

An improved MVDR beamformer's noise reduction based on post - filtering

Nguyen Thi Huyen Chau¹, Quan Trong The², Nguyen Minh Thang³

¹Faculty of Information Technology, Thang Long University, Hanoi, Vietnam

²Faculty of Information Security, Posts and Telecommunications Institute of Technologies (PTIT), Hanoi, Vietnam

³Faculty of Information Technology, Ho Chi Minh City University of Economics and Finance, Ho Chi Minh City, Vietnam

Article Info

Article history:

Received May 22, 2025

Revised Jul 27, 2026

Accepted Aug 19, 2026

Keywords:

Microphone array beamforming
Minimum variance distortion less response

Noise reduction

Post - filtering

The signal-to-noise ratio

ABSTRACT

Nowadays, microphone array (MA) technology has been installed into almost acoustic equipments, such as surveillance device, mobile phone, voice controlled - device, teleconference system. Due to its high diversity, high directivity index towards the sound source, MA allows obtaining the original useful of desired talker, the clean speech data or necessary component while attenuating the effect of interferences, annoying background noise, third - party talker with promising spatial diversity, directivity index or different acoustic factors. Minimum variance distortionless response (MVDR) beamformer based on the optimum constrained formulation of mitigating complicated environment, background noise field while saving the speech component. However, in realistic recording scenario, because of the error of estimating the direction of arrival, the different microphone quality, the error of sampling frequency, the displacement of microphone geometry, the existence of internal electrical noise, the MVDR beamformer's evaluation degraded. Consequently, the existence of the remaining noise or speech distortion usually occur. This contribution introduced an additive post - Filtering, which blocks the noise level and increases the speech quality from 10.8 to 12.7 dB and reduces the noise level to 11.2 dB. The numerical experiment has confirmed the effectiveness of the author's suggested technique in improving the robustness of MVDR beamformer in complex environment.

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 Technologies (PTIT)

Hanoi, Vietnam

Email: theqt@ptit.edu.vn

1. INTRODUCTION

In real-life scenario, automatic speech recognition are significantly degraded by the sound of transport vehicle, television, washing machine, mechanical equipment, aircraft with several types of noise, non-directional noise, coherent, incoherent, diffuse noise field or diffuse noise field. Consequently, these factors has seriously corrupts the speech intelligibility, different objective acoustic measurement, the perceptual metric listener and speech quality of the recorded signals. Therefore, the essential in almost acoustic speech applications is addressing the problem of noise reduction and improving the speech quality. To deal with this task, speech enhancement or separation is usually implemented as a front-end to remove annoying and complex

noise. Single-channel approach often based on subtraction method, which uses the criteria about stationary noise. Therefore, in several non-stationary noise or in the complex recording situation with non-directional noise, diffuse noise field, single-channel algorithm causes the speech distortion of desired target speaker. To avoid this drawback, microphone array (MA) [1]-[3] technology can be used to achieve advantage of noise suppression, dereverberant and speech enhancement simultaneously with perspective results without speech distortion or musical noise. MA technology exploit the designed MA geometry to steer frequency response at specified direction while removing signal from other directions.

With the advantage of high spatial diversity, high directivity factor and the ability of incorporating single-channel approach, MA technology installed into multi-channel system, such as: Smartphone, acoustic instrument, hearing aid, cochlear implant, human-machine interface, teleconference system to improving the speech quality of processed signal. The fixed beamformer with delay-and-sum (DAS) beamformer use assumed impinging incident of speaker to form the weights, which are applied to MA signals. Adaptive beamformer incorporate the prior information, the surrounding acoustic properties, the transfer function, the designed array configuration, the characteristics of background noise to obtain optimum beamformer's coefficient in preserving the desired talker while eliminating interference, background noise. The most useful representations of adaptive beamformer is differential microphone (DIF) [2], [4]-[6], generalized sidelobe canceller (GSC) [7]-[10], linearly constrained minimum variance (LCMV) [11]-[14], minimum variance distortionless response (MVDR) [15]-[22].

MVDR beamformer has capability of concerning the steerable beampattern at sound source, interest useful speech. However, the performance often interfered by complex and adverse environment. There are many efforts attempt to increase the overall evaluation under realistic scenario.

Wang *et al.* [15] introducing exploiting MVDR method for calculating spectral estimation based on the Kronecker product to achieve better evaluation under various recording scenarios. This article exploited the spectral estimation filters for achieving the more robust of speech enhancement than the traditional beamformer in realistic experiment.

Yamaoka *et al.* [16] time-frequency-bin-wise switching (TFS), which served as efficient voice activity detection (VAD) to switch implemented beamformer for processing array signals. The advantage of this approach is suppressing several particular interferences, non-directional noise sources by using MVDR beamformer. The numerical simulation has shown the improvement of speech enhancement in speech quality, perceptual metric listener underdetermined situations.

