y. Al-Naffouri, "On the use of blind source separation for peak detection in spectrum sensing", *Proceedings of IEEE International Conference on Control System, Computing and Engineering*, Penang, Malaysia, pp. 23 - 25 (2012).

[9] Z. Gu, J. Wang, F. Shen, K. Xu, D. Ye, and J. Huangfu, "Blind Separation of Doppler Human Gesture Signals Based on Continuous-Wave Radar Sensors", *IEEE Transactions on Instrumentation and Measurement*, 68, 7, pp. 2659-2661 (2019).

[10] Y. Zhang, X. Shi, and C. H. Chen, "A Gaussian mixture model for underdetermined independent component analysis", *Signal Processing*, 86, 7, pp. 1538-1549 (2006).

[11] M. Borga, "Learning multidimensional signal processing", *PhD dissertation*, Linkoping University, Sweden (1998).

[12] A. Savitzky and M. J. E. Golay, "Smoothing and differentiation of data by simplified least squares procedures", *Analytical Chemistry*, 36, 11, pp. 1627 - 1639 (1964).

[13] R. W. Schafer, "What is a Savitzky-Golay filter?", *IEEE Signal Processing Magazine*, 28, 4, pp. 111-117 (2011).

[14] D. Acharya, A. Rani, S. Agarwal, and V. Singh, "Application of Adaptive Savitzky-Golay Filter for EEG Signal Processing", *Perspectives on Science*, 8, pp. 677-679 (2016).

[15] J. Chen, P. Jönsson, M. Tamura, Z. Gu, B. Matsushita, and L. Eklundh, "A Simple Method for Reconstructing a High-Quality NDVI Time-Series Data Set Based on the Savitzky-Golay Filter", *Remote Sensing of Environment*, 91, 3, pp. 332-344 (2004).

[16] E. Vincent, R. Gribonval, and C. Fevotte, "Performance measurement in blind audio source separation", *IEEE Transactions on Audio, Speech, and Language Processing*, 14, 4, pp. 1462-1469 (2006).

Received 17 May 2021

Pengfei Xu received the MS degree in communication and information systems from the Xidian University,China,in 2010. He is now working towards the PhD degree in computer science and technology at Hohai University. His research interests include signal processing, machine learning and compressive sensing.He authored and coauthored eleven articles. (xpf@hhu.edu.cn)

Yinjie Jia received the MS degree in communication and information systems from the Xi'an University of Science and Technology, China,in 2011. She is now working towards the PhD degree in computer science and technology at Hohai University. Her research interests include signal processing, machine learning and compressive sensing. She authored and coauthored six articles. (jiayinjie@hhu.edu.cn)

Mingxin Jiang received her BS degree in Measurement & Control Technology and Instrument and the MS degree in Communications and Information System from Jilin University, Changchun, China, in 2002 and 2005. She received a PhD degree in Signal and information processing, Dalian University of Technology, China, in 2013. She was a post-doctoral researcher with the Department of Electrical Engineering in Dalian University of Technology from 2013 to 2015. She is currently an associate professor in Faculty of Electronic information Engineering at Huaiyin Institute of Technology. Her research interests include multi-object tracking, video content analysis and vision sensors for robotics. (jmx@hyit.edu.cn)