Differential Microphone Array Speech Enhancement Based on Post - Filtering

Quan Trong The

Digital Agriculture Cooperative, Cau Giay, Ha Noi, Viet Nam.

Abstract

Noise suppression has become an essential requirement for numerous acoustic devices, mobile phones and surveillance equipment. One standard criterion for almost digital signal processing is the saving of the target speech component while eliminating all background noise and interferences. Dual - microphone system is one of the most basic elements, which is widely commonly installed in speech applications. However, the designed signal processing faces many complex challenges in extracting the directional sound source. In this paper, the author proposed using an additive post-filtering to improve the author's previous research's performance. The evaluated experiment has confirmed the desired noise reduction to 5.5 (dB) and increasing the speech quality in terms of the signal-to-noise ratio from 3.2 (dB) to 5.7 (dB).

Keywords 1

microphone array, the signal-to-noise ratio, speech enhancement, noise reduction, dual - microphone system, post - filtering

1. Introduction

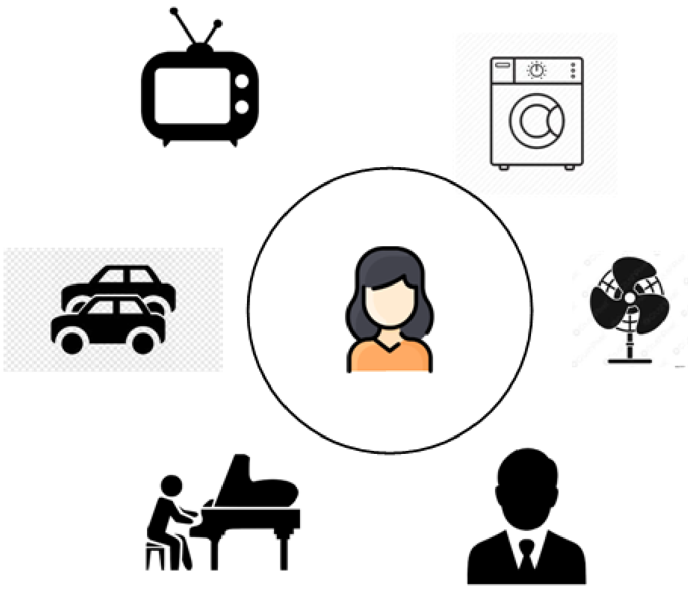

Figure 1: The complicated preserving of the target speaker real-life

The demand of noise suppression in almost speech processing applications like speech recognition, hearing aids, cell phone, surveillance equipment, smart home applications to work anytime and

COLINS-2023: 7th International Conference on Computational Linguistics and Intelligent Systems, April 20–21, 2023, Kharkiv, Ukraine
EMAIL: quantrongthe1984@gmail.com
ORCID: 0000 - 0002 - 2456 - 9598

anywhere, makes them has the capacity of reducing the effect of annoying disturbances, like background noise, interferences, third - party or surrounding vehicle transport. To decrease the degradation of desired target talker in terms of speech quality, speech intelligibility, the single and multi - microphone noise reduction approach are often applied for signal processing. Spectral estimation technique is the most widely implemented in single - channel techniques, such as: subspace method, Wiener filter and spectral subtraction, are based on calculation of the noise power with assumption that the background noise are stationary, or under the situation with ambient speech noise.

