Journal of Information and Computing Science
J. Info. Comput. Sci., Vol. 19, No. 1, pp. 53-64

DOI: 10.4208/JICS-2024-004
June 2024

Frequency Domain Acoustic Echo Cancellation Using Auxiliary Function Based Independent Component Analysis

Lifu Wu[1,2,*], Lei Wang[1], Xinnian Sun[1] and Shuaiheng Sun[1]

[1] *School of Electronics and Information Engineering, Nanjing University of Information Science and Technology, Nanjing 210044, China*

[2] *Collaborative Innovation Center of Atmospheric Environment and Equipment Technology, Nanjing University of Information Science and Technology, Nanjing 210044, China*

Abstract. The performance of traditional Acoustic Echo Cancellation (AEC) is restricted due to the double-talk detector it used to determine the double-talk and single-talk scenarios. While Blind Source Separation (BSS) signal model is a full duplex model with both far-end and near-end signals, thus the BSS-based AEC does not need the double-talk detector. This paper adopts Auxiliary function based Independent Component Analysis (Aux-ICA) algorithm to realize acoustic echo cancellation in frequency domain, in which the object function is minimizing the mutual information, and the auxiliary function technique is used for optimization. Simulation results show that this method has lower computational complexity and better performance in acoustic echo cancellation under continuous double-talk scenarios.

AMS subject classifications: 62H25, 68W40

Key words: Echo cancellation, Auxiliary function, Independent Component Analysis (ICA), Blind source separation, Double-talk

1 Introduction

In network conference, hands-free call and other applications, there are different degrees of acoustic echo problems. The existence of echo affects the quality of communication, and the communication system cannot work normally if it is serious. Therefore, effective

Translated from *Journal of Nanjing University of Information Science & Technology*, 2023, 15(3): 323-329.
*Corresponding author. *Email addresses:* **wulifu@nuist.edu.cn** (L. Wu).
©2024 by the author(s). Licensee Global Science Press. This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY) License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

measures must be taken to suppress the echo and eliminate its influence. Echo cancellation is a method usually used. Its basic idea is to estimate the echo path, get the estimate of the echo signal, subtract the estimated signal from the microphone signal, and realize the echo cancellation.

Adaptive filtering [1] is one of the common methods of acoustic echo cancellation. Normalized Least Mean Square (NLMS) algorithm [2-3] is a typical algorithm for echo cancellation, which minimizes the mean square error between the estimated echo and the microphone signal through gradient descent. To prevent filter divergence, additional use of a Double-Talk Detector (DTD) [4] or an adaptive step strategy [5] is required to slow or stop adaptive filter adjustment during double-ended talk. Recursive Least Square (RLS) [6] is also an AEC algorithm, and compared to the NLMS algorithm, the RLS algorithm has a faster convergence rate, but its computational complexity is also higher. Speex MDF [7] is a widely used adaptive filter echo cancellation algorithm. Based on NLMS algorithm, it is implemented by Multi Delay block Frequency domain (MDF) filtering algorithm, and the optimal step size estimation is derived. The advantage is that the filter coefficient is based on block update.

The AEC method mentioned above has some shortcomings. The method based on gradient descent has a balance problem between convergence speed and stability [8]. Although DTD and adaptive step strategies work well in both one-way and occasionally two-ended call scenarios, their performance can degrade in continuous two-ended call scenarios where the near-end signal is always present [9]. Blind source separation [10-11] is a technique that separates the desired signal from the observed mixed signal to achieve signal separation or enhancement. Independent Component Analysis (ICA) [12] and Independent Vector Analysis (IVA) [13] are typical BSS techniques. AEC can be thought of as a semi-blind source separation problem, where the goal is to separate the echo from the near-end signal from the microphone (microphone) signal.

In recent years, although the echo cancellation method based on Deep Learning [14-15] has shown good performance, this data-driven method has two main shortcomings: First, it needs enough data for training. Although there are some open-source audio databases, these databases are usually not enough to build robust neural networks; The second is that the parameters of the deep neural network cannot be interpreted, which is unacceptable to engineers or actual users who want to manipulate and adjust the performance of the echo cancellation system to their own needs.

