Indonesian Journal of Electrical Engineering and Computer Science
Vol. 44, No. 1, October 2026, pp. 31∼40
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v44.i1.pp31-40

Microphone array post-filtering in non-stationary environment

Nguyen Thi Huyen Chau[1], Quan Trong The[2], Nguyen Duc Toan[3]
[1]Laboratory of Artificial Intelligence, Thang Long University, Hanoi, Vietnam
[2]Faculty of Information Security, Posts and Telecommunications Institute of Technology (PTIT), Hanoi, Vietnam
[3]Information Technology Institute, Vietnam Women's Academy, Hanoi, Vietnam

Article Info

Article history:

Received Jul 7, 2026
Revised Aug 29, 2026
Accepted Sep 10, 2026

Keywords:

Microphone array
Minimum variance distortionless response
Noise reduction
Speech enhancement
The signal-to-noise ratio

ABSTRACT

Microphone array (MA) has confirmed the capability of suppressing the background noise, interferences, musical and residual noise at MA beamformer's output signal. Minimum variance distortionless response (MVDR) beamformer is an efficient technique for saving the original speech component while removing the background noise with high spatial directivity index and speech quality. However, in realistic recording situation, due to the existence of non-directional noise, the complex and annoying environment, the remained noise component at MVDR beamformer's output signal degrades the speech intelligibility and perceptual metric listener. In this article, the author proposed an effective post-Filtering for removing the unwanted residual or transient noise and increasing the speech quality in the term of signal-to-noise ratio (SNR) ratio in realistic scenario. This method based on the modified the covariance of noise field by considering the speech presence probability and taking the difference between auto power spectral densities between two observed MA signals for generating post-filtering. The demonstrated experiment has shown the improvement of reducing background noise to 10.9 dB and increasing the SNR from 8.5 to 12.3 dB. The promising result has proven that the author's approach can be integrated into multi-channel system for addressing other complicated problems.

This is an open access article under the CC BY-SA license.

Corresponding Author:

Quan Trong The
Faculty of Information Security, Posts and Telecommunications Institute of Technology (PTIT)
Hanoi, Vietnam
Email: theqt@ptit.edu.vn

1. INTRODUCTION

Nowadays, the using of microphone array (MA) technology [1]-[4] has been widely popular due to its capability of enhancing extracting the desired target speech component and eliminating background noise. Due to there are many types of noise source, such as, coherent/incoherent, diffuse noise field, complex and annoying environment and different noise sources, such as, the sound of television, fan, transport vehicle, the baby cries, the third - party talker, the ability of noise reduction of single-channel approach is degraded. Single-channel method usually based on the spectral subtraction, therefore, in rapid changed environment,the estimation of covariance of noise component is not exactly and spectral subtraction causes the speech distortion.

The prior designed MA configuration, the characteristic of surrounding noise, the preferred impinging angle of useful signal, the analyzed observed MA signals - based MA beamforming. Beamformer adaptively steers the impulse response at specified sound location while attenuating the interference, non-directional, background noise and signals from other directions.

Minimum variance distortionless response (MVDR) beamformer has been attractive more scholars and engineering to study the properties of recording scenario, the advantage of designed MA geometry to obtain an necessary parameter and implement an efficient algorithm of signal processing to mitigate surrounding noise. The promising benefit of using MA beamforming is preserving the clean speech data and suppressing the environmental surrounding situation. However, in real-life situation, because of the error of overlap, the imprecise frequency sampling, inaccurate estimation of transfer function, the error of configured recording speech component, the incorrect calculation of covariance matrix of received array signals, MVDR beamformer's usually corrupted. The unwanted consequence is the unaccepted remained residual noise and distortion at the beamformer's output signal.

There are many efforts to overcome this drawback, which based on the exploiting the speech presence probability [5]-[9], the coherence of MA signals [10]-[12], the characteristic of situation [13]-[16], the properties of received signals or using neural network model [17]-[20], for extracting the distinct features for enhancing the overall evaluation in removing unwanted residual noise and surrounding acoustic factors.

However, these approaches can not deal the realistic problems with low computation and low complexity in real-life applications. In this article, the author proposed an efficient post-filtering for enhancing the MVDR beamformer's evaluation under complex and annoying environment.

2. MVDR BEAMFORMER

In this section, the author will present the working scheme of MVDR beamformer under realistic recording environment in the frequency domain. In general case, the author used dual-microphone system (DMA2) for illustrating the representation of observed MA signals on two microphones.