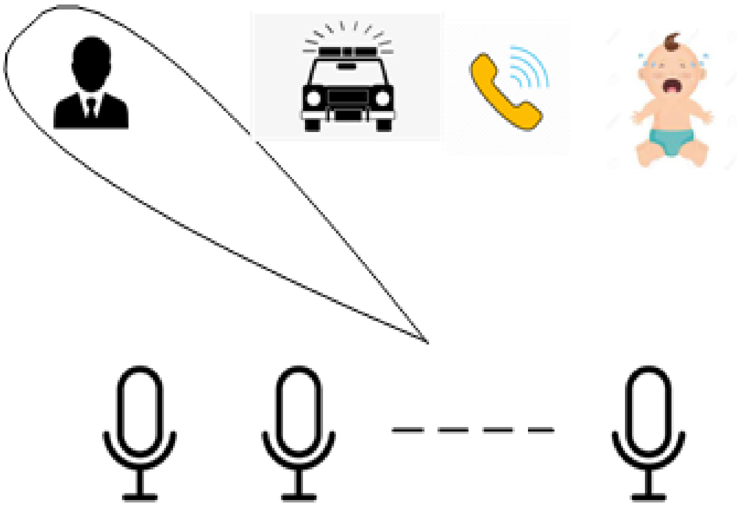

Figure 2: The using microphone array for extracting the desired talker

The microphone array beamforming [1-20] exploits the priori spatial information of the direction of speaker, the properties of recording environment, the direction of arrival of useful interest signal and the geometry of MA's configuration. Due to the possibility to perform spatial beamforming, MA's algorithm has a better ability to alleviate the noise level and save the speech component. For multi-microphone noise suppression, using spatial information to obtain the advantage of preserving the target signal while eliminating noisy environment. Besides, MA beamforming can be applied the pre - processing, post - filtering methods to achieve noise reduction and decrease the speech distortion.

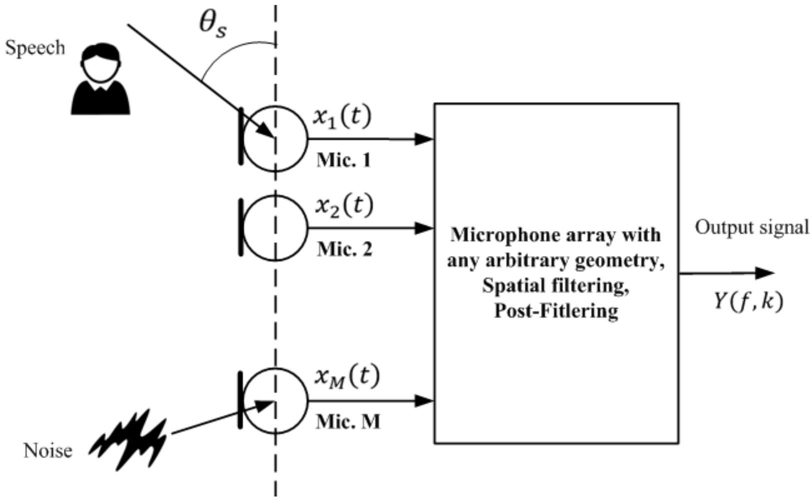

Figure 3: The implementation of microphone array in time - frequency domain

Dual - microphone system (DMA2) has numerous advantages for signal processing. DMA2 owns compact, easy to implement MA digital signal processing. In the previous work [23], the author suggested an effective method for separating each target, which stands at two opposite directions. However, in the real-life world, due to its undetermined recording situations, the performance also corrupted. In this contribution, the author proposed using an additive post - filtering to block the remaining noisy component after using [23]. The numerical results have rated the better performance of noise suppression to 5.5 (dB) and increasing of the speech quality from 3.2 to 5.7 (dB). The purpose of this article is demonstrating a new effective technique to enhance speech enhancement, which based on DMA2.

The remaining section of this article is organized as follows. The next section describes the signal model of DMA2 and the author's previous evaluation. Section III demonstrated additive post - Filtering, which uses MMSE estimator. Section IV shows the illustrated experiment and Section V concludes the purpose of this paper.

2. The signal model

A scheme of principal working of a certain differential microphone array (DMA2) [24 - 30] is illustrated in Figure 4. DMA2 owns high directional beampattern, high noise reduction and easy implemented to form a beampattern toward the sound source. DMA2 has a compact size, and is very suitable for microphone array technique, digital signal processing method to block surrounding noise environment and save the target desired speaker.

