Speech Source Separation Based on Dual – Microphone System

Quan Trong The[a]

[a] *Digital Agriculture Cooperative, Cau Giay, Ha Noi, Viet Nam*

Abstract

The fundamental intention of extracting the desired target speech source is saving the speech components of useful interesting signal at a certain direction while suppressing interference and unwanted signal, which come from other directions. Microphone array is one of the most widely technology used for improving speech enhancement in scenarios with multiple speech sources. Microphone arrays use a priori information of the direction of arrival to extract wanted signal, while eliminating background noise or different speech signal. In this article, dual-microphone system was used for recording two speech sources at two opposite directions; the authors proposed a method, which developed from the author's previous work for extracting the correct only target speaker at each direction, uses the ratio of power between two directions. The experiments have proven the capability of suggested method in comparison with the author's previous work.

Keywords

Microphone arrays, dual - microphone system, power, direction, speech source separation

1. Introduction

One of the most complicated tasks in speech communication system is speech signal acquisition without distortion and suppression background noise. This challenging problem still exist in real-life and requires scholar to investigate and solve it in such applications: speech recognition, hearing aids, telephony, human-machine interface, hands-free instrument. High quality of obtained output signal processing system is prerequisite order to restrict the effect of surrounding noise from a mixture of noisy signals and extract the clean desired speech (figure 1).

Figure 1: Multiple sound sources in human life

ICST-2023: Information Control Systems & Technologies, September 21-23, 2023, Odesa, Ukraine.
EMAIL: theqt@nongnghiepso.com (Q.T.The)
ORCID: 0000-0002-2456-9598 (Q.T.The)

Ambiguous speech enhancement is more clearly in a complex complicated acoustic surrounding environment when using a single-microphone approach, due to observed signal is corrupted by unwanted interference, diff use noise, coherence noise or complex condition. In order to obtain high speech quality and eliminate background noise, the technology of microphone array (MA) [1-4] has been developed for extracting satisfactory useful signal. This approach uses spatial information, such as distributed geometry MA or direction of arrival of desired signal for enhancing one signal in certain direction while suppressing other noise. Therefore, exploiting the properties of MA and environment background noise are the mainstream of current technology beamforming.

At present time, technology beamforming, which use microphone array, has three approaches, such as: fixed beamforming with fixed coefficient weights, adaptive beamforming algorithm, that adapts to observed array signals and post-filtering, which improve the speech enhancement. Adaptive beamforming and post-filtering have entrained many scholars to put some improvements for enhancing performance. These technologies have ability of saving desired signal and suppressing background noise while obtaining high speech quality. The combination with a justifiable post-filtering allows achieving an acceptable result in performance signal processing system and satisfied speech enhancement (figure 2).

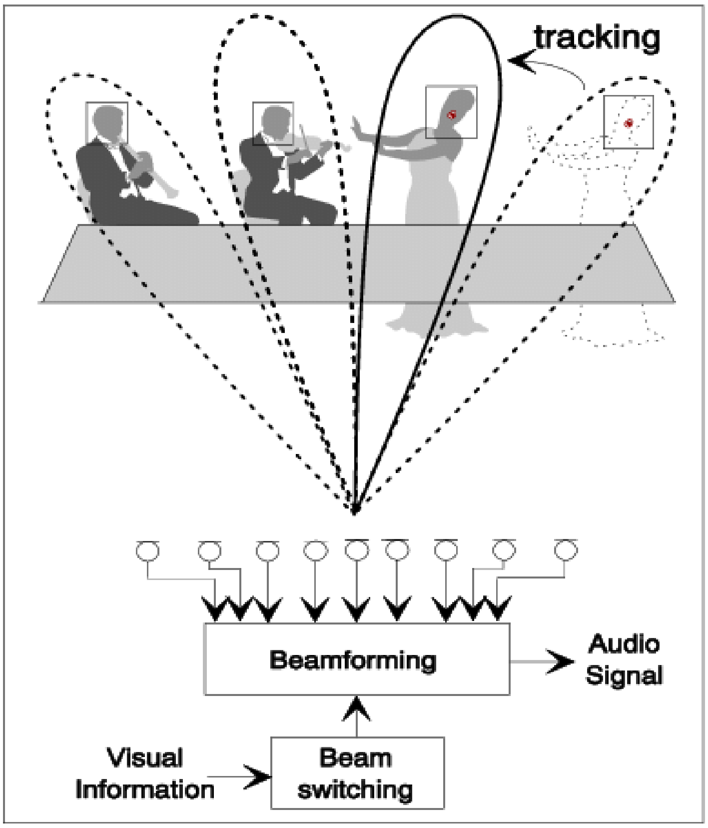

Figure 2: Microphone array beamforming technique

Because of possibly rapidly changing position of speaker, microphone array mismatch, or non-stationary acoustical environment; microphone array algorithm doesn’t sufficiently deliver effective performance in terms of speech distortion, signal-to-noise ratio.