At current considered frame k, frequency f, the expression of MA signals, $X_1(f,k)$, $X_2(f,k)$ can be formulated as,

$$X_1(f,k) = S(f,k)e^{j\Phi_s} + N_1(f,k) \tag{1}$$

$$X_2(f,k) = S(f,k)e^{-j\Phi_s} + N_2(f,k) \tag{2}$$

where $N_1(f,k)$, $N_2(f,k)$ are annoying situations, interference or third-party talker, undetermined noise sources; $S(f,k)$ is the clean speech data of original speaker; $\Phi_s = \pi f \tau_0 cos(\theta_s)$, θ_s is the preferred impinging angle of interest useful signal to the configured MA geometry, $\tau_0 = \frac{d}{c}$, d is the range between microphones, and $c = 343(\frac{m}{s})$ is the sound speed propagation in the fresh air.

With defined symbols, $\boldsymbol{N}(f,k) = [N_1(f,k) \quad N_2(f,k)]^T$, $\boldsymbol{X}(f,k) = [X_1(f,k) \quad X_2(f,k)]^T$ and steering vector $\boldsymbol{D}_s(f,\theta_s) = [e^{j\Phi_s} \quad e^{-j\Phi_s}]^T$, the expression of (1-2) can be rewritten as,

$$\boldsymbol{X}(f,k) = S(f,k)\boldsymbol{D}_s(f,\theta_s) + \boldsymbol{N}(f,k) \tag{3}$$

The most importance of almost acoustic equipment is recovering the speech component $\hat{S}(f,k)$ as approximately the original talker $S(f,k)$. The technique requires an efficient coefficients $\boldsymbol{W}(f,k)$, which satisfied the criteria $\hat{S}(f,k) = \boldsymbol{W}^H(f,k)\boldsymbol{X}(f,k)$.

In numerous MA beamformer, the problem of calculating the filter based on the constrained criteria. MVDR beamformer aims at preserving the clean speech component while attenuating the background noise field with minimizing the total noise output power. MVDR beamformer has the high spatial diversity of concerning the beampattern the specified direction. The constrained formulation of MVDR beamformer can be presented as the following way, The value of pattern at assumed angle θ_s means saving the clean speech component. The coefficients of beamoformer yields as,

$$\boldsymbol{W}(f,k) = \frac{\boldsymbol{\Phi}_{NN}^{-1}(f,k)\boldsymbol{D}_s(f,\theta_s)}{\boldsymbol{D}_s^H(f,\theta_s)\boldsymbol{\Phi}_{NN}^{-1}(f,k)\boldsymbol{D}_s(f,\theta_s)} \tag{4}$$

where $\boldsymbol{\Phi_{NN}}(f,k) = E\{\boldsymbol{N}^H(f,k)\boldsymbol{N}(f,k)\}$ is the covariance matrix of noise.

Under real-life scenario, due to the existence of internal electrical noise of equipments, the complex recording situation, the effect of different surrounding acoustic factors or heterogeneous environment, the computation of noise still difficult challenge in almost acoustic devices. Therefore, the using of received array signals $\Phi_{XX}(f,k)$ is applied for calculating the optimum MVDR beamformer.

The final formulation of MVDR beamformer is determined as,

$$\boldsymbol{W}(f,k) = \frac{\boldsymbol{\Phi}_{XX}^{-1}(f,k)\boldsymbol{D}_s(f,\theta_s)}{\boldsymbol{D}_s^H(f,\theta_s)\boldsymbol{\Phi}_{XX}^{-1}(f,k)\boldsymbol{D}_s(f,\theta_s)} \tag{5}$$

MVDR beamformer works properly in stationary and coherent noise field. However, because of the moving head of talker, which causes the error of steering vector, the presence of different types of noise, the significant effect of the annoying transport vehicle, washing machine, television, fan, the error of sampling rate, the difference of microphone sensitivities, the resulting signal of MVDR beamformer usually degraded. The speech distortion, musical noise or large ambient noise decrease the speech quality or perceptual metric listener.

In the next section, the paper will introduce an efficient post-filtering for addressing the problem of unacceptable residual noise at beamformer's output.

3. THE AUTHOR'S PROPOSED METHOD

The author's idea is computing the variance of speech $\sigma_s^2(f,k)$ and noise $\sigma_{nMVDR}^2(f,k)$ after applying MVDR beamformer. The post-filtering has the form,

$$H_{proPF}(f,k) = \frac{\sigma_s^2(f,k)}{\sigma_s^2(f,k) + \sigma_{nMVDR}^2(f,k)} \tag{6}$$

And the MVDR beamformer's output signal is filtered as the following equation,

$$\hat{y}_{MVDR}(f,k) = y_{MVDR}(f,k) \times H_{proPF}(f,k) \tag{7}$$

At the first stage, the paper illustrated how to measure variance of noise $sigma_{nMVDR}^2(f,k)$ after implementing MVDR beamformer.

$$\sigma_{nMVDR}^2(f,k) = \boldsymbol{W}_{MVDR}^H(f,k)\boldsymbol{\Phi}_{NN}(f,k)\boldsymbol{W}_{MVDR}(f,k) \tag{8}$$

Where $\Phi_{NN}(f,k)$ mean the noisy matrix covariance. And,

$$\boldsymbol{\Phi}_{NN}(f,k) = \sigma_n^2(f,k)\begin{bmatrix} 1 & \Gamma_{N_1N_2}(f,k) \\ \Gamma_{N_2N_1}(f,k) & 1 \end{bmatrix} \tag{9}$$

where $\Gamma_{N1N2}(f,k)$, $\Gamma_{X2X1}(f,k)$ mean the coherence between microphones and $\sigma_n^2(f,k)$ is original noise power, which can be estimated by minimum statistics algorithm [21] at each microphone.