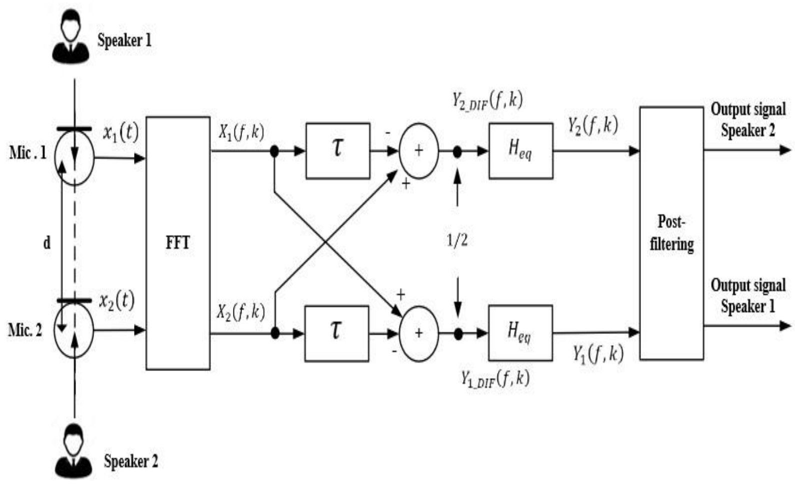

Figure 4: A certain differential microphone array

With the definition f, k is the frequency index and current considered frame. We denote the speech propagation of sound source is c $(343\ m/s)$, d is distance between two installed microphones, $\tau_0 = d/c$ is the sound delay, the direction of arrival of useful signal is θ, $\Phi_s = \pi f \tau_0 cos(\theta)$. The representation of two captured microphone array signals in the frequency-domain defined can be expressed as the following way:

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

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

With a defined time, delay τ is added, the directivity pattern of the processed signal is obtained by determined value of τ. DMA2 is used for extracting two directional different speakers at opposite directions. The output of DMA, which exploits subtraction signal between $X_1(f,k)$, $X_2(f,k)$ can be illustrated that:

$$Y_{1DIF}(f,k) = \frac{X_1(f,k) - X_2(f,k)e^{-j\omega\tau}}{2} \quad (3)$$

$$= jS(f,k)sin\left(\frac{\omega\tau_0}{2}\left(cos\theta + \frac{\tau}{\tau_0}\right)\right) \quad (4)$$

$$Y_{2DIF}(f,k) = \frac{X_2(f,k) - X_1(f,k)e^{-j\omega\tau}}{2} \quad (5)$$

$$= -jS(f,k)sin\left(\frac{\omega\tau_0}{2}\left(cos\theta - \frac{\tau}{\tau_0}\right)\right) \quad (6)$$

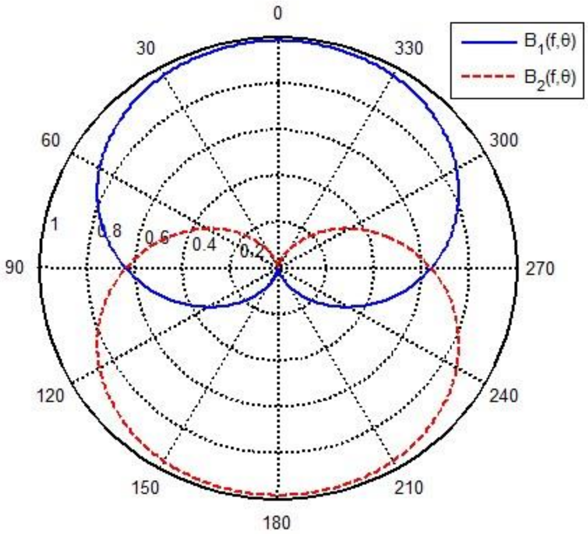

Figure 5: Shapes of beampattern. $\boldsymbol{d = 5(cm), f = 1500\ (Hz)}$

The resulting obtained beampattern toward two directional talkers have the high pattern, high resolution and can be expressed:

$$B_1(f,\theta) = \left|\frac{Y_1(f,k)}{S(f,k)}\right| = \left|sin\left(\frac{\omega\tau_0}{2}\left(cos\theta + \frac{\tau}{\tau_0}\right)\right)\right| \quad (7)$$

$$B_2(f,\theta) = \left|\frac{Y_2(f,k)}{S(f,k)}\right| = \left|sin\left(\frac{\omega\tau_0}{2}\left(cos\theta - \frac{\tau}{\tau_0}\right)\right)\right| \quad (8)$$

In the previous research, [10] proposed the use of an additive equalizer.

$$H_{eq}(f) = \begin{cases} 6 & 0\ (Hz) < f < 200\ (Hz) \\ \dfrac{1}{sin\left(\frac{\pi}{2}\frac{f}{f_c}\right)} & 200\ (Hz) < f \leq Fc \\ 1 & Fc < f \leq 2Fc \\ 0 & 2Fc < f \end{cases} \quad (9)$$

where $Fc = \frac{1}{4\tau_0}$.