Dual-microphone array (DMA2) [8-10], that is one of the most widely MA, has been exploited for many purposes, such as speech enhancement, estimation of direction of arrival (DOA), reverberation. Due to it’s compact, DMA2 exists in almost speech equipment. And problem of separation desired target speaker is an essence reality approach. In DMA2, the differential algorithm is often widely used to focusing on one certain direction along axis DMA2. Despite its simplicity, low computational performance, easily null forming towards noise source, the acoustic environments are often very complex and noisy. The noisy scenario severely reduces the effectiveness or speech quality of output signal. Therefore, the goal in this study is presented a new possible post-filtering to clearly pick-up speech in a certain direction.

In this paper, the authors considered problem of extracting desired target speaker when using DMA2 for recording two speakers, which their positions opposite each other. This work is the development of the author's paper for purpose of separation of target speaker. For more robust adaptive beamforming, the authors proposed a post-filtering, which based on estimation power each desired speech, for saving necessary speech while eliminating other. The experiments have illustrated the effectiveness and the ability of the proposed post-filtering in comparison with the author's previous work. The observed amplitude and spectrogram of processed signals has proven the capability and the received enhancement of the proposed additional post-filtering (figure 3).

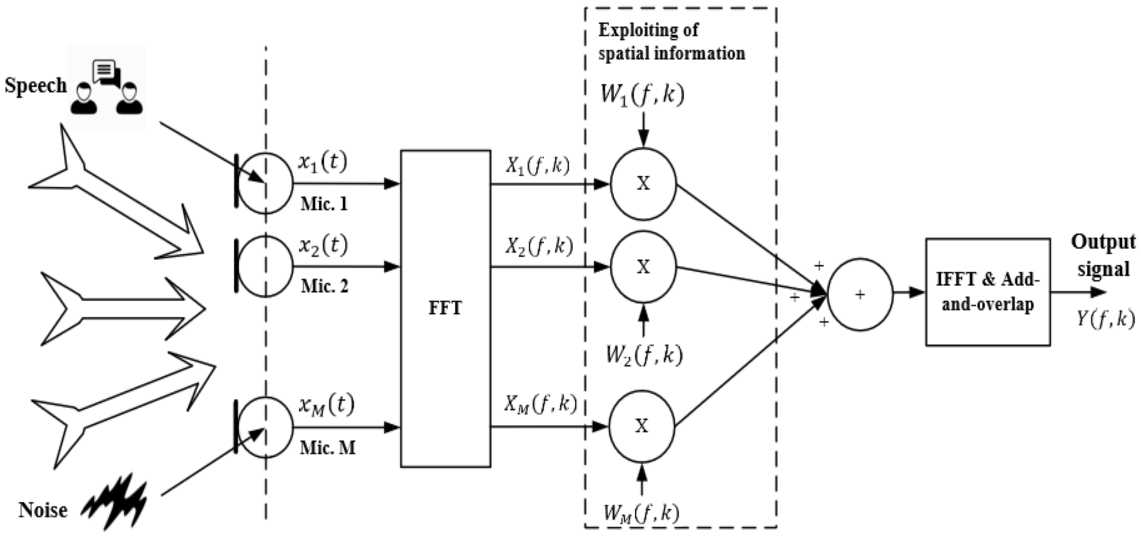

Figure 3: Microphone array structure on the frequency - domain

2. Separation of the Speakers

Natural Speech Processing (NSP) [18-19] methods has been utilized Automatic meeting transcription to aim extracting desired speaker in a long-duration recording speech signal. This challenging problem is considerable, because of using Speaker Identification and Diarization [20-21] and Automatic Speech Recognition (ASR) [22-24] technique to solving commonly speech applications. The existing available commercial method for speaker diarization can be categorized into two mainly approaches: (1) technology for distant speech acquisition and processing; (2) technology for close-talking processing. Distant acquisition processing tends to the scenarios, where a microphone array are distributed following an optimal geometry, situated at a long distance to record all attending desired speakers, and consequently interfering speech signals incoming from other directions and surrounding noise. Close-talking microphone uses each speech signal per each speaker to process and enhance the recording signal.

On the market, the promising automatic meeting transcription method are available, which ensure several perspective solutions for solving separation of speakers. Close-talking speech technique require less complicated and computational performance than distant speech processing. However, this situation almost is a trivial condition, that each attending speaker must be equipped each personal audio device.