In complex environment, the coherence contains characteristic of recording environments and direction of arrival of useful speaker θ_s. Under ideal conditions of diffuse noise field, the formulation of these relation can be expressed as,

$$\Gamma_{N_1N_2}(f,k) = \frac{sin(2\pi f\tau_0)}{2\pi f\tau_0}e^{j2\Phi_s} \tag{10}$$

$$\Gamma_{N_2N_1}(f,k) = \frac{sin(2\pi f\tau_0)}{2\pi f\tau_0}e^{-j2\Phi_s} \tag{11}$$

Unfortunately, the representation of diffuse noise field in the realistic recording environment is not as the same (5) and (6), according to [22], the author suggested using speech presence probability $spp(f,k)$ for illustrating accurate formulation of coherence between MA signals in adverse situation. The expression between two microphone array (MA) signals can be rewritten as,

$$\hat{\Gamma}_{N_1N_2}(f,k) = \frac{sin(2\pi f\tau_0)}{2\pi f\tau_0(1+(1-spp(f,k)\frac{\gamma_n}{P_{nn}})}e^{j2\Phi_s} \tag{12}$$

$$\hat{\Gamma}_{N_2N_1}(f,k) = \frac{sin(2\pi f\tau_0)}{2\pi f\tau_0 1+(1-spp(f,k)\frac{\gamma_n}{P_{nn}})}e^{-j2\Phi_s} \tag{13}$$

where γ_n means the uncorrelated noise component and P_{nn} is noise power floor. In our simulation, $\gamma_n = 0.03$.

Therefore, the covariance matrix of noise component can be determined as,

$$\hat{\mathbf{\Phi}}_{NN}(f,k) = \sigma_n^2(f,k) \begin{bmatrix} 1 & \hat{\Gamma}_{N_1N_2}(f,k) \\ \hat{\Gamma}_{N_2N_1}(f,k) & 1 \end{bmatrix} \tag{14}$$

So. the variance of noise $\sigma^2_{nMVDR}(f,k)$ yields as the formulation,

$$\sigma^2_{nMVDR}(f,k) = \boldsymbol{W}^H_{MVDR}(f,k)\hat{\mathbf{\Phi}}_{NN}(f,k)\boldsymbol{W}_{MVDR}(f,k) \tag{15}$$

The effectiveness of Wiener - based post-filtering depends on the precise calculation of speech's variance, $\sigma_s^2(f,k)$. In perfect scenario, we often use the assumption that speech component correlates with noise components, that $E\{S(f,k)N^*(f,k)\} = 0$. But under realistic recording situation, there is still exists small value of $E\{S(f,k)N^*(f,k)\} \neq 0$.

From (1) and (2), with assumption that $N_1(f,k) \approx N_2(f,k) \approx N(f,k)$, we can easily obtain these expressions of expectation of $E\{|X_1(f,k)|^2\}$, $E\{|X_2(f,k)|^2\}$ and $E\{X_1(f,k)X_2^*(f,k)\}$ as,

$$E\{|X_1(f,k)|^2\} = \sigma_s^2(f,k) + \sigma_n^2(f,k) + E\{S(f,k)N^*(f,k)\}e^{j\Phi_s} \tag{16}$$

$$E\{|X_2(f,k)|^2\} = \sigma_s^2(f,k) + \sigma_n^2(f,k) + E\{S(f,k)N^*(f,k)\}e^{-j\Phi_s} \tag{17}$$

$$\begin{aligned} E\{X_1(f,k)X_2^*(f,k)\} = \sigma_s^2(f,k)e^{j2\Phi_s} + \sigma_n^2(f,k) \\ + E\{S(f,k)N^*(f,k)\}e^{j\Phi_s} + E\{N(f,k)S^*(f,k)\}e^{j\Phi_s} \end{aligned} \tag{18}$$

And,

$$E\{X_1(f,k)X_2^*(f,k)\} = \sigma_s^2(f,k)e^{j2\Phi_s} + \sigma_n^2(f,k) + 2\Re\{S(f,k)N^*(f,k)\}e^{j\Phi_s} \tag{19}$$

where $\Re(.)$ is the operator takes the real part.