Compared with the traditional AEC algorithm, because the BSS signal model is a full-duplex model with both remote and near-end signals, the BSS based AEC algorithm has better echo cancellation capability in the continuous two-end call scenario. At the same time, the excellent performance of Speex MDF algorithm shows that the implementation of AEC in frequency domain has certain advantages. Therefore, this paper adopts independent component analysis based on auxiliary function to realize acoustic echo cancellation in frequency domain. Based on full-duplex characteristics, auxiliary function technology is used to avoid explicit step parameter selection and reduce the computational complexity of the algorithm.

2 Problem Description

2.1 Signal model

Regardless of nonlinear echoes, the signal model used in this paper is shown in Figure 1. Microphone signal x consists of two parts: linear echo e and near-end signal s:

$$x = s + e. \tag{1}$$

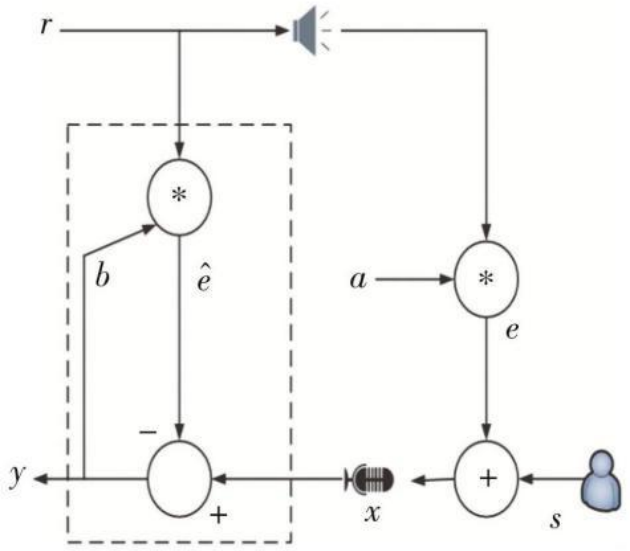

Figure 1: Signal model

If there is only one reference signal r in the system, the linear echo e can be regarded as the convolution of the reference signal r and the unknown echo path a, as shown in equation (2), where $*$ represents the linear convolution.

$$e = a * r. \tag{2}$$

As shown in the dotted box in Figure 1, AEC calculates the echo estimate $\hat{e}$ from the reference signal r and the estimated echo path b. Then subtract $\hat{e}$ from x to get the signal y, which is the estimated near-end signal. As shown in equations (3) and (4):

$$\hat{e} = b * r, \tag{3}$$

$$y = x - \hat{e}. \tag{4}$$

After the short-time Fourier transform, the time-domain signal can be converted into the frequency-domain signal model in equations (5) and (6):

$$\boldsymbol{X}(\tau) = \boldsymbol{S}(\tau) + \boldsymbol{A}(\tau)\boldsymbol{R}(\tau), \tag{5}$$

$$\boldsymbol{Y}(\tau) = \boldsymbol{X}(\tau) - \boldsymbol{B}(\tau)\boldsymbol{R}(\tau), \tag{6}$$

where τ is the frame index, and $\boldsymbol{X}, \boldsymbol{S}, \boldsymbol{R}, \boldsymbol{Y}, \boldsymbol{A}$ and $\boldsymbol{B}$ represent the microphone signal in the frequency domain, the near end signal, the reference signal, the estimated near end signal, the unknown echo path, and the estimated echo path, respectively.

2.2 BSS model

The signal model in equations (5) and (6) can also be expressed as the BSS mixing and separation model in equations (7) and (8), where L is the order of the filter and $(\ )^{\mathrm{H}}$ is the conjugate transpose.

$$\begin{bmatrix} \boldsymbol{X}(\tau) \\ \boldsymbol{R}^H(\tau) \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbf{1} & \boldsymbol{A}_1(\tau), \cdots, \boldsymbol{A}_L(\tau) \\ \mathbf{0}_{L\times 1} & \boldsymbol{I}_{L\times L} \end{bmatrix}}_{E(\tau)} \begin{bmatrix} \boldsymbol{S}(\tau) \\ \boldsymbol{R}^H(\tau) \end{bmatrix}, \tag{7}$$

$$\begin{bmatrix} \boldsymbol{Y}(\tau) \\ \boldsymbol{R}^H(\tau) \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbf{1} & [\boldsymbol{B}_1(\tau), \cdots, \boldsymbol{B}_L(\tau)] \\ \mathbf{0}_{L\times 1} & \boldsymbol{I}_{L\times L} \end{bmatrix}}_{F(\tau)} \begin{bmatrix} \boldsymbol{X}(\tau) \\ \boldsymbol{R}^H(\tau) \end{bmatrix}, \tag{8}$$

where $\mathbf{0}$ is the zero matrix, $\boldsymbol{I}$ is the identity matrix, $\boldsymbol{E}$ is the mixed matrix, and $\boldsymbol{F}$ is the separation matrix. Unlike traditional AEC algorithms, BSS can be viewed as a full-duplex model in which both remote and near-end signals coexist, with $\boldsymbol{S}$ as an independent component in the BSS signal model.