On the other hand, distant speech processing is a useful method to pose complex task for speaker diarization. All negatived unwanted effect: mixture of different signals; attenuation due to microphone mismatch, microphone sensitivities, long distant; reverberation and influence of environmental factors on the accuracy, effectiveness of separation of speakers. The state-of-the-art of in separation of speaker has many attractive achievements on speech acquisition quality and recognition, there still exits limitation of prominent extraction methods fitted for separating. An Artificial Neural Network (ANN) is used for training model by using acoustical, lexical features. Distribution of microphone array also ensure an additional technique to eliminate background noise and extract spatial diversity of target speaker. The biometric properties provide a prominent approach of research.

The expected application of the proposed method in this paper to extract useful speech signal, that from a determined direction of target speaker in scenario of presence of two active speakers. The conversation is recorded by using a DMA2, which place in a table or a desk. The recording audio device DMA2 is appropriate for discussion, interviews, service provision; because of it's compact, convenient, and low complexity.

The necessary task of speech enhancement by using DMA2 in realistic situation is extraction and estimation speech or target speaker, which incoming from a certain direction, each active speaker per each utterance. In this article, the authors primarily address the conversation, which involves speech of two active speakers, their speech doesn't appear simultaneously but continuously.

The rest of this article is organized as follows: the next section is introduced DMA2 and differential algorithm, which is used to forming beampattern to desired speech source. In section IV, the authors suggested using the estimation of speech power to form an additional post-filtering. Section V, the scheme and comparison's performance are shown. Final section VI, conclusion, and the author's future work.

3. Differential microphone array

Differential algorithm [5-7,11-12] is one of the most popular algorithms for processing microphone array signals. A structure of DMA2 is presented in Figure 4. Two microphones arranged toward two different speakers. DMA2 allow steering beampattern to a certain direction for obtaining high diversity and suppressing signals, interference, or noise, which come from other directions. Speech enhancement in DMA2 [13-15] is easily implemented cause of incorporating a suitable equalizer or post-filtering for attenuating background noise and saving the desired speech source.

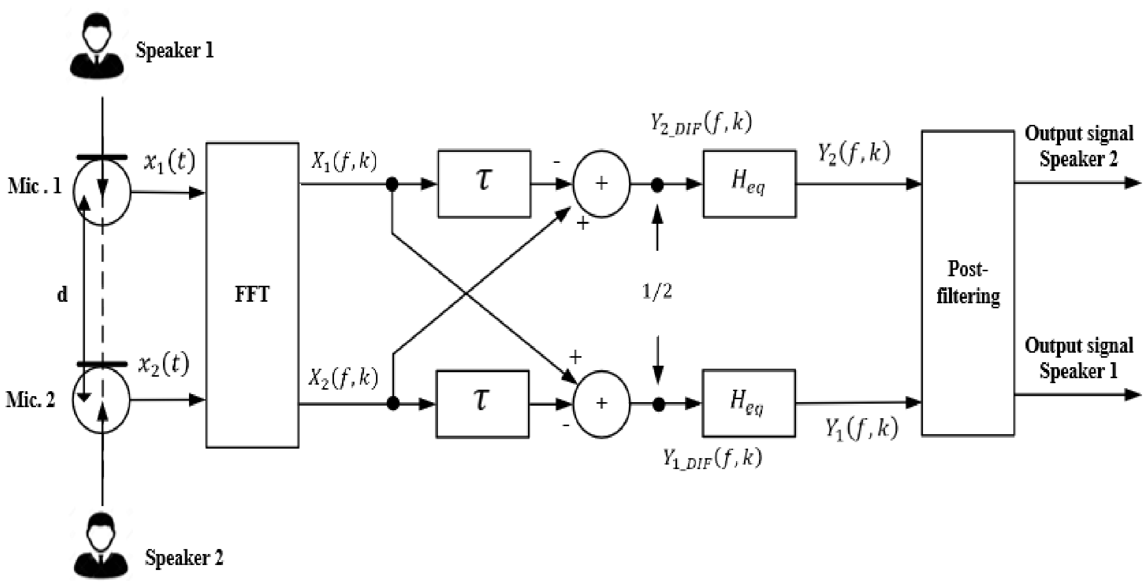

Figure 4: The scheme of proposed dual - microphone system

Here, the sound speech is $c(343\ m/s)$, distane between microphones is d, $\tau_0 = d/c$, useful signal or speech source propagates with direction θ_s, $\Phi_s = \pi f \tau_0 cos\theta_s$, f, k are the current frequency and current frame. Two noisy observed microphone signals in STFT-domain are represented as:

$$X_1(f,k) = S(f,k)e^{j\Phi_s} \quad (1)$$
$$X_2(f,k) = S(f,k)e^{-j\Phi_s} \quad (2)$$