From (16) and (17), the uncorrelated $E\{S(f,k)N^*(f,k)\}$ yields as the equation,

$$E\{S(f,k)N^*(f,k)\} = \frac{E\{|X_1(f,k)|^2\} - E\{|X_2(f,k)|^2\}}{2cos(\Phi_s) + \epsilon} \tag{20}$$

with additive small value of ϵ. According to our heuristic knowledge, $\epsilon = 0.025$ is suitable for MA - based speech applications.

From (16)-(19), the covariance of speech $\sigma_s^2(f,k)$ can be determined as,

$$\sigma_s^2(f,k) = \frac{A(f,k)}{1 - e^{j2\Phi_s} + \beta} \tag{21}$$

where $A(f,k) = E\{|X_1(f,k)|^2\} - E\{X_1(f,k)X_2^*(f,k)\} - (E\{S(f,k)N^*(f,k)\} - 2\Re\{S(f,k)N^*(f,k)\})e^{j\Phi_s}$ and β is an additive small constant for avoiding infinite division. In the next section, $\beta = 0.001$.

From (15) and (22), the optimum post-filtering can be achieved as,

$$H_{proPF}(f,k) = \frac{\frac{A(f,k)}{1-e^{j2\Phi_s}+\beta}}{\frac{A(f,k)}{1-e^{j2\Phi_s}+\beta} + \boldsymbol{W}^H_{MVDR}(f,k)\hat{\mathbf{\Phi}}_{NN}(f,k)\boldsymbol{W}_{MVDR}(f,k)} \tag{22}$$

In the next section, the author will perform an effective experiment for illustrating the advantage of suggested approach in reducing the musical or background noise field for increasing the speech quality of MVDR beamformer's output signal.

4. EXPERIMENT AND DISCUSSION

The purpose of the experiment is verifying the advantage of the presented method in improving the speech enhancement while suppressing the background noise, residual noise and musical noise in comparison with traditional MVDR, adaptive diagonal loading (ADL) [23] technique and GSC beamformer [24]. An objective method [25] was applied for computing the speech quality between these signals. In this experiment, the author used dual-microphone system (DMA2) to capture the clean speech data in living room ($3x4x5.5(m)$) with presence of washing machine, television, fan and third-party talker. The aim of this experiment is comparison the recorded MA signals, the processed signals by conventional MVDR beamformer and the author's suggested method (asm). The scheme of experiment is shown in Figure 1.

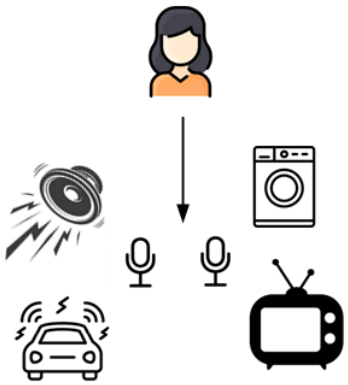
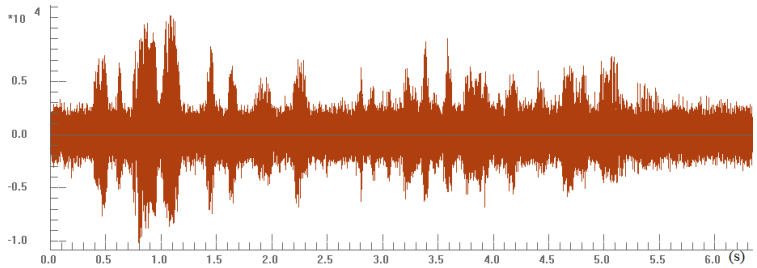
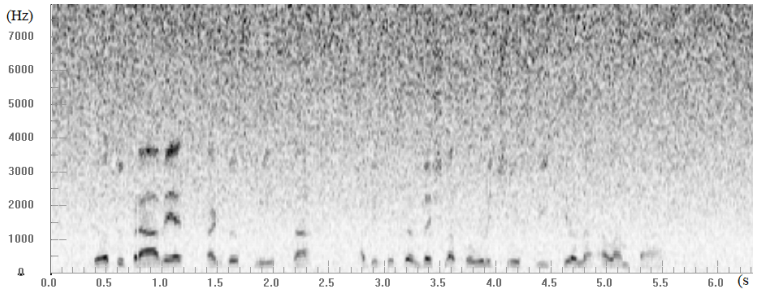


Figure 1. The demonstrated experiment in living room

The distance from a stand speaker to the axis of DMA2 is $L = 3(m)$, the range between two mounted microphones is $d = 4.35(cm)$. The preferred steering vector with impinging angle of useful signal is $\theta_s = 90(deg)$. In realistic recording environment, the captured speech component is seriously affected by interference, third-party speaker and non-directional noise source. For recording, the authors set parameters: Sampling frequency $Fs = 16kHz$, overlap 50%. The original MA signals can be shown in Figure 2.