3 Aux-ICA algorithm

3.1 Algorithm derivation

According to Aux-ICA algorithm [8,16], the weighted correlation matrix $\boldsymbol{C}$ [17] is first updated by nonlinear β in equation (10), as shown in equation (9) and equation (10). Where γ is the sparse parameter, α is the forgetting factor, and δ is a small number set to prevent the denominator from being zero.

$$\boldsymbol{C}(\tau) = \alpha\boldsymbol{C}(\tau-1) + \beta(\tau)\boldsymbol{X}(\tau)\boldsymbol{X}^H(\tau), \tag{9}$$

$$\beta(\tau) = (1-\alpha)(\|\boldsymbol{Y}(\tau)\|^2 + \delta)^{(\gamma-2)/2}. \tag{10}$$

The separation filter $\boldsymbol{w} = \boldsymbol{F}^H\boldsymbol{i}_1 = [1, \boldsymbol{B}^T]^T$ can be calculated according to equations (11) and (12), where $\boldsymbol{i}_1 = [1,0,\cdots,0]^T$ is the $L+1$ dimensional vector and $(\ )^{\mathrm{T}}$ represents the transpose. Unlike the blind source separation problem, there is no ambiguity problem in BSS based AEC. Therefore, $\boldsymbol{w}$ is normalized according to equation (12) to ensure that the separation matrix $\boldsymbol{F}$ maintains the structure in equation (8).

$$\boldsymbol{w}(\tau) = [\boldsymbol{F}(\tau-1)\boldsymbol{C}(\tau)]^{-1}\boldsymbol{i}_1 = \boldsymbol{C}^{-1}(\tau)\boldsymbol{F}^{-1}(\tau-1)\boldsymbol{i}_1, \tag{11}$$

$$\boldsymbol{w} \leftarrow \boldsymbol{w}/\boldsymbol{w}_1. \tag{12}$$

As can be seen from equation (8), the separation matrix $\boldsymbol{F}$ is an upper triangular matrix with diagonal elements 1, the first-row elements non-zero, and other elements 0, so $\boldsymbol{F}^{-1}$ and the separation matrix $\boldsymbol{F}$ have the same structure. And because $\boldsymbol{i}_1 = [1,0,\cdots,0]^T$, the calculation of the matrix in equation (11) is only related to the first column elements of $\boldsymbol{C}^{-1}$ and $\boldsymbol{F}^{-1}$, so the inverse process of separating matrix $\boldsymbol{F}$ is avoided in equation (11), as shown in

$$\boldsymbol{w}(\boldsymbol{\tau}) = \boldsymbol{C}^{-1}(\boldsymbol{\tau})\boldsymbol{i}_1. \tag{13}$$

Thus, the separation filter $\boldsymbol{w}$ can be written in the form of

$$\boldsymbol{w} = [\mathbf{1}, \boldsymbol{B}^{\mathrm{T}}]^{\mathrm{T}} = \boldsymbol{C}^{-1}\boldsymbol{i}_1. \tag{14}$$

The correlation matrix $\boldsymbol{C}$ in equations (9) and (11) can be converted to the representation of equations (15) to (18), where $\boldsymbol{V}$ is the autocorrelation matrix of the reference signal and $\boldsymbol{P}$ is the cross-correlation matrix of the microphone signal and the reference signal:

$$\boldsymbol{C} = \begin{bmatrix} \boldsymbol{C}_{11} & \boldsymbol{P}^{H} \\ \boldsymbol{P} & \boldsymbol{V} \end{bmatrix}, \tag{15}$$

$$\boldsymbol{C}_{11}(\tau) = \alpha \boldsymbol{C}_{11}(\tau - 1) + \beta(\tau)\boldsymbol{X}(\tau)\boldsymbol{X}^{H}(\tau), \tag{16}$$

$$\boldsymbol{P}(\tau) = \alpha \boldsymbol{P}(\tau - 1) + \beta(\tau)\boldsymbol{R}(\tau)\boldsymbol{X}^{H}(\tau), \tag{17}$$

$$\boldsymbol{V}(\tau) = \alpha \boldsymbol{V}(\tau - 1) + \beta(\tau)\boldsymbol{R}(\tau)\boldsymbol{R}^{H}(\tau). \tag{18}$$