The value of $H_{eq}(f)$ is limited with a determined threshold 12(dB). This equalizer ensures deriving desired signal.

So finally, the received signals are:

$$Y_1(f,k) = Y_{1DIF}(f,k) \times H_{eq}(f) \quad (10)$$
$$Y_2(f,k) = Y_{2DIF}(f,k) \times H_{eq}(f) \quad (11)$$

3. The suggested post - Filtering

The central ideal of suggested post - Filtering is based on the estimation of noise power. The MMSE estimator [21] is used for estimation a spectral gain:

$$G_{H_1}(f,k) = \frac{\sqrt{v(f,k)}}{\gamma(f,k)}\left[\Gamma\left(1+\frac{\alpha}{2}\right)M\left(-\frac{\alpha}{2};1;-v(f,k)\right)\right]^{\frac{1}{\alpha}} \quad (12)$$

Where $v(f,k) \triangleq \gamma(f,k)\frac{\xi(f,k)}{\xi(f,k)+1}$; $M(\alpha;c;x)$ is the confluent hypergeometric function. The author proposed the calculation of *a priori SNR* $\xi(f,k)$ and a posteriori SNR $\gamma(f,k)$ as the following equations:

$$\xi(f,k) = \frac{E[|Y_1(f,k)|^2]}{E[|Y_{2DIF}(f,k)|^2]} \quad (13)$$
$$\gamma(f,k) = \frac{E[|X_1(f,k)|^2]}{E[|Y_{2DIF}(f,k)|^2]} \quad (14)$$

With a defined appropriate value α, the obtained gain function by MMSE estimator can be used as an effective post - Filtering. In single - channel approach, the $G_{H_1}(f,k)$ is applied to gain the desired speech component while suppressing noise level. In the next section, this post - Filtering has the ability of preserving the target speaker while decreasing the background noise and enhancing the overall performance.

4. Experiments

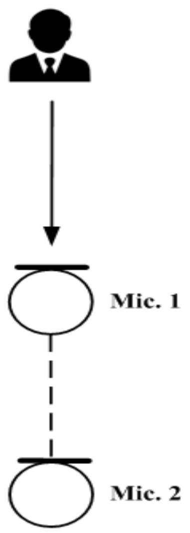

Figure 6: The illustrated experiment with DMA2

In this section, the author illustrated an experiment to enhance the performance of DMA2 in noise reduction. This experiment is conducted in a living room, where in presence of annoying background noise, diffuse noise field. The speaker in stand at $L = 2(m)$ to the DMA2. For further to rate the performance, an objective measurement [22] is used for calculating the speech quality of the previous work and additive post - Filtering. Two microphone array signals are sampled at $Fs = 16\ kHz$, and transformed in the frequency domain with these parameters: $NFFT = 512$, overlap 50%.

The waveform of the original microphone array signal can be expressed in Figure 7. From 0 - 1.4 (s), the speech component of desired talker exits, and from 1.6 - 3 (s), there are only direction noise source.

