In-car speech enhancement based on source separation technique

Jeyasingh Pathrose, Mohamed Ismail
Department of Electronics and Communication Engineering, B. S. Abdur Rahman Crescent Institute of Science and Technology, Chennai, India

Article Info

Article history:

Received Feb 20, 2023
Revised Apr 24, 2023
Accepted May 28, 2023

Keywords:

Inside the car
Microphone array
Noise reduction
Source separation method
Speech enhancement

ABSTRACT

This work addresses the solution of localizing and enhancing hands-free speech inside the car environment. Cars have different types of sounds from outside, co-passengers dialogue and noise. To provide better-quality speech, a microphone array-based beamforming technique is used. This research work proposes the method for selected source localization, source separation, and enhancement. An estimation of the direction of arrival (DOA) to localize the signal direction and preferred direction is selected for speech enhancement. The spiral and sine-cosine algorithm (SSCA) algorithm is combined with an adaptive least mean square to adapt the system for different environments. The algorithm is implemented in hardware and tested in a real-time car environment. The results showed significant improvement in signal-to-noise ratio (SNR) of 5.2 dB and perceptual evaluation of speech quality (PESQ) of 2.3. Finally, the model is fine-tuned for the car to get better quality. The proposed technique is efficient, and results are compared with existing methods.

Corresponding Author:

Jeyasingh Pathrose
Department of Department of Electronics and Communication Engineering, B. S. Abdur Rahman Crescent Institute of Science and Technology
Chennai, 600048, India
Email: gpjs@hotmail.com

1. INTRODUCTION

Speech quality is the major requirement for any communication. The speech signal is now used in a variety of systems including speaker location identification, system control using voice commands, speech-to-text systems, voice over Internet protocol (VoIP), speech recognition, interactive voice response system (IVRS) services, and other communication activities [1], interaction [2], sound aids [3], and coding of speech all require speech preprocessing and enhancement (SE) [4]. The SE is a difficult operation when the noisy environment also has the same speech spectrum [5]. The quality of the speech signal should not be sacrificed while designing a speech preprocessing system. However, speech signals can be corrupted in practice due to a variety of disturbances such as echo, noise in the background, babbling noise, clipping, and so on. Speech enhancement technology [6] can improve not just the signal-to-noise ratio (SNR) and audio perception of collected speech as well as the resilience of the speech. As a result, speech enhancement in noisy environments has gotten a lot of attention.

Engine noise and other noise sources such as airflow from air conditioners, air purifiers, outside wind noise, and environmental noise can affect the speech quality when using in-car speech applications. Inside the car, the reflection of speech waves can cause interference. The speech signals are picked up by the array of microphones and the microphones are placed in the front seat headrest position.

Beamformer arrays can be used in hands-free car systems and in-car speech recognition systems. Microphone array processing focuses on speech improvement and localization, particularly in noisy or reverberant environments [7]. A microphone array is used in the car to increase voice communication quality [8]. A microphone array may gather data in the spatial domain as well as the temporal and frequency domains. In this paper, two microphone array is used for noise reduction. The spiral and sine-cosine algorithm (SSCA) method provides noise reduction in the speech spectrum and enhances speech. For practical microphones are placed within 15 cm. These correlations reduce the algorithm's ability to suppress noise resulting in harmonic noise.

The main aim of the proposed system is to focus on the quality enhancement of speech in the car by using a source separation-based adaptive learning management system (LMS). It is possible to handle multiple speech signals as well as a wide range of interfering effects. Separating as well as improving the signals as a result of the proposed method is beneficial to speech recognition in the future. Source separation is the technique of extracting signals from their combination as a source without an existing understanding of the mixing models.

The framework of this research work is organized as the following: the literature survey is described in section 2, and the problem about the array position inside the car is addressed in section 3. The proposed source separation for the car is given in section 4, outcomes are shown in section 5. Finally, section 6 summarizes the conclusion.