Tuladhar *et al.* [17] proposed the unit circle rectification (UCR) method for satisfying a property of the ensemble MVDR beamformer. This technique incorporated the sample matrix inversion with MVDR beamformer to attenuating the effect of limited snapshots. Consequently, UCR satisfied the criteria of MVDR beamformer ensemble. The simulated results has shown the effectiveness of UCR in reducing the white surrounding noise field, interference than diagonal loading. The demonstrated experiment show the advantage of the author's suggested technique in improving the speech quality of MVDR beamformer's output signal.

As'ad *et al.* [18] proposed selection between monaural and binaural MVDR beamformer at each considered time-frequency bin. The decision depends on an accurate target activity detection system, which based on a binaural beamforming system for cancelling target source. The described approach applied binaural beamforming as speech presence probability in demanding of blocking the target speech component. The simulation results show the robustness of calculating the transfer function, related steering vector of impinging angle between existing algorithms.

In [19], the authors proposed using principal component analysis (PCA) to adaptively adjust covariance matrix to reduce ego-noise by applying MVDR beamformer. The promising result has confirmed the effectiveness of reducing the word error rate by 55% and enhancing the signal-to-distortion ratio by up to 10 dB under various types of recording scenarios.

Park *et al.* [20] used beamspace approximation, which use a combination of the spatial fast-fourier transform (FFT) to obtain linear-logarithmic complexity to enhance constructing and inverting the array covariance matrix. The conducted experiments studied the MVDR beamformer's tradeoffs between load computation and signal-plus-noise ratio (SINR) to examine the superiority for large array system in realistic environment.

Son *et al.* [21] calculated new frequency estimation with considering a variable parameter for computing the augmented covariance matrix to resolves a significant degradation of MVDR beamformer's evaluation. The improved performance is derived through computer simulations.

Zhang *et al.* [22] addressed the problem of source separation from the noisy mixture by using deep learning MVDR framework with recurrent neural networks for resolving the problem of matrix inversion and eigenvalue decomposition. The promising result is removing the residual noise field while keeping the target speech undistorted. The numerical results have confirmed the superiority with several metrics and automatic speech recognition under complex situations.

Jiang *et al.* [23] presented multistage forward/backward for reducing the complexity of sample covariance matrix's (SCM) inversion. The obtained advantage of this method is making the original complex-value of SCM progressively become real-valued block diagonal matrix. The numerical simulation demonstrates that the suggested technique own high computational robustness and speech quality in comparison with existing MVDR beamformer in extracting the desired target speaker at specified location while minimizing the total noise power.

Unfortunately, due to the complex and undetermined reasons about the complexity of surrounding noise field, the existence of different annoying factors, these mentioned technique can not deal the practical problem due to the displacement of MA distribution, inaccurate estimation of preferred steering vector, the error of sampling frequency, the different microphone sensitivities, the moving head of speaker during conversation, the complex and adverse environment, the microphone mismatches; MVDR beamformer's performance often corrupted, the existence of musical noise, unacceptable noise level or speech distortion cause the degradation of speech quality. These above works can not deal this problem. Therefore, in this paper, the authors proposed an efficient post-Filtering for alleviating the musical noise, background noise to improve the objective measurement, speech quality in the term of signal-to-noise ratio (SNR), speech intelligibility under real-life situations. The observed MA signals is shown in Figure 1.

2. MVDR BEAMFORMER

In this section, the author describes the principal working of MVDR beamformer in the frequency-domain as in Figure 2. MVDR beamformer based on the constrained criteria of minimizing the total output noise power at the output while preserving the original speech component at specified direction without speech distortion [15]-[18]. MVDR beamformer owns the high spatial diversity at certain sound source.

In general case, the author uses dual - microphone system (DMA2) to illustrate the principal working of MVDR beamformer. At the current frame k , current frequency f , $\omega = 2\pi f$, the observed MA signals $Y_1(j\omega, k)$, $Y_2(j\omega, k)$ in the short-time Fourier transform can be represented as the following equations,

$$Y_1(j\omega, k) = S(j\omega, k)e^{j\Phi_s} + N_1(j\omega, k) \quad (1)$$

$$Y_2(j\omega, k) = S(j\omega, k)e^{-j\Phi_s} + N_2(j\omega, k) \quad (2)$$

where $\Phi_s = \pi f \tau_0 \cos(\theta_s)$, $\tau_0 = \frac{d}{c}$, θ_s is the incident angle of interest useful signal to the axis of DMA2, d is the range between two mounted microphones, c is sound speed propagation in the air and $c = 343(\frac{m}{s})$. $N_1(j\omega, k)$, $N_2(j\omega, k)$ means the background noise, interference and competitive speaker, or non - directional noise, $S(j\omega, k)$ is the desired target data.