According to the inversion of the partitioned matrix, $\boldsymbol{C}^{-1}$ can be expressed in the form of

$$\boldsymbol{C}^{-1} = \begin{bmatrix} \frac{1}{\boldsymbol{C}_{11} - \boldsymbol{P}^{\mathrm{H}}\boldsymbol{V}^{-1}\boldsymbol{P}} & \frac{-\boldsymbol{P}^{\mathrm{H}}}{\boldsymbol{C}_{11}}\left(\boldsymbol{V} - \frac{\boldsymbol{P}\boldsymbol{P}^{\mathrm{H}}}{\boldsymbol{C}_{11}}\right) \\ \frac{-\boldsymbol{V}^{-1}\boldsymbol{P}}{\boldsymbol{C}_{11} - \boldsymbol{P}^{\mathrm{H}}\boldsymbol{V}^{-1}\boldsymbol{P}} & \left(\boldsymbol{V} - \frac{\boldsymbol{P}\boldsymbol{P}^{\mathrm{H}}}{\boldsymbol{C}_{11}}\right)^{-1} \end{bmatrix}. \tag{19}$$

Substituting equation (19) into equation (14) gives a solution as shown in

$$\begin{bmatrix} \mathbf{1} \\ \boldsymbol{B}^{\mathrm{T}} \end{bmatrix} = \begin{bmatrix} \frac{1}{\boldsymbol{C}_{11} - \boldsymbol{P}^{\mathrm{H}}\boldsymbol{V}^{-1}\boldsymbol{P}} \\ \frac{-\boldsymbol{V}^{-1}\boldsymbol{P}}{\boldsymbol{C}_{11} - \boldsymbol{P}^{\mathrm{H}}\boldsymbol{V}^{-1}\boldsymbol{P}} \end{bmatrix}. \tag{20}$$

The simplified solution shown in equation (21) below is obtained by simplifying equation (20):

$$\boldsymbol{B}(\tau) = \boldsymbol{V}^{-1}(\tau)\boldsymbol{P}(\tau). \tag{21}$$

Because the BSS separation model in equation (8) can be interpreted as adding a negative echo estimate to the microphone signal, the negative sign in equation (20) through equation (21) can be omitted, instead of subtracting it as in the traditional AEC model.

3.2 Discussion

The objective function of Aux-ICA AEC is obtained by minimizing mutual information measured by KL divergence (Kullback-Leibler divergence) [18] and optimized by auxiliary function technique. In the ICA model, the near-end signal is explicitly modeled as an independent component, with the nonlinear parameter β in ICA as the weighted value. The use of nonlinear parameter β improves the performance of speech separation. Because the BSS signal model is a full-duplex model in which both remote and near-end signals coexist, Aux-ICA AEC has good echo cancellation capability in continuous two-end call scenarios. Since equation (21) includes matrix inversion and is not suitable for online application, QRD-RLS (QR Decomposition-RLS) algorithm [19] can be used to reduce the computational complexity.

During signal processing in the frequency domain, in order to prevent large errors in the front end of the signal due to the zero matrix of the echo path of the first frame, the first frame of the microphone signal needs to be preprocessed in the simulation, that is, all points of the first frame are iterated according to the algorithm in this paper, so that

the echo path of the first frame is a non-zero matrix. The remaining frames are iterated according to the first frame. The echo elimination process of Aux-ICA AEC algorithm is shown in Table 1.

Table 1: The Aux-ICA AEC algorithm

initialize：$\boldsymbol{B}(0)=0$, $\boldsymbol{V}(0)=0$, $\boldsymbol{P}(0)=0$, $a=0.999$, $\gamma=0.2$, $\delta=1\times10^{-6}$
input：$X(\tau)$，$R(\tau)$，output：$Y(\tau)$
1）$Y(\tau)= X(\tau)-B^{H}(\tau\text{-}1)\,R(\tau)$；
2）by equation (10), β(τ) is calculated；
3）update P(τ) and V(τ) by equations (17) and (18)；
4）update B(τ) by equation (21)