2. LITERATURE SURVEY

The survey shows the details of research conducted on localization and speech enhancement. This paper reviews and provides an overview of developments in microphone beamforming. Gentet *et al.* [9] deal with the speech enhancement algorithm that increases the SNR. The result showed the technique for noise reduction and low computational complexity. Speech intelligibility optimization problem with a fixed perceived loudness restriction is a major drawback. Alkaher and Cohen [10] presented the dual microphone speech enhancement for enhancing speech communication in cars. The Pareto optimization decreases the overall speech distortion and relative gain reduction. The result demonstrates the dual-microphone system enhanced howling detection sensitivity. Lei *et al.* [11] presented a wavelet analysis and blind source separation to enhance the performance of the voice control system. The experimental outcome demonstrates that the suggested technique successfully separates various speech signals in a demanding automotive setting without the need for prior information. Low performance of vehicle speech recognition. Wang *et al.* [12] presented the speaker identification algorithm was used to test the speech augmentation using improved nonnegative matrix factorization (ImNMF). The results showed that the suggested ImNMF can significantly improve noise speech while also increasing the robustness of the signal in electric car noise environment. Tao *et al.* [13] presented the enhanced speech signal source localization and enhancement system which reduces microphone cost and reduces complexity. Dual-microphone sound algorithm effectively identifies the sound location, as well as the speech quality improvement, is more resilient and adaptive than the previous method, according to experimental data. Li *et al.* [14] presented the car speech enhancing method based on the distributed microphone. The dispersed microphone improves speech that has been distorted by noise in the car. The result showed that the suggested technique is more adaptable, and it significantly improves the SNR. Speech enhancement using distributed microphones is unusual. Panda [15] presented the stacked recurrent neural network used to create a robust speech enhancement system. The traffic noise is canceled in the car speech recognition system. The simulated result shows that the suggested model has higher complexity with an optimum number of layers. Qian *et al.* [16] presented the car speech-enhancing system based on a combination of a deep belief network and wiener filtering. The deep belief network parameters are optimized by using the quantum particle swarm optimization algorithm. The result showed that the method can effectively eliminate the noise signal of the input signal and enhance the speech signal. Krause *et al.* [17] presented that fast learning is possible using echo state networks and concepts for a variety of applications, including speech recognition, and detecting car driving actions.

Several works consider the difficulties in the optimization of speech intelligibility. Unstable speech signal, the low performance of microphone, cost and design requirement, and other factors were proposed, none of them is successful in speech enhancement system to overcome the above challenges this research proposed a novel source separation based adaptive LMS and it is a detailed process is presented in the next section.

3. MICROPHONE POSITION INSIDE THE CAR

Positioning a microphone array inside a car needs to obey some important requisites related to response quality, installation costs, and user-friendliness. Several works dealing with in-car speech separation and enhancement adopt two main microphone dispositions: microphones spread throughout the whole car interior. Although presenting some interesting results, this type of disposition presents some drawbacks that can make it difficult for its adoption in commercial systems. First, spreading the microphones throughout the car interior makes the receivers experiment with different levels of noise, produced by the horn sound, car engine sound, driver's voice, the interruption from other passengers, and so on. This can lead to degradation in the speech signal response. Moreover, considering that the objective is to enhance the speech of the passenger, microphones are placed on the headrests of the front seats. The driver position, front-seat passenger position, and back-seat passenger position are shown in Figure 1.

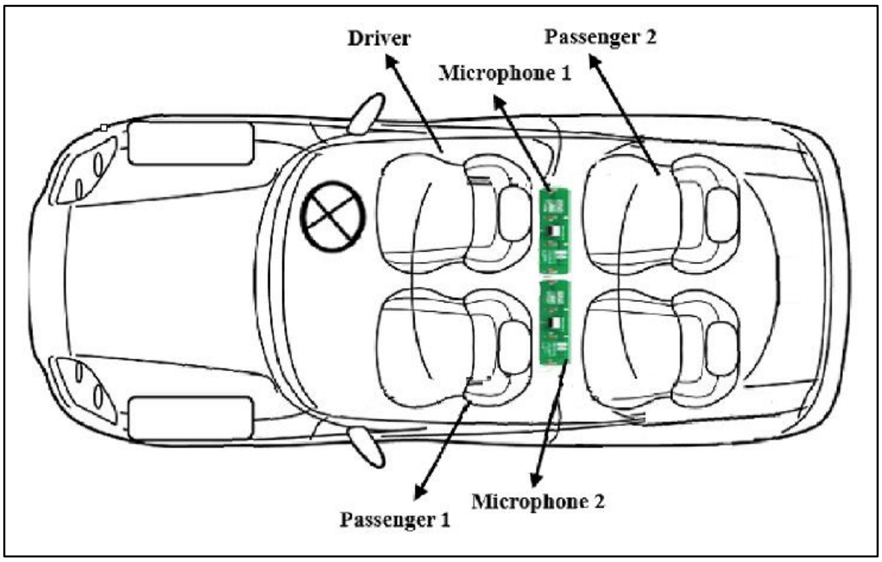


Figure 1. Microphone position inside the car

Let X and Y denote the number of source signals and microphones, respectively. The signals from the source are then referred to as (1).

$$\mathrm{si}[\mathrm{t}] = (\mathrm{si}_1[\mathrm{t}], \mathrm{si}_2[\mathrm{t}], \dots, \mathrm{s}_{\mathrm{iN}}[\mathrm{t}])^{\mathrm{T}} \tag{1}$$

The discrete time is denoted by t and the signals received can be labelled as (2).

$$\mathrm{x}[\mathrm{t}] = x_1[t], x_2[t], \dots, x_M[t])^T \tag{2}$$

The combinations at the receiving end constitute a more complicated mixing process known as convolution mixing namely a sum of signals with various weights due to the delay and echo between the microphone and the source.

$$\mathrm{x}_{\mathrm{m}}[\mathrm{t}] = \sum_{\mathrm{n}-1}^{\mathrm{N}} \sum_{\mathrm{d}} \mathrm{a}_{\mathrm{mnd}} \mathrm{s}_{\mathrm{n}}[\mathrm{t}-\mathrm{d}] \tag{3}$$

where s_n [t] is the nth signal, $x_m[t]$ is the mth microphone input signal and d is the discrete-time delay and reflects the source n to microphone m impulse response. Although these factors may fluctuate over time in practice, they are commonly considered to be stationary to simplify the model. The noise can be described as (4) and (5).