With the definition, $\mathbf{Y}(j\omega, k) = [Y_1(j\omega, k) \ Y_2(j\omega, k)]^T$, $\mathbf{N}(j\omega, k) = [N_1(j\omega, k) \ N_2(j\omega, k)]^T$, $\mathbf{h}(j\omega, \theta_s) = [e^{j\Phi_s} \ e^{-j\Phi_s}]^T$, the (1), (2) can be rewritten in the following way,

$$\mathbf{Y}(j\omega, k) = S(j\omega, k)\mathbf{h}(j\omega, \theta_s) + \mathbf{N}(j\omega, k) \quad (3)$$

The essential problem of speech enhancement is finding an optimum coefficient $\mathbf{W}(j\omega, k)$ to obtain the final signal $\hat{S}(j\omega, k)$ approximately the original clean speech data.

$$\hat{S}(j\omega, k) = \mathbf{W}^H(j\omega, k) * \mathbf{Y}(j\omega, k) \quad (4)$$

where H is conjugate transposition.

The mathematical formulation of MVDR beamformer can be expressed as,

$$\begin{aligned} \min_{\mathbf{W}(j\omega, k, \theta_s)} \quad & \mathbf{W}^H(j\omega, k, \theta_s) \Phi_{NN}(j\omega, k) \mathbf{W}(j\omega, k, \theta_s) \\ \text{s.t.} \quad & \mathbf{W}^H(j\omega, k, \theta_s) \mathbf{h}(j\omega, \theta_s) = 1 \end{aligned} \quad (5)$$

with $\Phi_{NN}(j\omega, k) = E\{\mathbf{N}^H(j\omega, k)\mathbf{N}(j\omega, k)\}$ mean the captured noisy covariance matrix.

Hence, the MVDR beamformer's filter coefficients can be derived as,

$$\mathbf{W}(j\omega, k, \theta_s) = \frac{\Phi_{NN}^{-1}(j\omega, k)\mathbf{h}(j\omega, \theta_s)}{\mathbf{h}^H(j\omega, \theta_s)\Phi_{NN}^{-1}(j\omega, k)\mathbf{h}(j\omega, \theta_s)} \quad (6)$$

However, in realistic recording scenario, the information about noise is not always available, therefore, we use the observed MA instead of. So, the final optimum MVDR beamformer's weights can be computed as,

$$\mathbf{W}_{MVDR}(j\omega, k, \theta_s) = \frac{\Phi_{YY}^{-1}(j\omega, k)\mathbf{h}(j\omega, \theta_s)}{\mathbf{h}^H(j\omega, \theta_s)\Phi_{YY}^{-1}(j\omega, k)\mathbf{h}(j\omega, \theta_s)} \quad (7)$$

where $\Phi_{YY}(j\omega, k) = E\{\mathbf{Y}^H(j\omega, k)\mathbf{Y}(j\omega, k)\}$ is the spectral covariance matrix of captured MA signals.

The auto - cross power spectral densities (PSD) between two received MA signals are recursively calculated by equations,

$$P_{Y_i Y_i}(j\omega, k) = (1 - \alpha)P_{Y_i Y_i}(j\omega, k - 1) + \alpha Y_i^*(j\omega, m)X_i(j\omega, k) \quad (8)$$

$$P_{Y_i Y_j}(j\omega, k) = (1 - \alpha)P_{Y_i Y_j}(j\omega, k - 1) + \alpha Y_i^*(j\omega, k)Y_j(j\omega, k) \quad (9)$$

with $i, j = 1, 2$, α is smoothing parameter in range of $\{0..1\}$.

The output signal of MVDR beamformer is derived as,

$$y_{MVDR}(j\omega, k) = \mathbf{W}_{MVDR}^H(j\omega, k, \theta_s) \times \mathbf{Y}(j\omega, k) \quad (10)$$

Due to many reasons, such as, the complex and annoying environment, the error of MA distribution, the error of sampling rate, the difference of MA quality, or imprecise estimation of preferred steering vector, the performance of MVDR beamformer usually degraded. The requirement of an efficient signal processing technique is an essential part of numerous speech applications. Consequently, in the next section, the author proposed an effective post - Filtering for suppressing the musical noise, remained noise at the MVDR beamformer's output to increase the speech quality, speech intelligibility.

3. THE AUTHOR'S PROPOSED POST - FILTERING

The author's approach is modifying the Wiener filter for suppressing the noise level at MVDR beamformer's output signal. The tradition Wiener filter $H_{Wiener}(j\omega, k)$ can be calculated as the following equation,

$$H_{Wiener}(j\omega, k) = \frac{\sigma_s^2(j\omega, k)}{\sigma_s^2(j\omega, k) + \sigma_n^2(j\omega, k)} \quad (11)$$

where $\sigma_s^2(j\omega, k) = \frac{1}{\mathbf{h}^H(j\omega, \theta_s)\Phi_{YY}^{-1}(j\omega, k)\mathbf{h}(j\omega, \theta_s)}$ is variance of original clean speech component at the specified direction θ_s and $\sigma_n^2(j\omega, k)$ is the covariance of noise at output of MVDR beamformer and can be computed by minimum statistic (MS) [24].