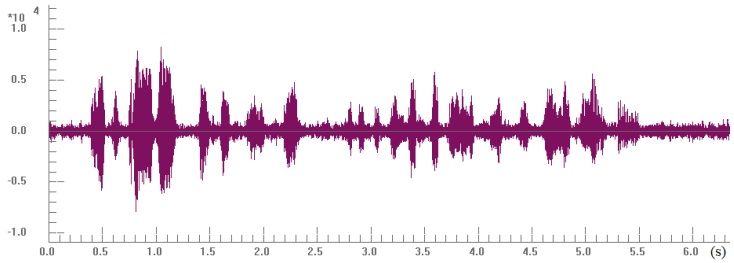
(a)



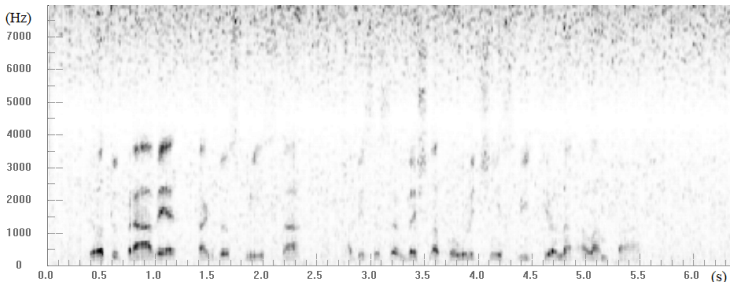
(b)

Figure 2. The captured MA signals, (a) waveform and (b) spectrogram

By applying the conventional MVDR beamformer, the beamformer's output signal is derived in Figure 3. Because of the undetermined reasons, the error of sampling frequency, the inaccurate estimation of steering vector, the different microphone sensitivities, the moving head of talker during the conversation, the performance of MVDR beamforemr's often corrupted. Hence, at the output of beamformer, the speech distortion or residual noise component degrades the speech quality and reduces the perceptual metric listener. Therefore, by applying the efficient post-filtering, the promising result has shown in Figure 4.

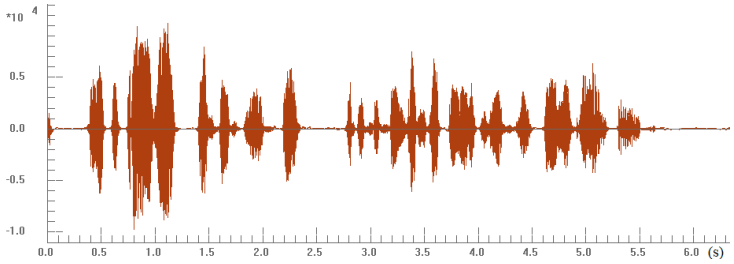


(a)

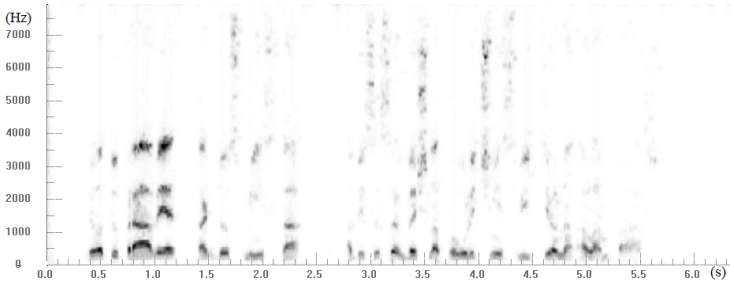


(b)

Figure 3. Performance of conventional MVDR beamformer, (a) The waveform and (b) spectrogram of conventional MVDR beamformer's results



(a)



(b)

Figure 4. The obtained promising signal of the proposed post-filtering, (a) waveform and (b) spectrogram

As we can see from Figure 5, the residual and musical noise level were suppressed to 10.9 dB. Using the objective measurement, the SNR of perspective result was increased from 8.5 to 12.3 dB in comparison with traditional MVDR beamformer. The Table 1 and Figure describes the energy of MA signals, the processed signal by MVDR beamformer and asm. asm has the capability of removing the surrounding noise while saving the original speech component.

The author's approach has the appealing properties is utilizing maximum likelihood to compute the covariance of noise and speech according to the rapid changed environment and complex situation. The speech presence probability $spp(f,k)$ was used for tracking and updating the formulation of coherence between two points sound source and presents exactly the characteristic of background noise, which derives computing variance of noise component $\sigma^2_{nMVDR}(f,k)$ at MVDR beamformer's output signal. In addition, the suggested technique measured the amount of uncorrelated component between speech and noise for estimating precise variance of speech.

Therefore, an additive post-filtering can be applied to further remove the remained noise component, residual noise. The effectiveness of post-filtering is not only attenuating background noise but also preserving the original speech component of desired target speaker. The illustrated experiment was performed in realistic recording living room in scenario of presence with third-party talker, interference and non-directional noise has approved the capability of speech enhancement and noise reduction of suggested post-filtering.