$$\mathrm{Ni}[\mathrm{t}] = (\mathrm{ni1}[\mathrm{t}], \mathrm{ni2}[\mathrm{t}], \dots, \mathrm{ni\ M}[\mathrm{t}]) \tag{4}$$

$$\tilde{\boldsymbol{x}}_{\boldsymbol{m}}[\boldsymbol{t}] = \sum_{n-1}^{N} \sum_{d} a_{mnd} s_n[t-d] + \ noise\ m[t] \tag{5}$$

4. PROPOSED SPEECH ENHANCEMENT METHOD

This paper describes a speech enhancement technique that is based on source separation and adaptive LMS to improve passenger voice commands derived from a signal containing varied interferences and different speech sounds. To begin, the microphone captures the necessary signals, which are then passed to the source separation, which separates the denoised signals and removes the permutation ambiguity. The frequency domain is used in the separation process. As a result, the convolutive mixes are converted to the frequency domain using a short-time Fourier transform (STFT). Finally, we use inverse short-time Fourier transform (ISTFT) to convert the unmixed signals to the time domain. The adaptive LMS receives the source separation outputs. Figure 2 depicts the specific procedure.

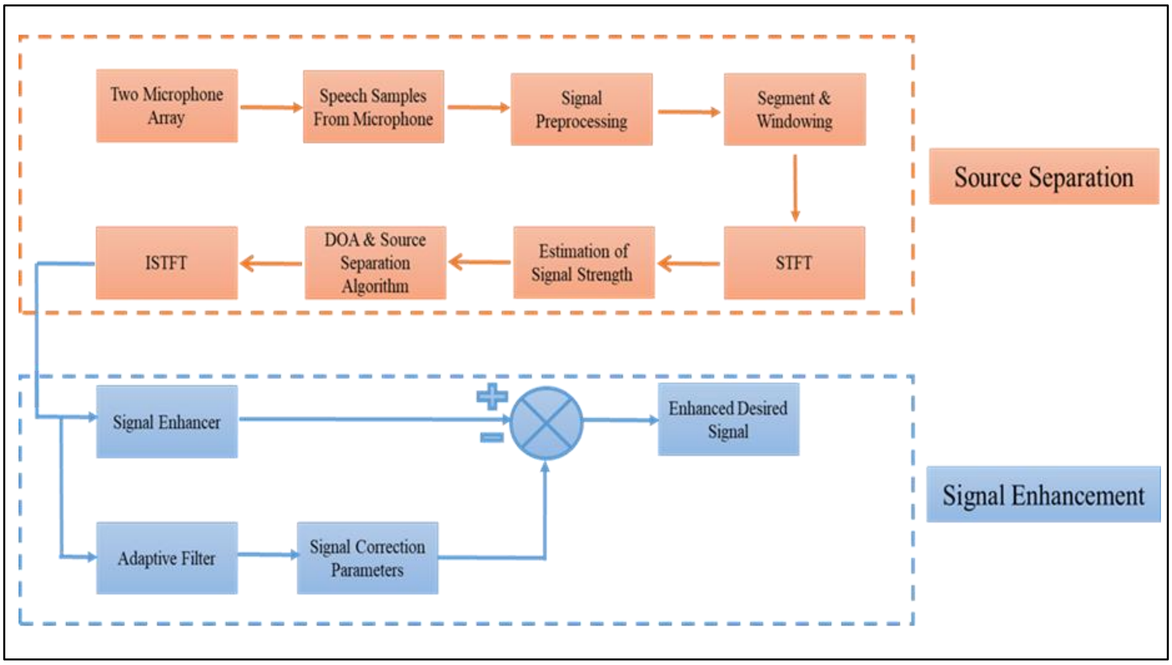


Figure 2. Block diagram of proposed speech enhancement technique

4.1. Input source mixture and preprocessing

An array microphone of a linear array shape with two microphones has been used in the analyzed method. The array system can be increased to eight microphones, The processing complexity and processing time will be more. Let us consider the input source matrix as (6),

$$\mathrm{X(n)} = \begin{bmatrix} \mathrm{x}_{1,1} & \mathrm{x}_{1,2} & \cdots & \mathrm{x}_{1,n} \\ \mathrm{x}_{2,1} & \mathrm{x}_{2,2} & \cdots & \mathrm{x}_{2,n} \\ \vdots & \vdots & \vdots & \vdots \\ \mathrm{x}_{N,1} & \mathrm{x}_{N,2} & \cdots & \mathrm{x}_{N,n} \end{bmatrix} \tag{6}$$

where X(n) is input source mixture, $\mathrm{x}_1, \mathrm{x}_2 \ldots \mathrm{x}_n$ is individual channel data, N is number of channels, n is number of samples per frame.