The adjustable value τ is added for controlling null - pattern when extracting the useful signal. So DMA2 is used for reducing surrounding noise and extracting the desired signal. Two output signals, which derived from operator subtraction signal between two recored microphone signals, are calculated by the following equation:

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

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

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

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

Equations (3,4,5,6) yield two steered beampattern, which can be expressed (figure 5):

$$B_1(f,\theta) = \left|\frac{Y_{1DIF}(f,k)}{S(f,k)}\right| = \left|sin\left(\frac{\omega\tau_0}{2}\left(cos\theta + \frac{\tau}{\tau_0}\right)\right)\right| \tag{7}$$

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

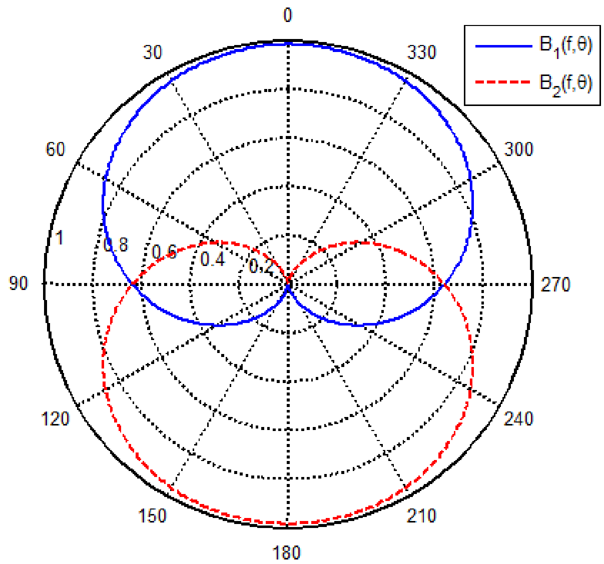

Figure 5: Shapes of beampattern. d = 5 (cm), f = 1500 (Hz)

An additional equalizer [17] is added for improving the received signal:

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

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 the output two desired target speakers are:

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

4. The proposed Post-Filtering

The designed post-filtering is proposed for microphone array system for increasing implementation in various types of background noise. Post-filtering method often uses a priori information of environment's characteristic or coherence between microphones. The authors proposed a post-filtering, which uses a determined power speech at a determined direction for forming post-filtering.

We have two directions: $\theta = \theta_1 = 0^0$ and $\theta = \theta_2 = 180^0$. With certain two steering vector $\boldsymbol{D}_s(f,\theta_1) = [e^{j\Phi_{s1}} \quad e^{-j\Phi_{s1}}]^T$ and $\boldsymbol{D}_s(f,\theta_2) = [e^{j\Phi_{s2}} \quad e^{-j\Phi_{s2}}]^T$, where $\Phi_{s1} = \pi f \tau_0 cos(\theta_1), \Phi_{s2} = \pi f \tau_0 cos(\theta_2)$, we can easily calculated power speech at two directions as the following equations:

$$\phi_{s1s1}(f,k) = \left(\boldsymbol{D}_s^H(f,\theta_1)\boldsymbol{P}_{XX}^{-1}(f,k)\boldsymbol{D}_s(f,\theta_1)\right)^{-1} \tag{12}$$
$$\phi_{s2s2}(f,k) = \left(\boldsymbol{D}_s^H(f,\theta_2)\boldsymbol{P}_{XX}^{-1}(f,k)\boldsymbol{D}_s(f,\theta_2)\right)^{-1} \tag{13}$$

Where $\boldsymbol{P}_{XX}(f,k) = E\{\boldsymbol{X}(f,k)\boldsymbol{X}^*(f,k)\}$ can be determined as:

$$\begin{Bmatrix} P_{X_1X_1}(f,k) * 1.001 & P_{X_1X_2}(f,k) \\ P_{X_2X_1}(f,k) & P_{X_2X_2}(f,k) * 1.001 \end{Bmatrix} \tag{14}$$

Where $P_{X_iX_i}(f,k), P_{X_iX_j}(f,k), i,j \in \{1,2\}$ calculated as:

$$P_{X_iX_j}(f,k) = (1-\alpha)P_{X_iX_j}(f,k-1) + \alpha X_i(f,k)X_j^*(f,k) \tag{15}$$

Where α is the smoothing parameter, which in the range {0 ... 1}.