4 Simulation experiment

4.1 Experimental environment

In the simulation, a 5 m × 7 m × 2.4 m room was established, and the position and direction of the sound source and microphone were randomly simulated. Room Impulse Response (RIR) is generated according to the mirror source method [20]. A set of signals is randomly selected from a dataset of 100 sound sources and reference signals with a sampling rate of 16 kHz, with a time length of approximately 24 s. The reference signal includes both music and speech types (that is, the echo signal can be either music or speech) and is always present during the experiment to simulate a continuous two-terminal call scenario. According to equations (1) and (2), the reference Signal is convolved with the simulated Echo path, and then mixed according to the Signal to Echo Radio (SER) ratio to obtain the experimental data. In this experiment, two algorithms are compared: Speex MDF and Aux-ICA AEC. Among them, the filter length of Speex MDF algorithm is set to 512, and the frame length is set to 256. In Aux-ICA AEC algorithm, the length of short-time Fourier transform is 8 192, and the frame shift is 4 096. In this paper, several experiments are conducted by setting different values of reverberation time (RT_{60}, representing the time from sudden sound stop to sound pressure level reduction of 60 dB) and signal echo ratio (SER). The algorithm running time T under the same hardware platform is used to compare the computational complexity of different algorithms, and the near-end Speech Quality after echo cancellation is evaluated by Perceptual Evaluation of Speech Quality (PESQ) [21]. The simulation parameter Settings of the experiment are shown in Table 2.

4.2 Results and discussion

Tables 3 and 4 show the algorithm running time obtained when SER values are different in the continuous two-terminal call scenario under different reverberation environments (RT_{60} = 0.2 s, 0.4 s, 0.8 s), when the echo signal is music and voice respectively.

Table 2: Simulation environment parameters setting

argument	value
Room size /m	5×7×2.4
Sampling rate /kHz	16
Frame length	8192
Frame shift	4096
RT_{60}/s	0.2、0.4、0.8
SER/dB	−10、−5、0、5、10

In Tables 3 and 4, T_{Aux} represents the running time of Aux-ICA AEC algorithm, and T_{Speex} represents the running time of Speex MDF algorithm. It can be observed from Tables 3 and 4 that in a continuous two-terminal call scenario with the same degree of reverberation, the running time of Aux-ICA AEC algorithm is about 6.87% lower than that of Speex MDF algorithm on average when the echo signal is music, and the running time of AUX-ICA AEC algorithm is about 6.87% lower than that of SPEEX MDF algorithm when the echo signal is voice. The running time of Aux-ICA AEC algorithm is about 11.91% lower than that of Speex MDF algorithm, which shows that Aux-ICA AEC algorithm has lower computational complexity. Compared with Speex MDF algorithm, Aux-ICA AEC algorithm is based on the full-duplex characteristics of BSS and does not require additional use of DTD, so Aux-ICA AEC algorithm has lower computational complexity, which is also verified by simulation experiments.

Table 3: Algorithm running time in different degrees of reverberation environment when echo signal is music

SER/dB	T_{Aux}/s			T_{Speex}/s		
	RT_{60} = 0.2 s	RT_{60} = 0.4 s	RT_{60} = 0.8 s	RT_{60} = 0.2 s	RT_{60} = 0.4 s	RT_{60} = 0.8 s
−10	12.53	13.05	13.21	13.57	13.68	14.19
−5	13.16	13.31	12.71	13.36	13.69	13.87
0	12.78	13.06	13.11	14.48	14.27	15.07
5	13.18	13.41	13.51	13.87	13.58	14.07
10	13.20	13.26	13.72	14.45	14.12	15.45

Table 4: Algorithm running time in different degrees of reverberation environment when echo signal is voice

SER/dB	T_{Aux}/s			T_{Speex}/s		
	RT_{60} = 0.2 s	RT_{60} = 0.4 s	RT_{60} = 0.8 s	RT_{60} = 0.2 s	RT_{60} = 0.4 s	RT_{60} = 0.8 s
−10	11.96	12.03	12.03	13.82	14.60	13.87
−5	12.05	13.30	13.18	14.88	15.07	14.58
0	13.32	13.26	13.34	15.02	14.01	13.82
5	11.96	12.21	13.32	13.87	15.08	13.86
10	11.94	12.22	13.22	14.95	13.83	13.97