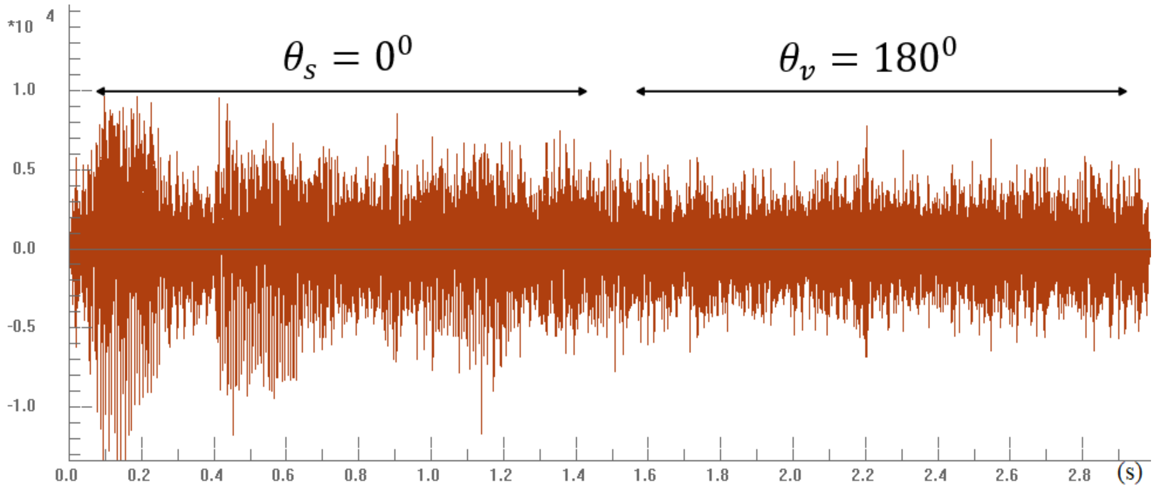

Figure 7: The waveform of microphone array signal

Using [23], the obtained waveform is shown in Figure 8.

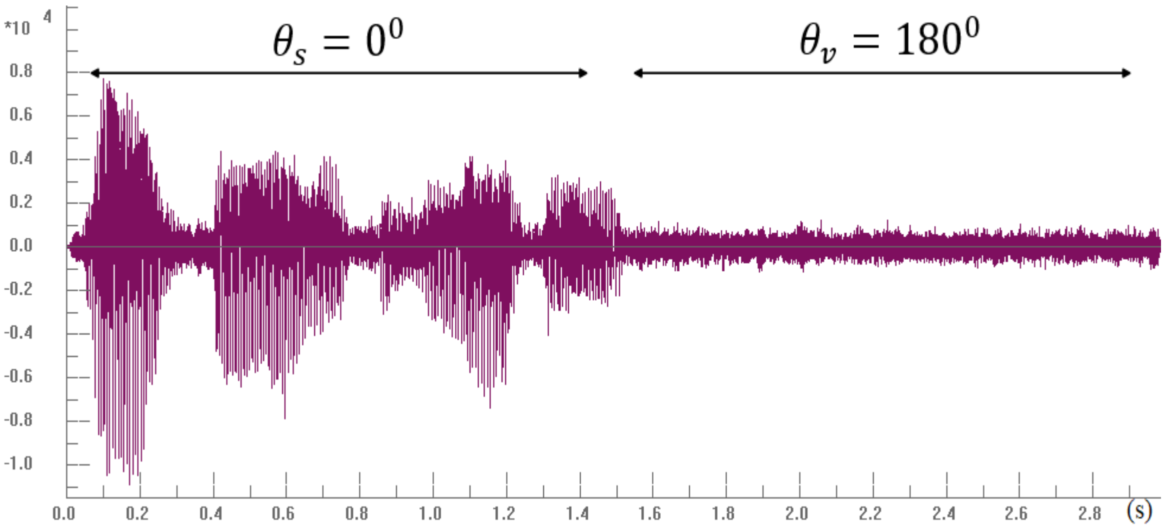

Figure 8: The waveform of processed by the author's previous research [23]

By using the post - Filtering, the effectiveness of noise reduction can be obtained. The processed signal is shown in Figure 9.