Due to the complex and rapidly changed recording scenario, the effectiveness of Wiener filter often degraded and there is still exists musical noise, remained noise component at the MVDR beamformer's output. The author proposed modifying Wiener filtering with incorporating the prior information of the phase difference between two captured MA signals $Y_1(j\omega, k)$, $Y_2(j\omega, k)$.

In [25], the prior information about the direction of arrival of interest useful speech component allows deriving the spectral masking $G_k^{BM}(j\omega) = 1 - \exp(j\phi_{12,k}^{norm}(j\omega))$ and $G_k^{DSB}(j\omega) = 0.5 + 0.5 * \exp(j\phi_{12,k}^{norm}(j\omega))$ and where $\phi_{12,k}^{norm}(j\omega) = \frac{\phi_{12,k} - c}{f.d}$, $\phi_{12,k}$ is the phase difference between $Y_1(j\omega, k)$, $Y_2(j\omega, k)$.

The modified Wiener filter is given by,

$$\hat{H}_{Wiener}(j\omega, k) = \beta \frac{\sigma_s^2(j\omega, k)}{\sigma_s^2(j\omega, k) + \sigma_{nMVDR}^2(j\omega, k)} + (1 - \beta) \frac{G_k^{DSB}(j\omega)}{G_k^{DSB}(j\omega) + G_k^{BM}(j\omega)} \quad (12)$$

where β is an appropriate parameter, which in $\{0..1\}$. In the experiment, the value of $\beta = 0.35$ was based on our heuristic knowledge.

The processed signal by applying modified Wiener filter is computed as,

$$\hat{y}_{MVDR}(j\omega, k) = y_{MVDR}(j\omega, k) * \hat{H}_{Wiener}(j\omega, k) \quad (13)$$

By adding the spectral masking gain, which based on the prior of direction of arrival of interest useful signal, the modified Wiener filter is suitable under annoying acoustic environment to removing the remained background noise to increase the speech quality, perceptual metric listener.

The appealing properties of described technique is utilizing the gain function, which based on preferred direction of arrival of interest helpful signal. The method exploits the gain of superdirective beamformer and blocking matrix for determining the gain function. In practical scenario, the success of Wiener filter usually degraded due to the complicated problem of estimating covariance of speech $\sigma_s^2(j\omega, k)$ and $\sigma_n^2(j\omega, k)$. These necessities parameters significantly depends on the microphone sensitivities, the homogeneity of recording situation, the time of simultaneous capture, the sampling frequency. Hence, the estimation of covariance of speech and noise usually not exact under complex environment with presence several undetermined noise sources. This additive gain function allows increasing the capability of suppressing the remained noise level. The exploiting of additive post-Filtering will improve the overall performance of MVDR beamformer. In the next section, the author performed an experiment in realistic situation to verify the effectiveness of the mentioned post - Filtering.

4. EXPERIMENTS AND DISCUSSION

In this section, the author will illustrate the advantage of the author's proposed method in enhancing the speech quality and reducing the remained noise level. The experiment was performed in living room with the existence of washing machine, interference and annoying background noise. The author uses DMA2 to capture the original speech component of talker. The range between two mounted microphones is $d = 5(cm)$, the distance from talker to the axis of DMA2 is $L = 3(m)$, the direction of arrival of speaker is $\theta_s = 90(deg)$.