The source separation method is implemented for the real-time application so that input to the algorithm is processed frames by frames. Block-based processing is done, and the block size is 256 samples. a Hanning window is used in this process [18]. Wn is the windowing function.

$$X(n) = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,n} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,n} \\ \vdots & \vdots & \vdots & \vdots \\ x_{N,1} & x_{N,2} & \cdots & x_{N,n} \end{bmatrix} * Wn \tag{7}$$

Wn is the Hanning window coefficient. Enhancing hands-free speech inside the car environment, the input signal is filtered using second order IIR bandpass filter with a cutoff of 300 to 4,200 Hz. The filtered input signal matrix X(n) is given as (8),

$$X(n) = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,n} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,n} \\ \vdots & \vdots & \vdots & \vdots \\ x_{N,1} & x_{N,2} & \cdots & x_{N,n} \end{bmatrix} * h(n) \tag{8}$$

where h(n) is filter coefficients.

4.2. The direction of arrival (DOA)

Source direction is calculated to depend on the elapsed time of the signal arrived at the microphone array. The difference in time of arrival of signals received at microphones is physically separated by the cross-correlation function, which can be modeled as (9).

$$R_{ij}(\tau) = \sum_{n=0}^{N-1} x_i[n]\, x_j[n - \tau] \tag{9}$$

The two microphones are denoted by i and j. $x_i[n]$ and $x_j[n]$ are the signals received at microphones i and j respectively; n is the time-sample index, and τ is the signal correlation lag. The STFFT of the signal is given as (10),

$$R_{ij}(\tau) = \frac{1}{N}\sum_{K=0}^{N-1} X_i\,(k) X_j(k) * e^{\frac{j2\pi k\tau}{N}} \tag{10}$$

where $X_i(k)$, $X_j(k)$–FFT of $x_i(n), x_j(n)$. N is the number of FFT points. Time difference between two signals are represented by (11).

$$\tau_{delay} = \text{argmax}\left(R_{ij}(\tau)\right) \tag{11}$$

The time of arrival at the microphone is given by (12),

$$\tau_{delay} = \frac{\tau_{delay}}{C} \tag{12}$$

where C is the speed of sound in air. As a result, (ϴ) can be used to estimate the direction as (13),

$$\Theta = \text{Sin}^{-1}\left(\frac{\tau_{delay}}{d}\right) \tag{13}$$

where d is microphones separated by this distance.

Equation (13) gives the input signal direction. The signal flow of the SSCA algorithm is depicted in Figure 3. The linear array will act as the spatial filter to separate the input signal as desired and the interference signal. The linear array indicates the sound source direction from different seating positions in the car. Let us consider the direction angle as given in (14),

$$R_n = [r_1, r_2, r_3 \ldots\ldots r_{12}] \tag{14}$$

where r_1 is reference, r_2 represents d distance from r1. If the desired signal is from the left side user, then the then right will be masked.