In this paper, PESQ is used to evaluate the signal after echo cancellation in the

experiment, as shown in Figures 2-7 (Figures 2-4 is the experimental result when the echo is music, Figures 5-7 is the experimental result when the echo is speech). The higher the PESQ score, the better the echo cancellation ability of the algorithm. As can be seen from Figure 2 and Figure 5, when $RT_{60} = 0.2s$, SER=10dB, compared with $RT_{60} = 0.2s$, SER= -10dB, Aux-ICA AEC algorithm has a higher PESQ score due to the reduction of feedback amplitude. As the degree of reverberation intensifies, $RT_{60} = 0.8s$ and SER= -10dB in Figure 4. Compared with $RT_{60} = 0.2s$ and SER= -10dB in Figure 2, the echo cancellation capability of Aux-ICA AEC algorithm decreases, but its echo cancellation performance is still significantly better than that of Speex MDF algorithm.

As can be seen from the data of each group in Figures 2-7, in the same continuous two-terminal call scenario, the PESQ score of Speex MDF algorithm is lower than that of Aux-ICA AEC algorithm, which proves that Aux-ICA AEC algorithm has better echo cancellation performance in continuous two-terminal call scenario.

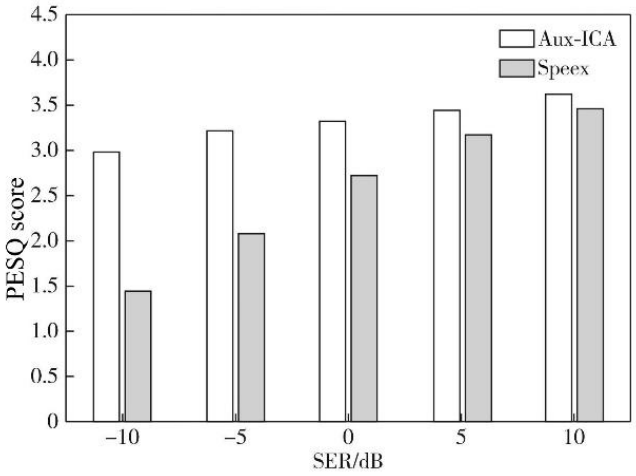

Figure 2: PESQ score of the algorithm when RT_{60} is 0.2 s (the echo signal is music)

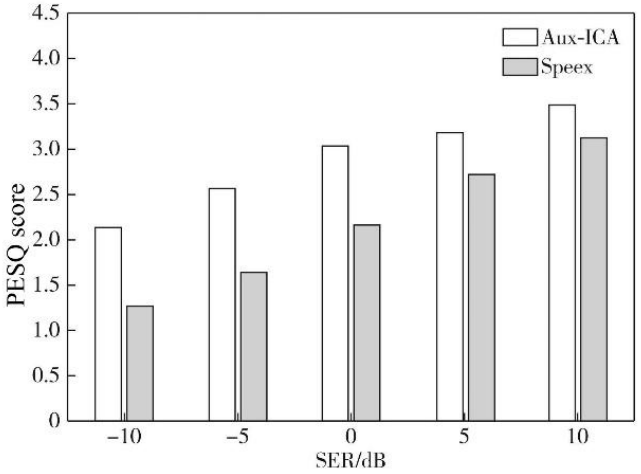

Figure 3: PESQ score of the algorithm when RT_{60} is 0.4 s (the echo signal is music)

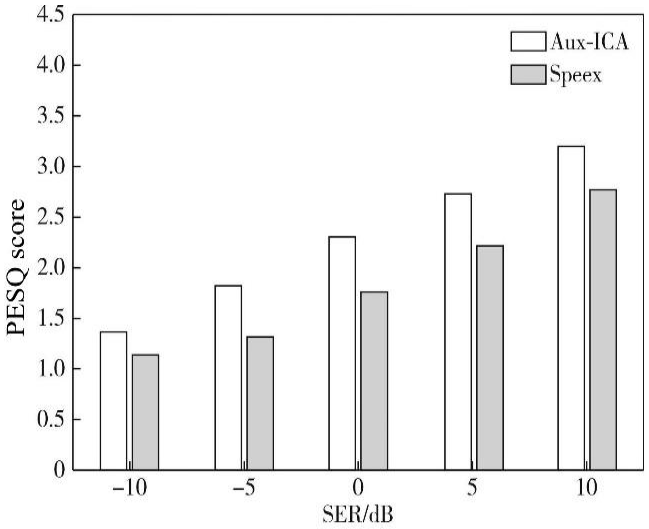

Figure 4: PESQ score of the algorithm when RT_{60} is 0.8 s (the echo signal is music)