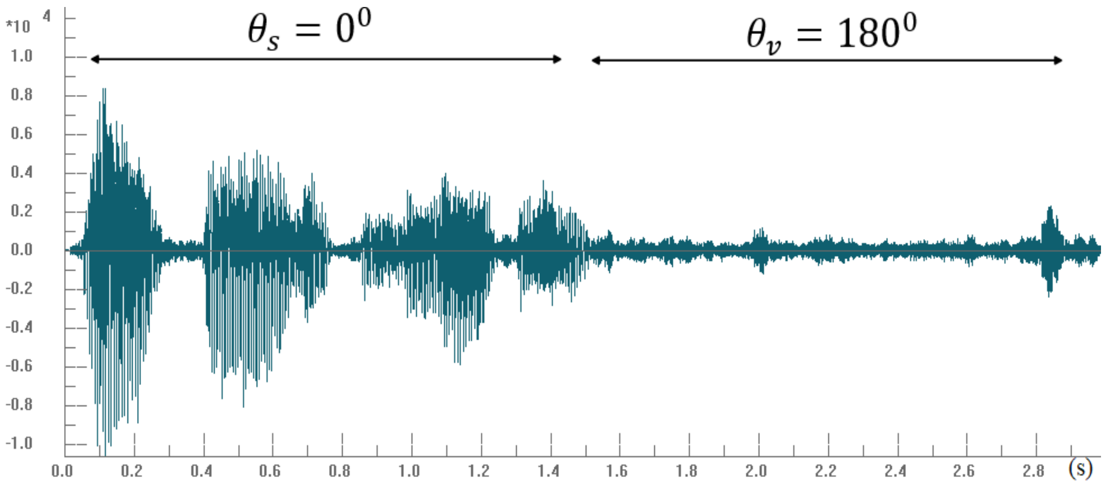

Figure 9: The waveform of processed by the additive post - Filtering

The overall energy of microphone array signal, the processed signals by [23], and post – Filtering is depicted in Figure 10.

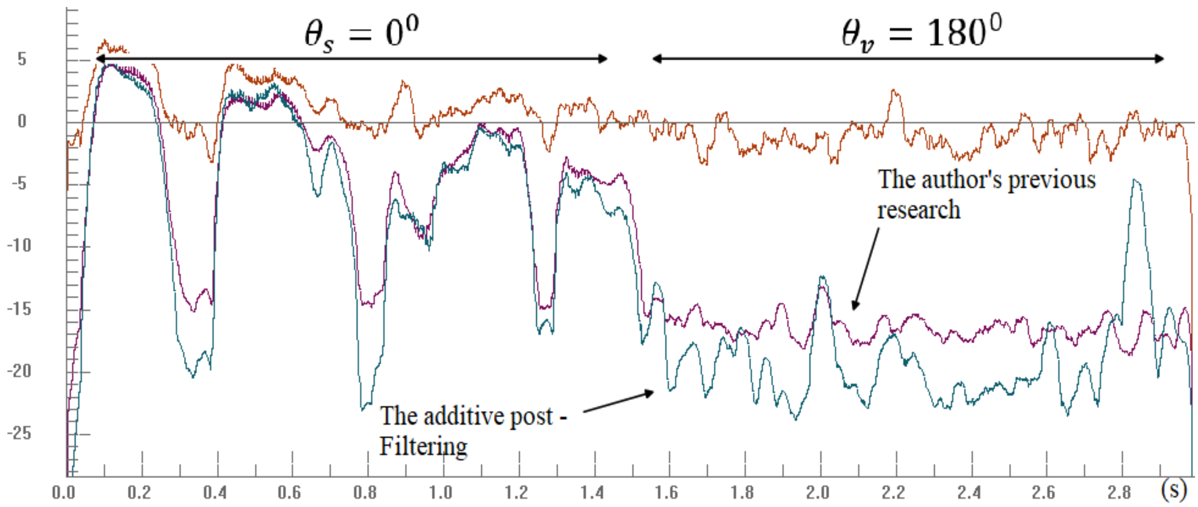

Figure 10: The energy of microphone array signal, the processed signal by the author's previous work and the additive post - Filtering

As we can see that, the advantage of post - Filtering is presented. The achieved noise reduction is to 5.5 (dB), and the speech quality in the terms of the signal - to - noise ratio (SNR) is increased from 3.2 (dB) to 5.7 (dB).

Table 1.

The signal - to - noise ratio (dB)

Method Estimation	Microphone array signal	The author's previous work	The additive post - Filtering
NIST STNR	4.8	21.0	24.2
WADA SNR	2.8	16.1	21.8

An essential problem is almost signal processing is achievement of more robust noise reduction. DMA2 is one of the most widely common installed in numerous speech applications, such as mobile phone, surveillance equipment, smart home, teleconferencing due to its compact. Therefore, an effective post - Filtering for DMA2 is an attractive research direction. In this section, the improvement of noise reduction has been confirmed. The numerical results show that proposed post - Filtering allows obtaining better performance in DMA2.