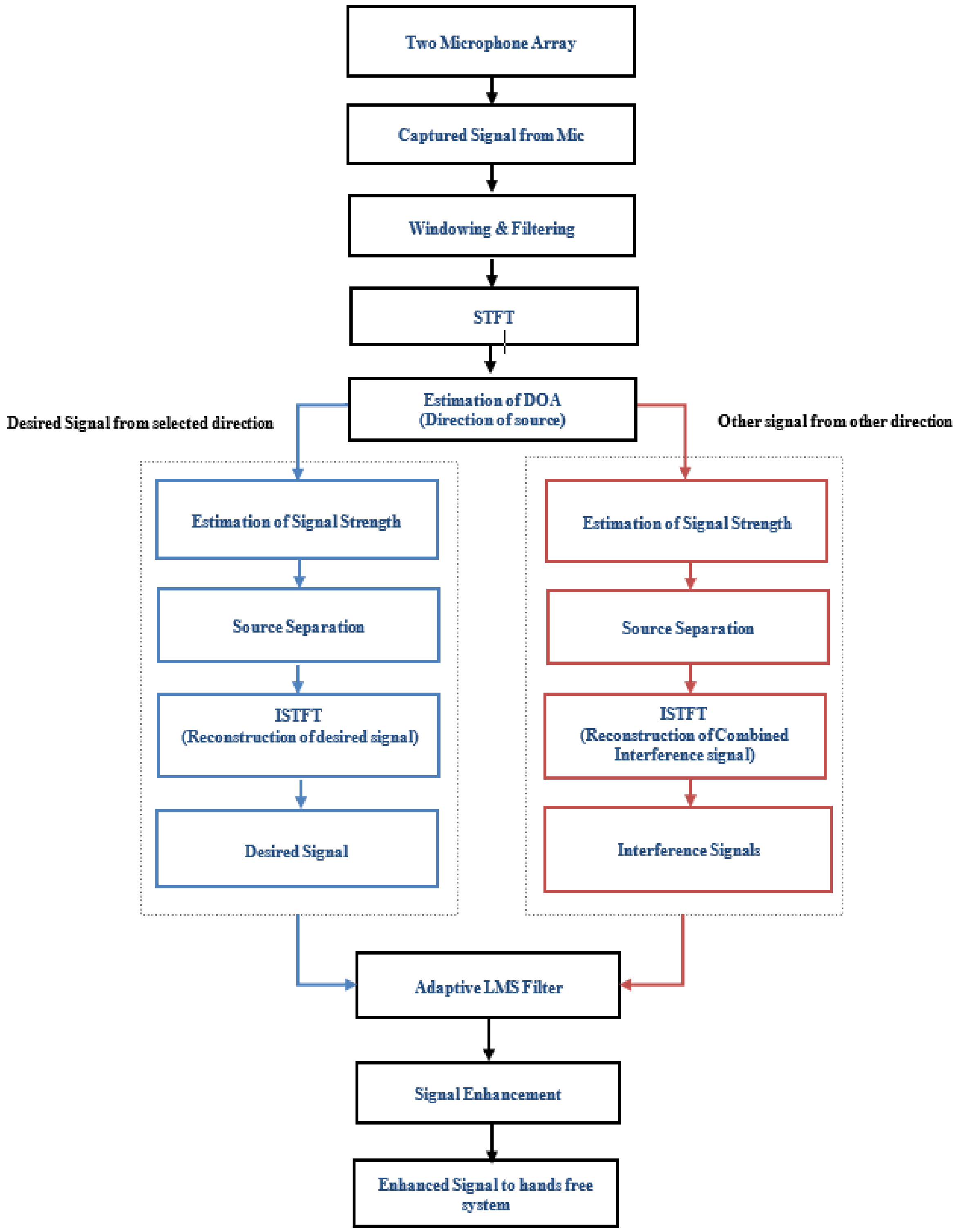


Figure 3. Flow chart of the proposed algorithm

4.3. Source separation (direction classification, masking, and reconstruction)

The source separations can be defined as in (15) and (16).

$$R_n(\text{desired}) = [r_1, r_2, r_3 \ldots\ldots r_{12}] \tag{15}$$

$$R_n(\text{interference}) = [r_1, r_2, r_3 \ldots\ldots r_{12}] \tag{16}$$

In a car, in the front, there will be a driver seat and a co-passenger seat, either one can select their position as the desired signal for the hands-free system. The direction of the signal from the desired direction will be processed, separated, and enhanced. The direction angle of the input is defined by (17),

$$Z_n = [i_1 , \quad i_2 \ \ldots . \quad i_n] \tag{17}$$

where n is number of samples per frame. The direction of arrival and the time delay are used to mask signal from other directions.

$$W_n(\text{desired}) = Z_n * R_n(\text{desired}) \tag{18}$$

$$W_n(\text{interference}) = Z_n * R_n(\text{interference}) \tag{19}$$

The interference and desired signal are obtained by convoluting the masking coefficients with the input signal. The separated interference and desired signal are derived by (20) and (21).

$$Y_n(\text{desired}) = X_i(k) * W_n(\text{desired}) \tag{20}$$

$$Y_n(\text{interference}) = X_i(k) * W_n(\text{interference}) \tag{21}$$

The inverse STFFT is used to reconstruct the signal in time domain. The processed interference and desired signals are given to the adaptive LMS filter for further enhancement.

4.4. Adaptive LMS filter

Adaptive filters are used to iterate the model using the input and expected output signal. The adaptive digital filter changes the coefficients based on iteration. The parameters are, u(n) is the signal from other directions, d(n) is the selected direction signal, y(n) is the output signal, and e(n) is the signal error of u(n) and y(n).

Finite impulse response (FIR) or infinite impulse response (IIR) can be used to create adaptive filters. The proposed method needs a linear phase, so an FIR-based adaptive filter is used. The error signal e(n) can minimize by the adaptive system using iterations.

The y(n) output signal from adaptive filter (22),

$$y(n) = w(n) \ * \ k(n)T \tag{22}$$

where k(n) is input signal, k(n)=$[u(n), u(n-1) \ldots \ldots \ldots u(n-N+1)]T$, w(n) is adaptive filter coefficients, $w(n) = [w_0(n), w_1(n) \ldots \ldots \ldots \ w_{N-1}(n)]T$. The error is derived using (23),

$$e(n) = d(n) - y(n) \tag{23}$$

The adaptive system coefficients are calculated using (24),

$$w(n+1) = \mu . e(n). k(n) + (1 - \mu c). w(n) \tag{24}$$

where μ is Filter step size w(n) is filter coefficients, and k(n) is input signal. The adaptive system gives an enhanced [19] output. The interference level is adaptively reduced in the enhanced output.

5. EXPERIMENTAL RESULT AND DISCUSSION

In this section, the experiment details and the simulation outcomes are given and analyzed. Figure 4 depicts a typical situation for the operation of speech enhancement in the car. There is a speech signal as well as interference signals in the car environment. Interference speeches are considered noisy sources, the desired source direction is selected by the user. The noise source signal in moving a car from the NOISEX-92 database [20]. The desired source is placed on the left or right to the device, and the speech is from another direction of the desired direction as shown in the car. A total of 50 sets of data were collected with different combinations of input sources in different directions. Microphone receivers include Mic1, Mic2, and so on.