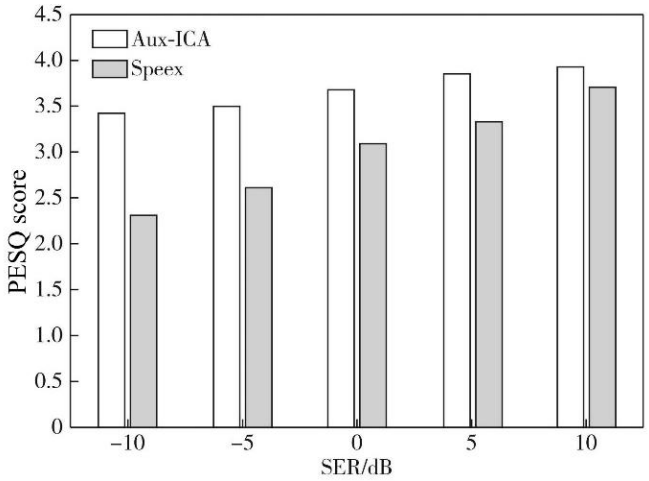

Figure 5: PESQ score of the algorithm when RT_{60} is 0.2 s (the echo signal is voice)

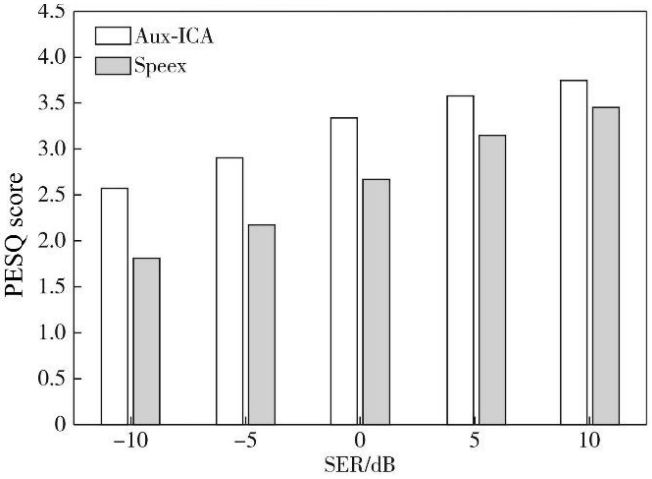

Figure 6: PESQ score of the algorithm when RT_{60} is 0.4 s (the echo signal is voice)

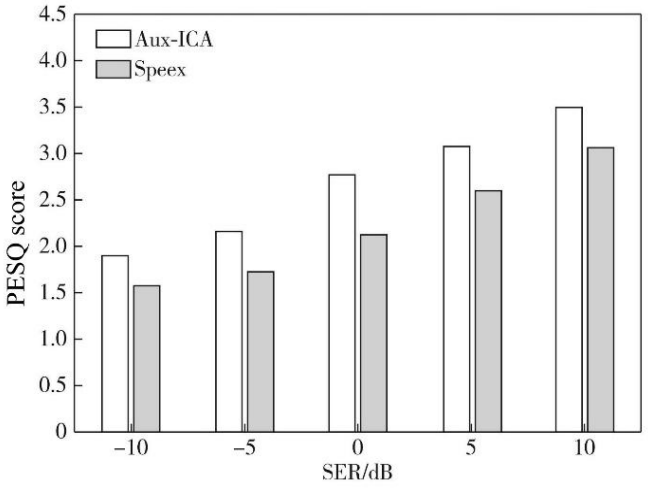

Figure 7: PESQ score of the algorithm when RT_{60} is 0.8 s (the echo signal is voice)

5 Conclusions

In this paper, an ICA algorithm based on auxiliary function is studied to realize acoustic echo cancellation in frequency domain. Based on the full-duplex feature, the explicit step parameter selection and the double-ended call detector can be omitted by using the auxiliary function technique, which reduces the computational complexity of the algorithm. Simulation results show that the proposed method has lower computational complexity and better echo cancellation performance.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (Grant No. 12074192).

Conflicts of Interest

The authors declare no conflict of interest.

References

[1] Q. Zhang, X. Wang, L. Wang, et al., Variable step-size NLMS algorithms in echo cancellation, *J. Data Acquis. Process.*, 2013, 28(1): 64-68.