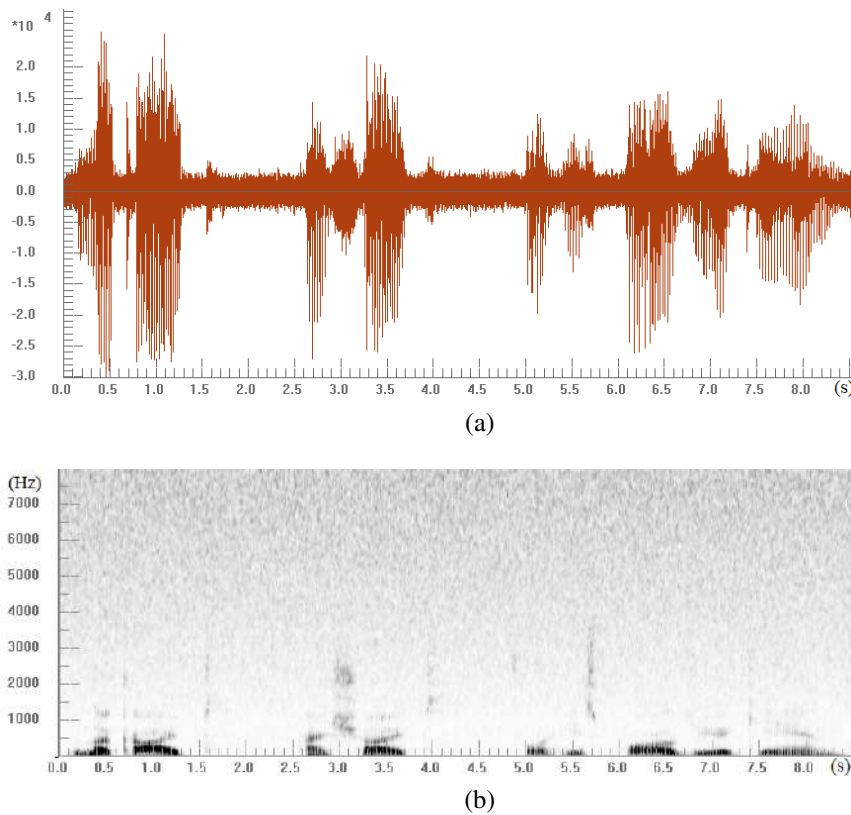


Figure 1. Original speech signals captured by DMA2 in the noisy experiment: (a) the waveform and (b) spectrogram of captured MA signals

An objective measurement was implemented for calculating the SNR for comparing the speech quality between the captured MA signals, the processed signal by tradition MVDR beamformer (dioMVDR) and the author's suggested post - Filtering (arsgpF). For capturing the talker's speech component, the sampling frequency $F_s = 16\text{kHz}$, the overlap 50% were used. The observed MA signals is shown in Figure 1. For performing MVDR beamformer, these parameters $nFFT = 512$, the smoothing parameter $\alpha = 0.1$. The obtained MVDR beamformer's output signal can be derived in Figure 2.

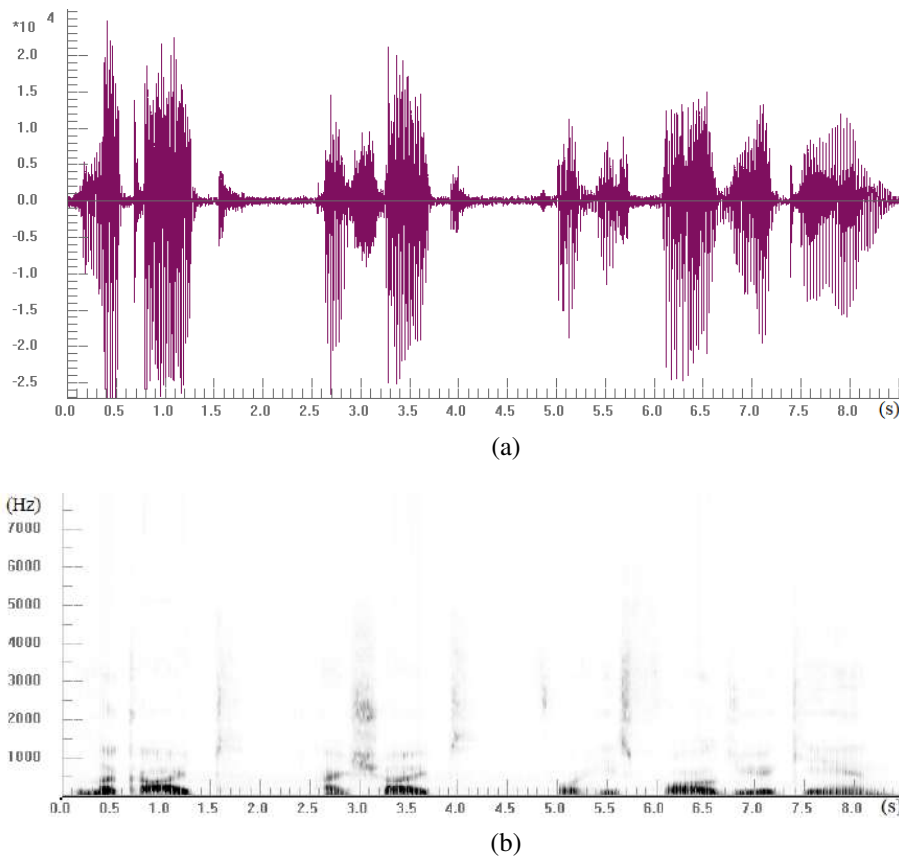


Figure 2. The promising signal of MVDR beamformmer's performance, (a) waveform and (b) spectrogram

In many complex environment, due to the undetermined reasons, such as the microphone mismatches, the inaccurate estimation of preferred steering vector, the displacement of MA distribution, the different microphone sensitivities, the error of sampling rate, the MVDR beamformer's performance often corrupted. The speech distortion or existence of remaining noise level is a challenge task. Therefore, the author proposed an additive post - Filtering for suppressing the noise while preserving the original speech component. By applying arsgpF, the processed signal yields as in Figure 3.