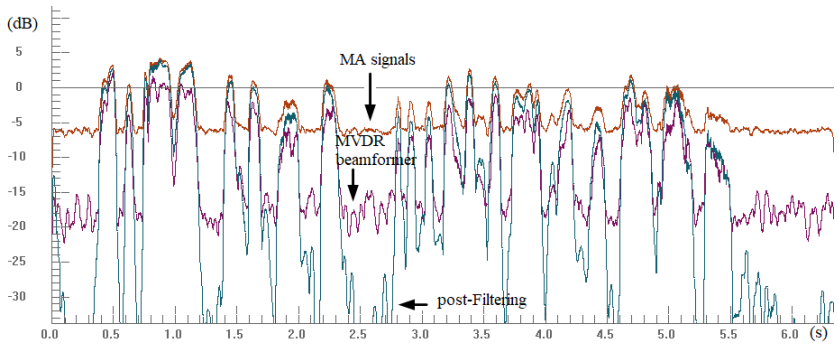


Figure 5. The comparison of energy between signals

Table 1. The obtained SNR (dB)

Method estimation	Microphone array signal	MVDR beamformer	ADL	GSC beamformer	Post-filtering
NIST STNR	3.5	11.8	13.4	13.9	20.3
WADA SNR	2.1	12.2	14.1	14.7	24.5

From the obtained Figures and Table 1, this article can conclude some results about experiments,

- MA beamformer has the ability of removing surrounding noise field, interference and different types of noise. MVDR beamformer based on optimum constrained criteria of suppressing unwanted noise while preserving the original desired target talker.
- Due to the complex and annoying environment, the influence of undetermined noisy recording scenario, the microphone mismatches, the error of sampling frequency, internal electrical acoustic equipment, the different time of recording between microphones, the inaccurate sampling frequency, the movement of speaker and heterogeneous environment, MVDR beamformer often corrupted. The speech distortion, ambient noise, residual or musical noise distor the speech quality of beamformer's output signal. In our experiment, the speech quality of beamformer's performance is not unsatisfied. At the speech frame, musical noise occurs.
- ADL used an appropriate value to increase the robust of MVDR beamformer in reducing the white noise gain. The essential GSC beamformer core is an efficient adaptive noise canceller to suppress the surrounding noise field while saving the target speaker. As a results, these improved methods derived more noise level reduction. However, ADL often depends on choosing the diagonal loading factor, which is an constant value and GSC beamformer is very sensitive with the error of computing transfer function or microphone quality. So the effectiveness of mentioned technique is limited.
- The author's idea is expanding the model Wiener filter - based post - Filtering. This approach uses the speech presence probability and modified formulation of coherence diffuse/coherent noise field for adaptively

coinciding the rapidly changed environment. The novelty is estimating the uncorrelated value between speech and noise component. Consequently, this contribution computed the exact the original variance of speech data. As a results, an effective post-filtering was derived for suppressing residual, musical and ambient noise component after performing MVDR beamformer.

- The demonstrated simulation was conducted under living room with presence different types of noises, the experiment applied dual-microphone system for evaluating MVDR beamformer. The numerical simulation has shown the improvement of eliminating the background noise level while preserving the clean speech data. Consequently, the speech quality in the term of signal-to-noise ratio (SNR). The perspective results shown the average noise reduction is 10.9 dB, the obtained speech quality was improved from 8.5 to 12.3 dB.
- post-filtering is an efficient solution for alleviating surrounding noise to enhance the objective measurement of beamformer's output signal. The paper's method owns the low computation and easy installation into various speech equipments for resolving complicated problems, such as, automatic speech recognition, voice-controlled device, distant speech acquisition, smart-home.
- The effectiveness of described method depends on the direction of arrival of useful signal θ_s. So under realistic recording environment, due to the movement of talker, the influence of internal electrical noise, the microphone mismatches, the presence of multiple undetermined noise sources, the presented technique can not deal practical problems.

5. CONCLUSION

In this contribution, the author illustrated an efficient post-filtering for separating the original speech component while suppressing transient and residual noise. The author's idea is exploiting the speech presence probability to obtain the precise estimation of coherence and the maximum likelihood approach for computing the exact observed MA signals, which according to the rapid changed recording scenario and complex situation. The author proposed method includes the estimation of covariance matrix of recorded MA signals by applying maximum likelihood. The effectiveness of described idea is confirmed by increasing the speech quality in the terms of SNR and reducing the background noise level. The suggested technique can be applied into multi-channel system for dealing other complicated problems, such as, speech recognition, reverberation and audio surveillance device. In the future, the contribution will study and incorporate the properties of surrounding noise field, accurate estimation of time delay for improving the results of the proposed post-filtering.

ACKNOWLEDGMENT

This work was supported by Posts and Telecommunications Institute of Technology (PTIT), Hanoi, Vietnam.