Data captured by the two microphones serve as input to the algorithm. The spectrogram and time domain of the input signal desired, interference, and enhanced output signals are depicted in Figures 5 to 8.

Table 1 lists the parameters that were used in the SSCA algorithm. Figure 5 shows the speech and different kinds of noise-mixed signals in the time domain and time-frequency domain, respectively. Human speech and ambient noise have a wide frequency distribution. They can provide a great deal of information about speech characteristics. Figures 6(a) and 6(b) illustrates the separation result of the desired signal in the time domain and spectrogram [21]. The noise signal of a moving car in the time domain and spectrogram is shown in Figure 6. The frequency of car sound is primarily distributed in 50-500 Hz with no apparent regularity across time. Figure 6 shows that the time-domain processed signals are significantly different from the sources, indicating interference signals can be found in the voice spectrum. Furthermore, the speech spectrogram contains noise-desirable signal information at 270 Hz that affects the speech enhancement, and the ambient signal has a high frequency which indicates the SSCA approach is dependent on the adaptive LMS algorithm. Figure 7(a) shows the separated interference speech signal and Figure 7(b) is the spectrogram correspondingly. The proposed technique improves the speech signal, as shown in Figures 8(a) and 8(b) shows the spectrogram. As the result is shown in Table 2 the proposed method improves speech more effectively. According to the performance results, an SSCA method gives a good SNR and perceptual evaluation of speech quality (PESQ) at each input source combination, when compared to the other methods.

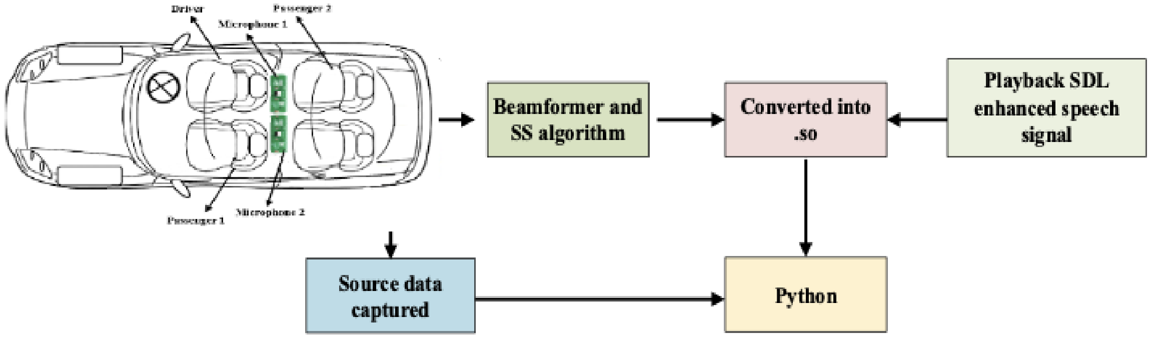
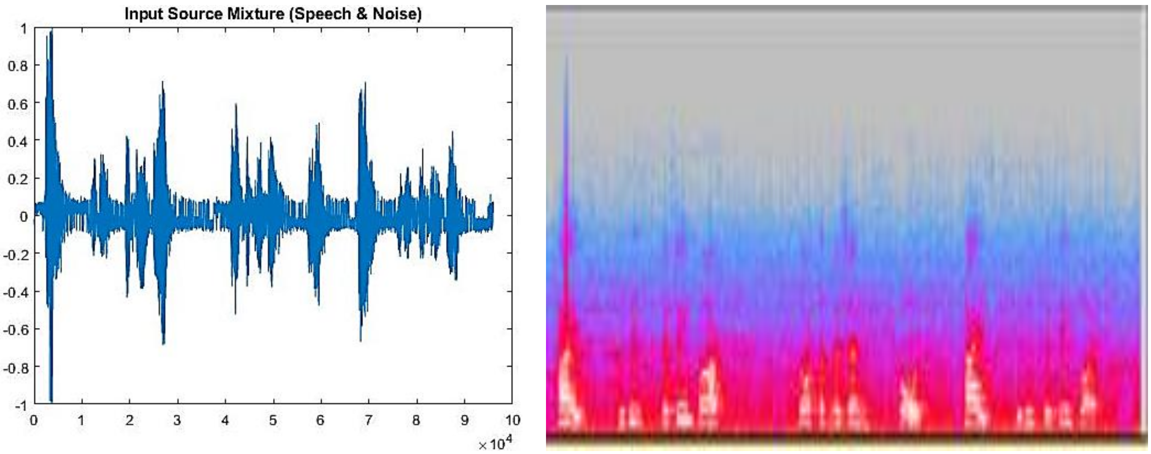