The suggested post-filtering for enhancing speech quality is formulated as:

$$H_{PF1}(f,k) = \frac{\phi_{s1s1}(f,k)}{\phi_{s1s1}(f,k) + \phi_{s2s2}(f,k)} \tag{16}$$

$$H_{PF2}(f,k) = \frac{\phi_{s2s2}(f,k)}{\phi_{s1s1}(f,k) + \phi_{s2s2}(f,k)} \tag{17}$$

And the final signals of interest are given by:

$$Y_{out1}(f,k) = Y_1(f,k) \times H_{PF1}(f,k) \tag{18}$$
$$Y_{out2}(f,k) = Y_2(f,k) \times H_{PF2}(f,k) \tag{19}$$

In the next section, the authors will compare the effectiveness of using the proposed post-filtering (DIF-PF-Pro) and the known algorithm [17].

5. Experiments

In this section, a DMA2 system was placed in an anechoic room to capture two microphone array signals from two independent speakers, which stand opposite each other. The distance from speaker to

DMA2 is 3 (m), the distance between microphones is $d = 2{,}5\ (cm)$. So, the directions of two interest useful signals are opposite (figure 6).

All two noisy microphone array signals are sampled 16 (kHz). For processing the method [17] and suggested method (DIF-PF-Pro), a Hamming window, overlap 50%, $NFFT = 512$ and a constant smoothing parameter $\alpha = 0{,}1$ were chosen to compute PSD and additive post-filtering.

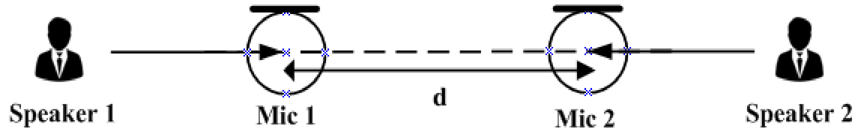

Figure 6: The scheme of experiment

For estimating speech quality, an objective measure NIST SNR [16] was utilized to verify the effectiveness of the proposed method (DIF-PF-Pro) in comparison with known differential algorithm [17] (figure 7).

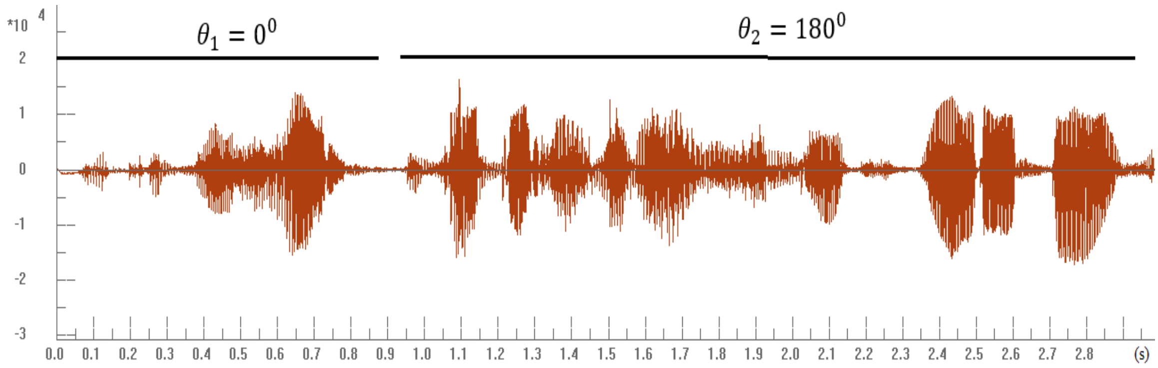

Figure 7: The original signal, which captured with microphone array

Figure 7 shows the captured microphone array signals with two directions of two desired speakers. The next figure presents the obtained signal by using method [17] for extracting speech of speaker, who stand at direction $\theta_2 = 180\ (deg)$. As we can see that, the large amount of first speaker's speech still exits.

By using the suggested method, the target useful of second speaker was saved while reducing the amount of first speaker's speech. Result of DIF-PF-Pro is shown in Figure 9, that verifies the effectiveness of the proposed additive post-filtering.

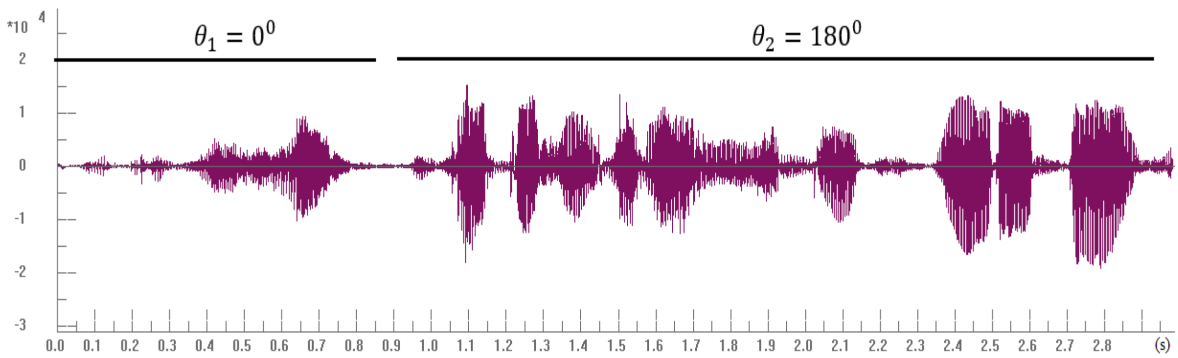

Figure 8: The obtained signal by [17].