REFERENCES

[1] R. Fang et al., "A frequency invariant broadband beamformer based on directional microphone array," in 2022 14th International Conference on Signal Processing Systems (ICSPS), Jiangsu, China, 2022, pp. 623-627, doi: 10.1109/ICSPS58776.2022.00115.

[2] G. Huang, I. Cohen, J. Benesty, and J. Chen, "Kronecker product beamforming with multiple differential microphone arrays," in 2020 IEEE 11th Sensor Array and Multichannel Signal Processing Workshop (SAM), Hangzhou, China, 2020, pp. 1-5, doi: 10.1109/SAM48682.2020.9104333.

[3] M. T. Hossein, M. S. Hossain, M. Rashid, and N. Ahmed, "Robust microphone beamformer in the presence of multiple interfering signals," in 2021 3rd International Conference on Electrical Electronic Engineering (ICEEE), Rajshahi, Bangladesh, 2021, pp. 173-176, doi: 10.1109/ICEEE54059.2021.9718861.

[4] F. Zhang, J. Benesty, C. Pan, and J. Chen, "A universal linear beamformer for microphone arrays," in 2025 IEEE 15th International Conference on Signal Processing, Communications and Computing (ICSPCC), Hong Kong, 2025, pp. 1-5, doi: 10.1109/ICSPCC66825.2025.11194607.

[5] S. Tao, K. Yang, S. Kindt, J. R. Jensen, M. G. Christensen, and N. Madhu, "Array agnostic multi-channel speech presence probability estimation," in 2025 33rd European Signal Processing Conference (EUSIPCO), Palermo, Italy, 2025, pp. 176-180, doi: 10.23919/EUSIPCO63237.2025.11226120.

[6] J. Bartolewska and K. Kowalczyk, "Frame-based maximum a posteriori estimation of second-order statistics for multichannel speech enhancement in presence of noise," in 2021 IEEE 23rd International Workshop on Multimedia Signal Processing (MMSP), Tampere, Finland, 2021, pp. 1-6, doi: 10.1109/MMSP53017.2021.9733520.

[7] S. Tao, P. Mowlaee, J. R. Jensen, and M. G. Christensen, "Learning-based multi-channel speech presence probability estimation using a low-parameter model and integration with MVDR beamforming for multi-channel speech enhancement," in 2024 18th International Workshop on Acoustic Signal Enhancement (IWAENC), Aalborg, Denmark, 2024, pp. 100-104, doi: 10.1109/IWAENC61483.2024.10694149.