Figure 4. Operation of speech enhancement in car

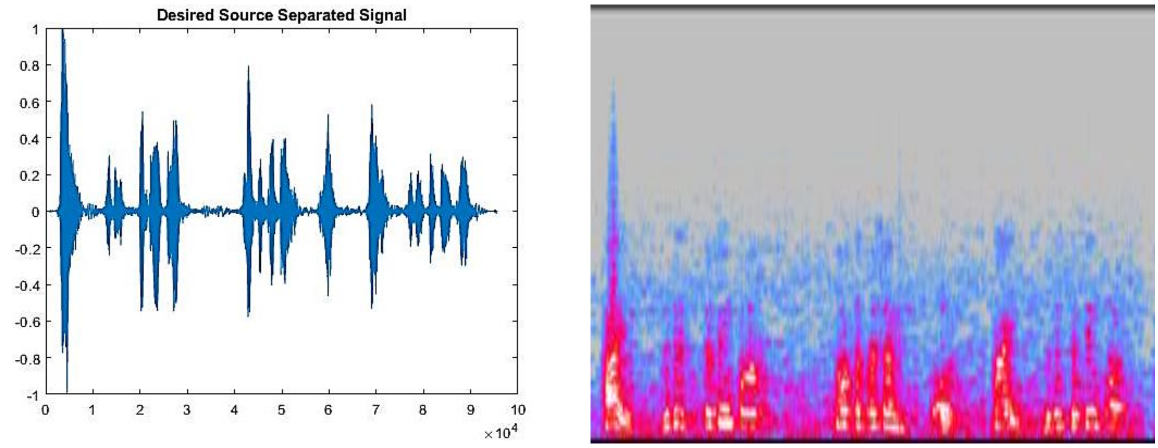
Table 1. Parameters used in source separation method

Characteristic	Parameter
Numbers of sources	2
Source classifications	source direction
Number of mics	2
Sampling rate	8 kHz
Size of the FFT window	512
Adaptive LMS filter step size	0.01
Filter length	32



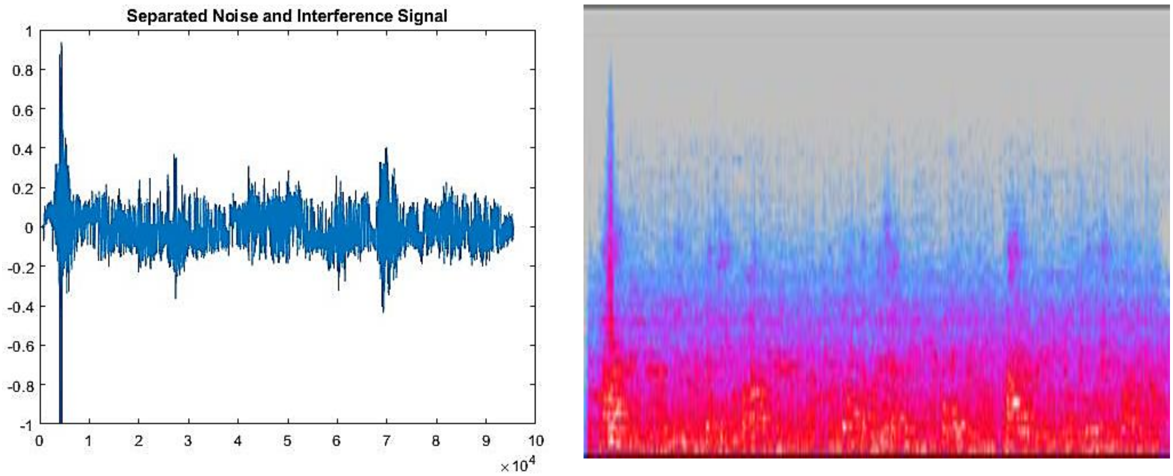
(a) (b)

Figure 5. The (a) time domain and (b) spectrogram of input signal



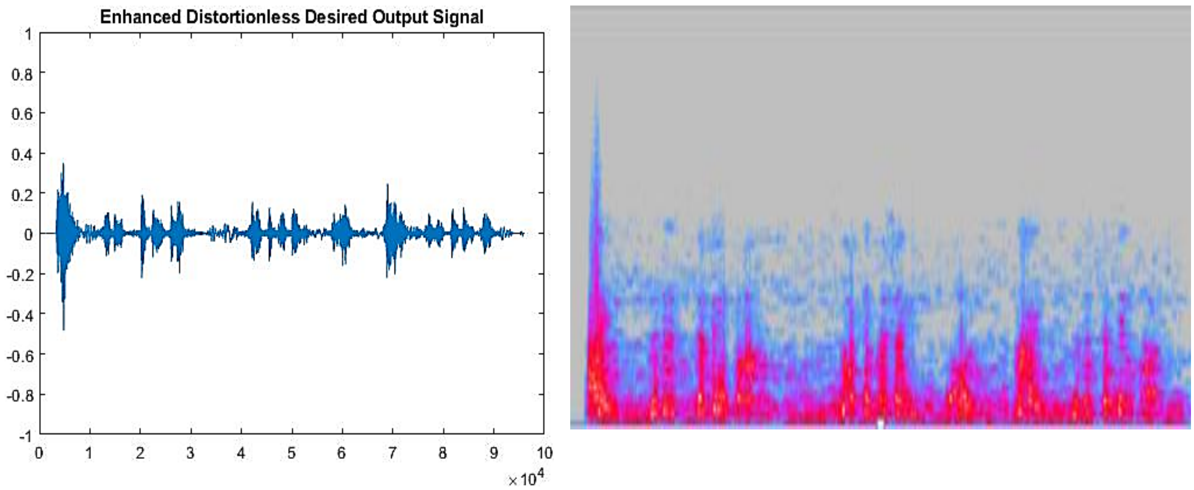
(a) (b)