5. Conclusion

DMA2 has the capacity of compact arrangements, low size and promise high directional beampattern, high gain the output signal in comparison with other MA beamforming. However, DMA owns its drawback that even the complex noisy environment can corrupt its performance. For many several speech applications, decreasing speech distortion or noise suppression is always a considered problem. In this research, the author has presented and demonstrated a post - Filtering for enhancing the DMA2's performance in realistic recording scenario. The author has shown how the noisy component at certain direction can be suppressed, and the speech quality of DMA2's evaluation was increased in comparison with the author's previous work. This post - Filtering can be integrated into multi-microphone system, which use other different MA beamforming.

6. Acknowledgements

This research was supported by Digital Agriculture Cooperative. The author thanks our colleagues from Digital Agriculture Cooperative, who provided insight and expertise that greatly assisted the research.

7. References

[1] Albertini D., Bernardini A., Borra F., Antonacci F., Sarti A. Two-Stage Beamforming With Arbitrary Planar Arrays of Differential Microphone Array Units. IEEE/ACM Transactions on Audio, Speech, and Language Processing. Pp: 590 - 602. DOI: 10.1109/TASLP.2022.3231719.

[2] Huang W., Feng J. Robust Steerable Differential Beamformer for Concentric Circular Array With Directional Microphones // Proc 2022 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC). DOI: 10.23919/APSIPAASC55919.2022.9980184.

[3] JWang J., Yang F., Yang J. Insights Into the MMSE-Based Frequency-Invariant Beamformers for Uniform Circular Arrays. IEEE Signal Processing Letters. Pp(s): 2432 – 2436. DOI: 10.1109/LSP.2022.3224687.

[4] Ueno N., Kameoka H. Multiple Sound Source Localization Based on Stochastic Modeling of Spatial Gradient Spectral // Proc 2022 30th European Signal Processing Conference (EUSIPCO). DOI: 10.23919/EUSIPCO55093.2022.9909524.

[5] Yan L., Huang W., Kleijn W.B., Abhayapala T. D. Phase Error Analysis for First-Order Linear Differential Microphone Arrays // Proc 2022 International Workshop on Acoustic Signal Enhancement (IWAENC). DOI: 10.1109/IWAENC53105.2022.9914748.

[6] Itzhak G., Cohen I. Differential and Constant-Beamwidth Beamforming with Uniform Rectangular Arrays. 2022 International Workshop on Acoustic Signal Enhancement (IWAENC). DOI: 10.1109/IWAENC53105.2022.9914769.

[7] Huang G., Benesty J., Chen J. Fundamental Approaches to Robust Differential Beamforming With High Directivity Factors. IEEE/ACM Transactions on Audio, Speech, and Language Processing. Pp(s): 3074 - 3088. DOI: 10.1109/TASLP.2022.3209935.

[8] Jin J., Benesty J., Huang H., Chen J. On Differential Beamforming With Nonuniform Linear Microphone Arrays. IEEE/ACM Transactions on Audio, Speech, and Language Processing. Pp(s): 1840 - 1852. DOI: 10.1109/TASLP.2022.3178229.

[9] Yang X., Wei J. DMANET: Deep Learning-Based Differential Microphone Arrays for Multi-Channel Speech Separation. ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). DOI: 10.1109/ICASSP43922.2022.9747725.

[10] Wang X., Cohen I., Benesty J., Chen J. Study of the Null Directions on The Performance of Differential Beamformers. ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). DOI: 10.1109/ICASSP43922.2022.9746462.

[11] Tu Q., Chen H. Theoretical Lower Bounds on the Performance of the First-Order Differential Microphone Arrays With Sensor Imperfections. IEEE/ACM Transactions on Audio, Speech, and Language Processing. Pp(s): 785 – 801. DOI: 10.1109/TASLP.2022.3145317.

[12] Shi C., Du F., Liu C., Li H. Differential Error Feedback Active Noise Control With the Auxiliary Filter Based Mapping Method. IEEE Signal Processing Letters. Pp(s): 573 – 577. DOI: 10.1109/LSP.2022.3144839.

[13] Itzhak G., Cohen I., Benesty J. Robust Differential Beamforming with Rectangular Arrays // Proc 2021 29th European Signal Processing Conference (EUSIPCO). DOI: 10.23919/EUSIPCO54536.2021.9616085.