The compared energy between MA signals, the processed signal by dioMVDR, arsgpF is shown in Figure 4, arsgpF allows removing the noise level to 11.2 dB and consequently, the signal-to-noise was increased from 10.8 to 12.7 dB. The signal-to-noise ratio (dB) shown in Table 1.

Table 1. The signal-to-noise ratio (dB)

Method estimation	Microphone array signal	dioMVDR	arsgpF
NIST STNR	6.8	13.2	24.0
WADA SNR	1.5	9.4	22.1

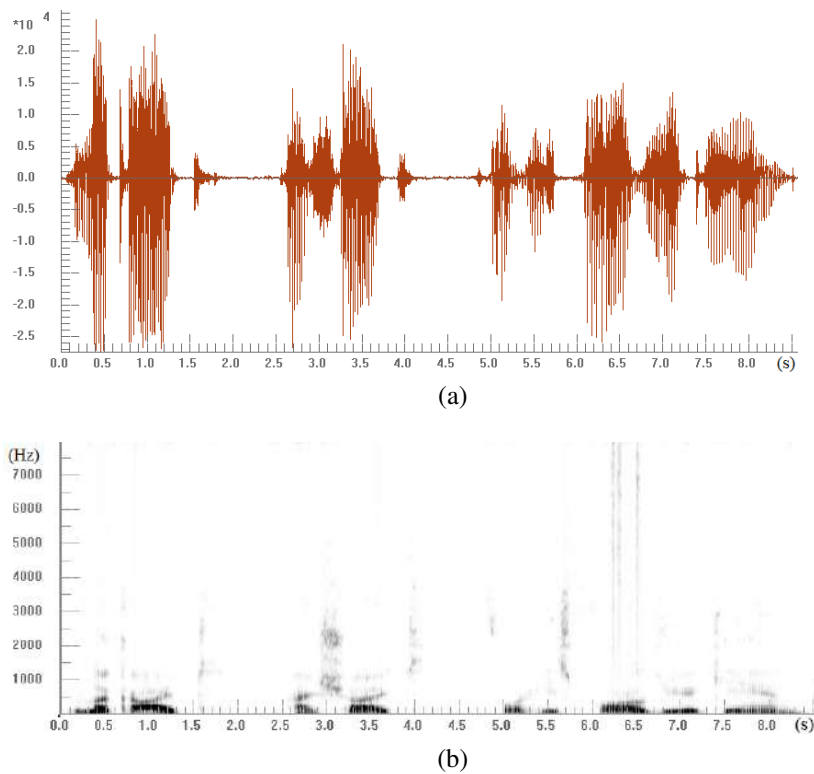


Figure 3. The perspective results by applying arsgpF, (a) waveform and (b) spectrogram

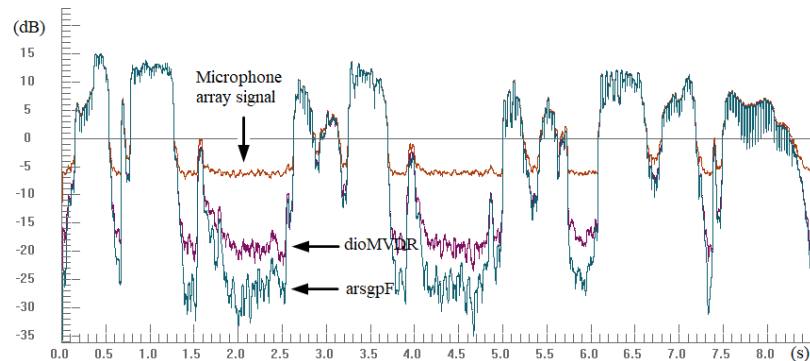


Figure 4. The comparison of energy between MA signals and processed signals by dioMVDR, arsgpF

From the conducted experiments, this paper can conclude some opinions,

- MA beamformer has the capability of suppressing the surrounding noise field, interference or the signals from other directions while preserving the clean speech data at specified location. In this experiment, the direction of target speaker at $90(deg)$ and the advantage of MVDR beamformer in preserving the clean speech data while attenuating the effect of background noise.
- However, due to complex and annoying situations, the displacement of microphones, the internal electrical noise of acoustic equipments, the effect of reverberation, the existence of several time-varying noise sources, the overall MVDR beamformer's evaluation often corrupted. The speech distortion or large ambient noise cause the unsatisfied speech quality. We can easily realize the large amount of residual noise component.
- Post - Filtering is an efficient solution for addressing this problem. Post - Filtering is constructed for saving the original speech data and reducing the surrounding noise field. In this contribution, the described post-Filtering was inherited from the model Wiener filter and combined with spectral gain function,

which based on the prior information about direction of arrival of interest useful signal. The hybrid integration between Wiener filter and gain function has increased the effectiveness of post-Filtering.