[2] D. R. Morgan, and S. G. Kratzer, On a class of computationally efficient, rapidly converging,

generalized NLMS algorithms, *IEEE Signal Process. Lett.*, 1996, 3(8): 245-247.
[3] I. Albu, C. Anghel, and C. Paleologu, Adaptive filtering in acoustic echo cancellation systems: a practical overview, 2017 9th International Conference on Electronics, Computers and Artificial Intelligence (ECAI), June 29-July 1, 2017, Targoviste, Romania, IEEE, 2017: 1-6.
[4] J. M. Valin, On adjusting the learning rate in frequency domain echo cancellation with double-talk, *IEEE Trans. Audio, Speech, Lang. Process.*, 2007, 15(3): 1030-1034.
[5] J. Yang, Multilayer adaptation based complex echo cancellation and voice enhancement, *2018 IEEE International Conference on Acoustics, Speech and Signal Processing*, April 15-20, 2018, Calgary, AB, Canada, IEEE, 2018: 2131-2135.
[6] F. Ding, Auxiliary model based identification methods, Part E: least squares identification, *J. Nanjing Univ. Inf. Sci. Technol. (Nat. Sci. Ed.)*, 2016, 8(5): 385-403.
[7] G. Yang, Y. J. Wang, and J. Wang, Analysis and evaluation of echo cancellation algorithm in speex codec, *Audio Eng.*, 2013, 37(9): 52-55.
[8] N. Ono, and S. Miyabe, Auxiliary-function-based independent component analysis for super-Gaussian sources, *Latent Variable Anal. Signal Separ.*, 2010: 165-172.
[9] J. Gunther, Learning echo paths during continuous double-talk using semi-blind source separation, *IEEE Trans. Audio, Speech, Lang. Process.*, 2012, 20(2): 646-660.
[10] M. Z. Ikram, Blind source separation and acoustic echo cancellation: a unified framework, 2012 IEEE International Conference on Acoustics, Speech and Signal Processing, March 25-30, 2012, Kyoto, Japan, IEEE, 2012: 1701-1704.
[11] Y. P. Zhang, and J. Li, Optimization and simulation of blind source separation in acoustic echo cancellation, *J. Nanjing Univ. Inf. Sci. Technol. (Nat. Sci. Ed.)*, 2013, 5(6): 539-543.
[12] H. T. Shi, Q. Zhou, Y. T. Wang, et al., Fault detection method based on relative-transformation ICA, *J. Electron. Meas. Instrum.*, 2017, 31(7): 1040-1046.
[13] R. Scheibler, and N. Ono, Independent vector analysis with more microphones than sources, 2019 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, October 20-23, 2019, New Paltz, NY, USA, IEEE, 2019: 185-189.
[14] Q. H. Lei, H. Chen, J. F. Hou, et al., Deep neural network based regression approach for acoustic echo cancellation, *Proc. 2019 4th Int. Conf. Multimedia Syst. Signal Process.*, 2019: 94-98.
[15] X. Y. Zhou, K. Wang, and L. Y. Li, Review of object detection based on deep learning, *Electron. Meas. Technol.*, 2017, 40(11): 89-93.
[16] T. Taniguchi, N. Ono, A. Kawamura, et al., An auxiliary-function approach to online independent vector analysis for real-time blind source separation, 2014 4th Joint Workshop on Hands-free Speech Communication and Microphone Arrays (HSCMA), May 12-14, 2014, Villers-les-Nancy, France, IEEE, 2014: 107-111.
[17] N. Ono, Auxiliary-function-based independent vector analysis with power of vector-norm type weighting functions, *Proc. 2012 Asia Pacific Signal and Inf. Process. Assoc. Annu. Summit Conf.*, December 3-6, 2012, Hollywood, CA, USA, IEEE, 2012: 1-4.
[18] B. Z. Wei, J. Gan, and Y. L. Yin, Medical image registration based on mutual information entropy combined with edge correlation feature, *J. Data Acquis. Process.*, 2018, 33(2): 248-258.
[19] B. Jiang, J. Yang, and E. Y. Zhang, QRD-RLS algorithm for sparse adaptive Volterra filtering, *Signal Process.*, 2008, 24(4): 595-599.
[20] J. B. Allen, and D. A. Berkley, Image method for efficiently simulating small-room acoustics, *J. Acoust. Soc. Am.*, 1979, 65(4): 943-950.
[21] P. C. Loizou, Speech enhancement: theory and practice, Florida: CRC Press, 2007: 492-495.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of Global Science Press and/or the editor(s). Global Science Press and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.