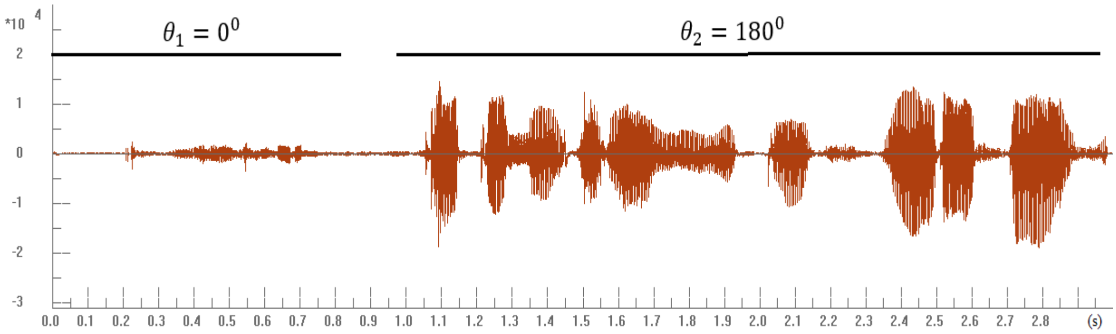

Figure 9: The resulting signal by using the proposed method [DIF-PF-Pro].

6. Conclusion

The goal of microphone array is to increase the evaluation of speech enhancement system with the purpose of saving useful target speech signal while eliminating interference and noise, which come from other directions. Therefore, improvement of MA's performance is prerequisite order in front-end almost speech applications. A lot of noise reduction and preserving speech signal methods are investigated, evaluated. Because of undetermined noise source, beamforming technique can not sufficiently exclude almost all noise component; so, an additive post-filtering can help further keep speech in desired direction and reduce considered noise component.

An important essence in signal processing is exploiting' characteristic of dual-microphone system to obtain useful speech target speaker and improve speech enhancement. The dual-microphone system always provides a high noise reduction, high spatial diversity for extracting desired speech in noisy complicated conditions. This paper proposes a new technology post-filtering to improve the capability of extracting the correct desired target speaker. With priori information of direction of arrival, the suggested method has proven in increasing performance dual-microphone system in comparison with other work. While the components of useful signal at a certain direction are saved, and the other speech source is suppressed. Dual - microphone is a basic element in microphone array, so the proposed method can be further integrated into system multi-microphone for resolving more complicated problem speech enhancement.

7. References

[1] Microphone Arrays / ed. by M. Brandstein, D. Ward. Heidelberg, Germany: Springer-Verlag, 2001. doi:10.1007/978-3-662-04619-7.

[2] J. Benesty, J. Chen, Y. Huang, Microphone Array Signal Processing, Berlin, Germany: Springer-Verlag, 2008. doi:10.1007/978-3-540-78612-2.

[3] J. Benesty, J. Chen, C. Pan, Fundaments of Diff erential Beamforming. Springer, 2016. doi:10.1007/978-981-10-1046-0.

[4] J. Benesty, I. Cohen, J. Chen, Fundamentals of Signal Enhancement and Array Signal Processing, Wiley, IEEE Press, 2017.

[5] G.W. Elko, A.-T.N. Pong, A steerable and variable fi rst-order diff erential microphone array, in: Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP). 1997, p. 223 – 226. doi: 10.1109/ICASSP.1997.599609.

[6] M. Buck, M. Rößler, First order diff erential microphone arrays for automotive applications, in: Proc. 7th International Workshop on Acoustic Echo and Noise Control, IWAENC. 2001, pp. 19–22.

[7] M. Buck, Aspects of first-order diff erential microphone arrays in the presence of sensor imperfections. European Transactions on Telecommunications, 13 2 (2002) 115–122. doi: 10.1002/ett.4460130206.

[8] H. Zhang, Q. Fu., Y. Yan, A frequency domain approach for speech enhancement with directionality using compact microphone array, in: Proc. 9th Annual Conference of the International Speech Communication Association (INTERSPEECH 2008), 2008. pp. 447–450.