- The appealing properties of the author's suggested method is exploiting the priori information of preferred direction of arrival of desired talker, the phase difference between two MA signals to form an efficient post - Filtering for eliminating background noise while saving the original clean speech data. The author's suggested method not only suppresses background noise but also saving the use signal without speech distortion. The promising result verified the ability of the post - Filtering in adverse environment for solving other complex problems.

5. CONCLUSION

This paper addresses the problem of reducing the speech distortion in MVDR beamformer's performance in many complex recording situations. The author's proposed method is an efficient post - Filtering, which extracts the desired target speaker while eliminating the remaining surrounding noise. The numerical simulations has confirmed the effectiveness of reducing noise level to 11.2 dB and improving the speech quality from 10.8 to 12.7 dB in realistic recording scenario. The appealing properties of the author's suggested technique is based on the priori information of the preferred direction of arrival of interest useful signal and combination with speech presence probability to form an appropriate post - Filtering. This method can be integrated into multi-channel system for addressing other complicated problems, such as noise reduction, speech separation, speech recognition.

REFERENCES

- [1] Y. Jia, B. Gray and R. Vaughan, "Measurements of microphone array phase and amplitude behavior towards controllable beamforming," 2020 IEEE SENSORS, Rotterdam, Netherlands, 2020, pp. 1-4, doi: 10.1109/SENSORS47125.2020.9278799.
- [2] G. Huang, J. Benesty, J. Chen and I. Cohen, "Robust and steerable kronecker product differential beamforming With rectangular microphone arrays," ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Barcelona, Spain, 2020, pp. 211-215, doi: 10.1109/ICASSP40776.2020.9052988.
- [3] V. Tourbabin, I. Malka and E. Tzirkel-Hancock, "Performance of fixed in-car microphone array beamformer under variations in car noise," 2017 Hands-free Speech Communications and Microphone Arrays (HSCMA), San Francisco, CA, USA, 2017, pp. 21-25, doi: 10.1109/HSCMA.2017.7895554.
- [4] K. Zhao, X. Luo, J. Jin, G. Huang, J. Chen and J. Benesty, "Design of robust differential beamformers with microphone arrays of arbitrary planar geometry," ICASSP 2025 - 2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Hyderabad, India, 2025, pp. 1-5, doi: 10.1109/ICASSP49660.2025.10888941.
- [5] X. Zhao, X. Luo, G. Huang, J. Chen and J. Benesty, "Differential beamforming with null constraints for spherical microphone arrays," ICASSP 2024 - 2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Seoul, Korea, Republic of, 2024, pp. 776-780, doi: 10.1109/ICASSP48485.2024.10446768.
- [6] X. Luo, J. Jin, G. Huang, J. Chen and J. Benesty, "Design of steerable linear differential microphone arrays with omnidirectional and bidirectional sensors," in IEEE Signal Processing Letters, vol. 30, pp. 463-467, 2023, doi: 10.1109/LSP.2023.3267969.
- [7] F. M. Ribas *et al.*, "Minimum variance generalized sidelobe canceller and eigenspace-based generalized sidelobe canceller beamformers combined with frost postfilter for medical ultrasound imaging," 2021 IEEE UFFC Latin America Ultrasonics Symposium (LAUS), Gainesville, FL, USA, 2021, pp. 1-4, doi: 10.1109/LAUS53676.2021.9639164.
- [8] J. Wang, F. Yang, J. Guo and J. Yang, "Robust adaptation control for generalized sidelobe canceller with time-varying gaussian source model," 2023 31st European Signal Processing Conference (EUSIPCO), Helsinki, Finland, 2023, pp. 16-20, doi: 10.23919/EUSIPCO58844.2023.10289801.
- [9] S. Dai, M. Li, Q. H. Abbasi and M. A. Imran, "A fast blocking matrix generating algorithm for generalized sidelobe canceller beamformer in high speed rail like scenario," in IEEE Sensors Journal, vol. 21, no. 14, pp. 15775-15783, 15 July 2021, doi: 10.1109/JSEN.2020.3002699.
- [10] S. S. Priyanka and T. K. Kumar, "GSC beamforming using different adaptive algorithms for speech enhancement," 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT), Kanpur, India, 2019, pp. 1-6, doi: 10.1109/ICCCNT45670.2019.8944415.
- [11] A. Frank, A. Ben-Kish and I. Cohen, "Constant-beamwidth linearly constrained minimum variance beamformer," 2022 30th European Signal Processing Conference (EUSIPCO), Belgrade, Serbia, 2022, pp. 50-54, doi: 10.23919/EUSIPCO55093.2022.9909899.
- [12] A. M. Narayanan and A. Bertrand, "Group-utility metric for efficient sensor selection and removal in LCMV beamformers," ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Barcelona, Spain, 2020, pp. 4950-4954, doi: 10.1109/ICASSP40776.2020.9053467.
- [13] V. W. Neo, E. d'Olne, A. H. Moore and P. A. Naylor, "Fixed beamformer design using polynomial eigenvalue decomposition," 2022 International Workshop on Acoustic Signal Enhancement (IWAENC), Bamberg, Germany, 2022, pp. 1-5, doi: 10.1109/IWAENC53105.2022.9914799.
- [14] J. Zhang, J. Du and L. -R. Dai, "Sensor selection for relative acoustic transfer function steered linearly-constrained beamformers," in IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 29, pp. 1220-1232, 2021, doi: 10.1109/TASLP.2021.3064399.