[14] Zhang P., Gao J., Bian Y., Huang Y. A Preliminary on the Sound Pickup for Unmanned Aerial Vehicles Using Differential Microphone Arrays. 2021 4th International Conference on Information Communication and Signal Processing (ICICSP). DOI: 10.1109/ICICSP54369.2021.9611979.

[15] Chen Z., Chen H., Tu Q. Sensor Imperfection Tolerance Analysis of Robust Linear Differential Microphone Arrays. IEEE/ACM Transactions on Audio, Speech, and Language Processing. Pp(s): 2915 – 2929. DOI: 10.1109/TASLP.2021.3110136.

[16] Yu G., Qiu Y., Wang N. A Robust Wavenumber-Domain Superdirective Beamforming for Endfire Arrays. IEEE Transactions on Signal Processing Page(s): 4890 – 4905. DOI: 10.1109/TSP.2021.3105754.

[17] Huang G., Wang Y., Benesty J., Cohen I., Chen J. Combined Differential Beamforming With Uniform Linear Microphone Arrays. ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). DOI: 10.1109/ICASSP39728.2021.9414189.

[18] Zhao X., Huang G., Benesty J., Chen J., Cohen I. On the Design of Square Differential Microphone Arrays with a Multistage Structure. ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). DOI: 10.1109/ICASSP39728.2021.9413759.

[19] Borra F., Bernardini A., Bertuletti I., Antonacci F., Sarti A. Arrays of First-Order Steerable Differential Microphones. ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). DOI: 10.1109/ICASSP39728.2021.9413476.

[20] Wang X., Huang G., Cohen I., Benesty J., Chen J. Robust Steerable Differential Beamformers with Null Constraints for Concentric Circular Microphone Arrays. ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). DOI: 10.1109/ICASSP39728.2021.9414119.

[21] Ephraim Y., Malah D. Speech enhancement using minimum mean-square error log-spectral amplitude estimator. IEEE Trans. ASSP, vol. ASSP-33, no. 2, pp. 443–445, 1985. DOI: 10.1109/TASSP.1985.1164550.

[22] https://labrosa.ee.columbia.edu/projects/snreval/

[23] Stolbov M., Tatarnikova M., The Q.T. (2018) Using Dual-Element Microphone Arrays for Automatic Keyword Recognition // Karpov A., Jokisch O., Potapova R. (eds) Speech and Computer. SPECOM 2018. Lecture Notes in Computer Science, vol 11096. Springer, Cham. https://doi.org/10.1007/978-3-319-99579-3-68.

[24] Ma D., Wang Y., He L., Jin M., Su D., Yu D. DP-DWA: Dual-Path Dynamic Weight Attention Network With Streaming Dfsmn-San For Automatic Speech Recognition. ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). DOI: 10.1109/ICASSP43922.2022.9746328.

[25] Xu X., Gu R., Zou Y. Improving Dual-Microphone Speech Enhancement by Learning Cross-Channel Features with Multi-Head Attention. ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). DOI: 10.1109/ICASSP43922.2022.9746359.

[26] Xiao Z., Chen T., Liu Y., Li J., Li Z. Keystroke Recognition with the Tapping Sound Recorded by Mobile Phone Microphones. IEEE Transactions on Mobile Computing. DOI: 10.1109/TMC.2021.3137229.

[27] Tan K., Zhang X., Wang D.L. Deep Learning Based Real-Time Speech Enhancement for Dual-Microphone Mobile Phones. IEEE/ACM Transactions on Audio, Speech, and Language Processing. DOI: 10.1109/TASLP.2021.3082318.

[28] Tan K., Zhang X., Wang D.L. Real-Time Speech Enhancement for Mobile Communication Based on Dual-Channel Complex Spectral Mapping. ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). DOI: 10.1109/ICASSP39728.2021.9414346.

[29] Shankar N., Bhat G.S., Panahi M.S. Real-time dual-channel speech enhancement by VAD assisted MVDR beamformer for hearing aid applications using smartphone. 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC). DOI: 10.1109/EMBC44109.2020.9175212.

[30] Li H., Zhang X., Gao G. Beamformed Feature for Learning-based Dual-channel Speech Separation. ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). DOI: 10.1109/ICASSP40776.2020.9054049.