[9] S. Fischer, K. Simmer, An adaptive microphone array for hands–free communication, in: Proc. IWAENC-95,1995, pp. 1–4.
[10] J. Bitzer, K. Kammeyer, K. Simmer, An alternative implementation of the superdirective beamformer, in: Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPLAA'99),1999, pp.7–9. doi: 10.1109/ASPAA.1999.810836.
[11] M. Buck, T. Wolff, T. Haulick, G. Schmidt, A compact microphone array system with spatial post-filtering for automotive applications, in: Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2009), 2009, pp. 221–224. doi: 10.1109/ICASSP.2009.4959560.
[12] Microphone Array Beamforming. Application Note AN–1140. URL: http://www.invensense.com/wp- content/uploads/2015/02/Microphone-Array-Beamforming.pdf.
[13] T. Lotter, P. Vary, Dual-channel speech enhancement by superdirective beamforming. Eurasip Journal on Applied Signal Processing, (2006) 1–14. doi: 10.1155/ASP/2006/63297.
[14] J. Bitzer, K.U. Simmer, K.-D. Kammeyer, Multi-microphone noise reduction techniques for hands-free speech recognition – a comparative study, in Proc. of Workshop on Robust Methods forSpeech Recognition in Adverse Conditions (ROBUST 99), 1999, pp. 171–174.
[15] Y. Ji, J. Byun, Y. Park, Coherence-based dual-channel noise reduction algorithm in a complex noisy environment, in: INTERSPEECH 2017. Stockholm, Sweden, August 20–24, 2017, doi:10.21437/INTERSPEECH.2017-1464.
[16] Snreval. URL: https://labrosa.ee.columbia.edu/projects/snreval/.
[17] M. Stolbov, M. Tatarnikova, The Q.T. Using dual-element microphone arrays for automatic keyword recognition. Lecture Notes in Computer Science (including subseries Lecture Notes in Artifi cial Intelligence and Lecture Notes in Bioinformatics). 11096 (2018) 667-675. DOI:10.1007/978-3-319-99579-3-68.
[18] S. Liu, Y. Yao, B. Liu, R. Zhu, J. Ren, Multi-band discriminant speech synthesis analysis based on Natural Language Processing, in: 2022 7th International Conference on Intelligent Computing and Signal Processing (ICSP), 2022, pp. 1933-1936. doi: 10.1109/ICSP54964.2022.9778506.
[19] Y. Lu, A. Harati, T. Rutowski, R. Oliveira, P. Chlebek, E. Shriberg, Robust Speech and Natural Language Processing Models for Depression Screening, in: 2020 IEEE Signal Processing in Medicine and Biology Symposium (SPMB), Philadelphia, PA, USA, 2020, pp. 1-5. doi: 10.1109/SPMB50085.2020.9353611.
[20] Z. Li, J. Whitehill, Compositional Embedding Models for Speaker Identifi cation and Diarization with Simultaneous Speech From 2+ Speakers, in: ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Toronto, ON, Canada, 2021, pp. 7163-7167. doi: 10.1109/ICASSP39728.2021.9413752.
[21] F. Roumiassa, Z. Chelali, Speaker Identifi cation and Verifi cation System for Arabic and Berber Language, in: 2020 1st International Conference on Communications, Control Systems and Signal Processing (CCSSP), El Oued, Algeria, 2020, pp. 242-247. doi: 10.1109/CCSSP49278.2020.9151633.
[22] K. Chu, L. Collins, B. Mainsah, Using Automatic Speech Recognition and Speech Synthesis to Improve the Intelligibility of Cochlear Implant users in Reverberant Listening Environments, in: ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Barcelona, Spain, 2020, pp. 6929-6933. doi: 10.1109/ICASSP40776.2020.9054450.
[23] M. Yousefi, J. Hansen, Speaker Conditioning of Acoustic Models Using Affine Transformation for Multi-Speaker Speech Recognition, in: 2021 IEEE Automatic Speech Recognition and Understanding Workshop (ASRU), Cartagena, Colombia, 2021, pp. 283-288. doi: 10.1109/ASRU51503.2021.9688231.
[24] M. Soleymanpour, M. T. Johnson, R. Soleymanpour, J. Berry, Synthesizing Dysarthric Speech Using Multi-Speaker Tts For Dysarthric Speech Recognition, in: ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Singapore, Singapore, 2022, pp. 7382-7386. doi: 10.1109/ICASSP43922.2022.9746585.
[25] D. Sundström, J. Lindström, A. Jakobsson, Recursive Spatial Covariance Estimation with Sparse Priors for Sound Field Interpolation, in: Proc 2023 IEEE Statistical Signal Processing Workshop (SSP), Hanoi, Vietnam, 2023, pp. 517-521, doi: 10.1109/SSP53291.2023.10208010.