- [15] X. Wang, J. Benesty, G. Huang and J. Chen, "A minimum variance distortionless response spectral estimator with kronecker product filters," 2022 30th European Signal Processing Conference (EUSIPCO), Belgrade, Serbia, 2022, pp. 2261-2265, doi: 10.23919/EUSIPCO55093.2022.9909584.
- [16] K. Yamaoka, N. Ono, S. Makino and T. Yamada, "Time-frequency-bin-wise Switching of Minimum Variance Distortionless Response Beamformer for Underdetermined Situations," ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Brighton, UK, 2019, pp. 7908-7912, doi: 10.1109/ICASSP.2019.8683528.
- [17] S. R. Tuladhar and J. R. Buck, "Unit circle rectification of the minimum variance distortionless response beamformer," in IEEE Journal of Oceanic Engineering, vol. 45, no. 2, pp. 500-510, April 2020, doi: 10.1109/JOE.2018.2876584.
- [18] H. As'ad, M. Bouchard and H. Kamkar-Parsi, "Robust minimum variance distortionless response beamformer based on target activity detection in binaural hearing aid applications," 2019 IEEE Global Conference on Signal and Information Processing (GlobalSIP), Ottawa, ON, Canada, 2019, pp. 1-5, doi: 10.1109/GlobalSIP45357.2019.8969387.
- [19] 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," 2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Detroit, MI, USA, 2023, pp. 3533-3538, doi: 10.1109/IROS55552.2023.10342193.
- [20] J. -J. Park, G. Rodriguez-Martinez, M. M. Bolding, R. S. Westafer and V. S. Sathe, "Investigation of scalable approximation of the minimum variance distortionless response beamformer," 2024 IEEE International Symposium on Phased Array Systems and Technology (ARRAY), Boston, MA, USA, 2024, pp. 1-6, doi: 10.1109/ARRAY58370.2024.10880424.
- [21] K. J. Son, G. S. An, Y. H. Kim and T. -G. Chang, "Improved "augmented MVDR spectrum-based frequency estimation algorithm" for Unbalanced Power Systems," in IEEE Transactions on Instrumentation and Measurement, vol. 70, pp. 1-5, 2021, Art no. 9003205, doi: 10.1109/TIM.2021.3079565.
- [22] Z. Zhang, Y. Xu, M. Yu, S. -X. Zhang, L. Chen and D. Yu, "ADL-MVDR: All deep learning MVDR beamformer for target speech separation," ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Toronto, ON, Canada, 2021, pp. 6089-6093, doi: 10.1109/ICASSP39728.2021.9413594.
- [23] S. Jiang, S. Liu, M. Jin and Z. Lin, "Efficient MVDR beamformer employing multistage forward/backward averaging scheme," in IEEE Transactions on Aerospace and Electronic Systems, vol. 61, no. 2, pp. 3963-3976, April 2025, doi: 10.1109/TAES.2024.3494758.
- [24] R. Martin, "Noise power spectral density estimation based on optimal smoothing and minimum statistics," in IEEE Transactions on Speech and Audio Processing, vol. 9, no. 5, pp. 504-512, July 2001, doi: 10.1109/89.928915.
- [25] S. M. Kim, "Hearing aid speech enhancement using phase difference-controlled dual-microphone generalized sidelobe canceller," in IEEE Access, vol. 7, pp. 130663-130671, 2019, doi: 10.1109/ACCESS.2019.2940047.

BIOGRAPHIES OF AUTHORS



Nguyen Thi Huyen Chau    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 Telecommunications Institute of Technology, Hanoi, Vietnam. His research interests are related to cybersecurity, speech enhancement, image processing, machine learning. He has authored or coauthored more than 30 publications: 27 proceedings and 03 journals. He can be contacted at email: theqt@ptit.edu.vn.



Quan Trong The    He received Engineering of Information Technology in 2017 from Ho Chi Minh City University of Technology, Master of Information Technology in 2022 from University of Information Technology, Vietnam National University at Ho Chi Minh City. His key areas of research include information security and applied AI for cybersecurity. He can be contacted at email: thangnm@uef.edu.vn.



Nguyen Minh Thang    He received Engineering of Information Technology in 2017 from Ho Chi Minh City University of Technology, Master of Information Technology in 2022 from University of Information Technology, Vietnam National University at Ho Chi Minh City. His key areas of research include information security and applied AI for cybersecurity. He can be contacted at email: thangnm@uef.edu.vn.