[8] S. Tao, P. Mowlaee, J. R. Jensen, and M. G. Christensen, "Multi-channel speech enhancement guided by learning-based a posteriori speech presence probability estimation," IEEE Transactions on Audio, Speech, and Language Processing, vol. 33, pp. 3838-3849, 2025, doi: 10.1109/TASLPRO.2025.3599782.
[9] S. -L. Pu, H. Guo, H. -W. Xie, C. Chen, P. Zhou, and G. -Q. Zhou, "Minimum variance optimizations with different covariance matrices in plane-wave ultrasound imaging," IEEE Transactions on Instrumentation and Measurement, vol. 73, pp. 1-13, 2024, Art no. 4509213, doi: 10.1109/TIM.2024.3476709.
[10] W. Lobato and M. H. Costa, "Worst-case-optimization robust-MVDR beamformer for stereo noise reduction in hearing aids," IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 28, pp. 2224-2237, 2020, doi: 10.1109/TASLP.2020.3009831.
[11] F. Zhang, C. Pan, J. Benesty, and J. Chen, "Approximate MVDR beamformer with automatic null steering and free of noise covariance matrix estimation," IEEE Transactions on Audio, Speech, and Language Processing, vol. 34, pp. 848-863, 2026, doi: 10.1109/TASLPRO.2026.3657247.
[12] P. -O. Lagacé, F. Ferland, and F. Grondin, "Ego-noise reduction of a mobile robot using noise spatial covariance matrix learning and minimum variance distortionless response," in 2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Detroit, MI, USA, 2023, pp. 3533-3538, doi: 10.1109/IROS55552.2023.10342193.
[13] Y. Wang, H. Xu, B. Wang, and M. Sun, "Broadband spatial spectrum estimation based on space-time minimum variance distortionless response and frequency derivative constraints," IEEE Access, vol. 11, pp. 27955-27962, 2023, doi: 10.1109/ACCESS.2023.3258978.
[14] R. Basumatary and G. Murmu, "Implementation of Sidelobe Canceller using MVDR Beamforming Technique," in 2025 International Conference on Wireless Communications Signal Processing and Networking (WiSPNET), Chennai, India, 2025, pp. 1-5, doi: 10.1109/WiSPNET64060.2025.11005384.
[15] M. R. Islam and S. S. Islam, "Comparative analysis of MVDR beamformer with conventional beamformer for the efficient interference reduction by null-forming for radar applications," in 2024 IEEE International Conference on Signal Processing, Information, Communication and Systems (SPICSCON), Khulna, Bangladesh, 2024, pp. 1-5, doi: 10.1109/SPICSCON64195.2024.10940083.
[16] Y. Luo, C. Yang, Y. Jiao, and L. Xu, "Deep MVDR: a high resolution DoA estimation method against array gain-phase error," in 2024 IEEE International Conference on Signal, Information and Data Processing (ICSIDP), Zhuhai, China, 2024, pp. 1-6, doi: 10.1109/ICSIDP62679.2024.10868173.
[17] Z. -W. Tan, A. H. T. Nguyen, L. T. T. Tran, and A. W. H. Khong, "A joint-loss approach for speech enhancement via single-channel neural network and MVDR beamformer," in 2020 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), Auckland, New Zealand, 2020, pp. 841-849.
[18] Y. Liang, C. Bao, and J. Zhou, "An implementation of the CNN-based MVDR beamforming for speech enhancement," in 2021 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC), Xi'an, China, 2021, pp. 1-5, doi: 10.1109/ICSPCC52875.2021.9564817.
[19] H. Kim, K. Kang, and J. W. Shin, "Factorized MVDR deep beamforming for multi-channel speech enhancement," IEEE Signal Processing Letters, vol. 29, pp. 1898-1902, 2022, doi: 10.1109/LSP.2022.3200581.
[20] R. Martin, "Noise power spectral density estimation based on optimal smoothing and minimum statistics," IEEE Transactions on Speech and Audio Processing, vol. 9, no. 5, pp. 504-512, Jul. 2001, doi: 10.1109/89.928915.
[21] J. Bitzer, K. -D. Kammeyer, and K. U. Simmer, "An alternative implementation of the superdirective beamformer," in Proceedings of the 1999 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA '99), New Paltz, NY, USA, 1999, pp. 7-10, doi: 10.1109/ASPAA.1999.810836.
[22] K. Chaudhari, M. Sutaone, and P. Bartakke, "Adaptive diagonal loading of MVDR beamformer for sustainable performance in noisy conditions," in 2020 IEEE Region 10 Symposium (TENSYMP), Dhaka, Bangladesh, 2020, pp. 1144-1147, doi: 10.1109/TENSYMP50017.2020.9230850.
[23] S. M. Kim and H. K. Kim, "Multi-microphone target signal enhancement using generalized sidelobe canceller controlled by phase error filter," IEEE Sensors Journal, vol. 16, no. 21, pp. 7566-7567, Nov. 2016, doi: 10.1109/JSEN.2016.2602375.
[24] D. P. W. Ellis, "SNREVAL: speech noise ratio and voice activity detection evaluation tools," Laboratory for the Recognition and Organization of Speech and Audio (LabROSA), Columbia University, 2007.
[25] SNRVAD [Online]. Available: https://labrosa.ee.columbia.edu/projects/snreval/

BIOGRAPHIES OF AUTHORS



Nguyen Thi Huyen Chau graduated with a Bachelor's degree in Information Technology from the University of Engineering and Technology, Vietnam National University, Hanoi, and a Master's and PhD in Computer Science from Ecole Nationale Superieure de Mechanique and d'Aerotechnique – ENSMA, France. Dr. Chau has been the Head of the Faculty of Information Technology at Thang Long University (TLU) since 2023. She has also been the Head of the Artificial Intelligence Laboratory, TLU, since 2019. She has supervised 7 master's students, authored or co-authored 25 scientific works, participated in 12 scientific research projects, including 1 completed ministerial-level project and 1 state-level project that is being approved. Her research interests include soft computing, machine learning, intelligent systems, schedulability analysis. She has been leading projects applying artificial intelligence (AI) for screening depression disorder symptoms, integrating AI in music teaching, developing virtual assistants for ATM and CRM devices, developing tools to evaluate learning outcomes and personalize student learning paths at universities. She can be contacted at email huyenchau@thanglong.edu.vn.



Quan Trong The received his Ph.D. in mathematics from the National Research University ITMO in 2021. He works as a lecturer at Faculty of Information Security of Posts and Telecommunica- tions Institute of Technology, Hanoi, Vietnam. His research interests are related to cybersecurity, speech enhancement, image processing, machine learning. Ha has authored or coauthored more than 30 publications: 27 proceedings and 03 journals. He can be contacted at email: theqt@ptit.edu.vn.



Nguyen Duc Toan is the Acting Vice Director of the Institute of Information Technology at the Vietnam Women's Academy. With extensive experience in higher education, academic management, and information technology, he has played an active role in advancing digital transformation initiatives in education and promoting the integration of emerging technologies into teaching, learning, and research activities. His research interests include cryptography, communications, and network security. He can be contacted at email: toannd@hvpnvn.edu.vn.