[26] A. Sarafnia, M. O. Ahmad, M. N. Swamy, A Beam Steerable Speaker Tracking-based First-order Differential Microphone Array, in: Proc *2023* 21st IEEE Interregional NEWCAS Conference (NEWCAS), Edinburgh, United Kingdom, 2023, pp. 1-5, doi: 10.1109/NEWCAS57931.2023.10198069.
[27] T. Liu, Z. Lu, T. Fei, A Hybrid Reverberation Model and Its Application to Joint Speech Dereverberation and Separation. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 31 (2023) 3000-3014. doi: 10.1109/TASLP.2023.3301227.
[28] J. Jin, J. Benesty, J. Chen, G. Huang, Differential Beamforming From a Geometric Perspective. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 31 (2023) 3042-3054. doi: 10.1109/TASLP.2023.3301245.
[29] Z. Wang, W. Zou, H. Su, Y. Guo, D. Li, Multiple Sound Source Localization Exploiting Robot Motion and Approaching Control. IEEE Transactions on Instrumentation and Measurement, (2021). doi: 10.1109/TIM.2023.3298406.
[30] D. M. Baruah, S. Senchowa, B. K. Das, Impact of Reference Microphone Selection on Binaural Cue Preservation for Directional Sources in Multi-microphone Hearing Aids, in: *Proc* 2023 IEEE Guwahati Subsection Conference (GCON), Guwahati, India, 2023, pp. 1-5. doi: 10.1109/GCON58516.2023.10183407.
[31] G. Li, Audio-Visual End-to-End Multi-Channel Speech Separation, Dereverberation and Recognition, IEEE/ACM Transactions on Audio, Speech, and Language Processing, 31 (2017) 2707-2723. doi: 10.1109/TASLP.2023.3294705.
[32] J. Wang, H. He, Y. Yu, Y. Zhou, L. Zhang, Robust Time-Delay Estimation for Speaker Localization Using Mutual Information Among Multiple Microphone Signals. IEEE Sensors Journal, (2023). doi: 10.1109/JSEN.2023.3293499.
[33] S. Chen, R. Tan, Z. Wang, X. Tong, K. Li, VoiceMap: Autonomous Mapping of Microphone Array for Voice Localization, IEEE Internet of Things Journal, (2021). doi: 10.1109/JIOT.2023.3294937.
[34] X. Fang, Y. Liu, X. Zhang, Q. Sun, Online Monitoring of Converter Station Using an Acoustic Signal Analysis Method Based on The Mobile Microphone Array, in: Proc 2023 IEEE 6th International Electrical and Energy Conference (CIEEC), Hefei, China, 2023, pp. 3918-3922. doi: 10.1109/CIEEC58067.2023.10166859.
[35] Y. Shen Research on the Design of Omni-Directional Mobile Robot for Sound Seeking, Positioning and Navigation, in: Proc 2023 4th International Conference for Emerging Technology (INCET), Belgaum, India, 2023, pp. 1-6. doi: 10.1109/INCET57972.2023.10170055.
[36] F. Yang, R. Song, A Review of Sound Source Localization Research in Three-Dimensional Space, in: Proc 2023 IEEE 12th Data Driven Control and Learning Systems Conference (DDCLS), Xiangtan, China, 2023, pp. 579-584. doi: 10.1109/DDCLS58216.2023.10165974.
[37] J. Wang, F. Yang, J. Yang, A General Approach to the Design of the Fractional-Order Superdirective Beamformer, IEEE Transactions on Circuits and Systems II: Express Briefs, (2022). doi: 10.1109/TCSII.2023.3287918.
[38] Y. Wakabayashi, K. Yamaoka, N. Ono, Sound Field Interpolation for Rotation-Invariant Multichannel Array Signal Processing, IEEE/ACM Transactions on Audio, Speech, and Language Processing, 31 (2023) 2286-2298. doi: 10.1109/TASLP.2023.3282098
[39] C. Zhang, J. Liu, H. Li, X. Zhang, Neural Multi-Channel and Multi-Microphone Acoustic Echo Cancellation, IEEE/ACM Transactions on Audio, Speech, and Language Processing, 31 (2023) 2181-2192. doi: 10.1109/TASLP.2023.3282103.
[40] J. Wang, F. Yang, J. Yang, A Perspective on Fully Steerable Differential Beamformers for Circular Arrays, IEEE Signal Processing Letters, 30 (2023) 648-652. doi: 10.1109/LSP.2023.3280852.
[41] X. Zhao, G. Huang, J. Chen, J. Benesty, Design of 2D and 3D Differential Microphone Arrays With a Multistage Framework, IEEE/ACM Transactions on Audio, Speech, and Language Processing, 31 (2023) 2016-2031. doi: 10.1109/TASLP.2023.3278182.