Figure 6. The (a) time domain and (b) spectrogram of separated desired signal



(a) (b)

Figure 7. The (a) time domain and (b) spectrogram of separated interference signal



(a) (b)

Figure 8. The (a) time domain and (b) spectrogram of enhanced desired signal

Table 2. Comparison of results of SNR and PESQ of SSCA method with other methods at multiple combinations of input signals

Combination of sources	Speech with interference		Speech with interference		Speech with environmental noise and interference speech	
Desired source direction	90 degree		270 degree		90 degree	
Methods	PESQ	SNR (dB)	PESQ	SNR (dB)	PESQ	SNR (dB)
ICA	1.2	3	0.9	2.1	0.6	1.6
MVDR	1.7	4.8	1.5	3.5	1.3	2.8
RLS	2	7.2	1.9	6.2	1.7	4.4
Proposed SSCA-adaptive LMS algorithm	2.8	8.2	2.6	7.4	2.3	5.2

6. CONCLUSION

In the car environment, a range of interference effects and varied speech signals provide a major difficulty for the operation of the hands-free system. In this paper, a proposed adaptive LMS and SSCA algorithm is used to process and enhance the speech signal from the desired direction. The input source direction is obtained using the TDOA between the microphones. The input signals are separated and reconstructed to get the desired interference signals. The separated signals are adaptively filtered and enhanced. The SSCA method is implemented and validated using real-time hardware with different combinations of input source mixtures, which has yielded the expected results. Real-time results show that the proposed technique can successfully separate various speech signals without the need for prior information in the car system. The efficiency of the system is then demonstrated using a two-microphone beamformer. The signal quality parameters of the input signal to the enhanced desired signal are SNR of 5.2 dB and PESQ of 2.3 achieved.

REFERENCES

[1] J. Benesty, "Speech enhancement via correlation coefficients," in *Fundamentals of Speech Enhancement*, Springer International Publishing, 2018, pp. 45–64. doi: 10.1007/978-3-319-74524-4_4.

[2] B. K. Khonglah, A. Dey, and S. R. M. Prasanna, "Speech enhancement using source information for phoneme recognition of speech with background music," *Circuits, Systems, and Signal Processing*, vol. 38, no. 2, pp. 643–663, Feb. 2019, doi: 10.1007/s00034-018-0873-x.

[3] Z.-X. Li, L.-R. Dai, Y. Song, and I. McLoughlin, "A conditional generative model for speech enhancement," *Circuits, Systems, and Signal Processing*, vol. 37, no. 11, pp. 5005–5022, Nov. 2018, doi: 10.1007/s00034-018-0798-4.

[4] P. Malathi, G. R. Suresh, M. Moorthi, and N. R. Shanker, "Speech enhancement via smart larynx of variable frequency for laryngectomee patient for Tamil language syllables using RADWT algorithm," *Circuits, Systems, and Signal Processing*, vol. 38, no. 9, pp. 4202–4228, Sep. 2019, doi: 10.1007/s00034-019-01055-8.

[5] T. K. Dash, S. S. Solanki, and G. Panda, "Improved phase aware speech enhancement using bio-inspired and ANN techniques," *Analog Integrated Circuits and Signal Processing*, vol. 102, no. 3, pp. 465–477, Mar. 2020, doi: 10.1007/s10470-019-01566-z.

[6] J. Yang, B. Xia, Y. Shang, W. Huang, and C. Mi, "Improved battery parameter estimation method considering operating scenarios for HEV/EV applications," *Energies*, vol. 10, no. 1, Dec. 2016, doi: 10.3390/en10010005.

[7] S. Gannot, E. Vincent, S. Markovich-Golan, and A. Ozerov, "A consolidated perspective on multimicrophone speech enhancement and source separation," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 25, no. 4, pp. 692–730, Apr. 2017, doi: 10.1109/TASLP.2016.2647702.

[8] M. Tammen and S. Doclo, "Deep multi-frame MVDR filtering for single-microphone speech enhancement," in *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Jun. 2021, pp. 8443–8447. doi: 10.1109/ICASSP39728.2021.9413775.

[9] E. Gentet, B. David, S. Denjean, G. Richard, and V. Roussarie, "Speech intelligibility enhancement by equalization for in-car applications," in *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2020, pp. 6934–6938. doi: 10.1109/ICASSP40776.2020.9053537.

[10] Y. Alkaher and I. Cohen, "Dual-microphone speech reinforcement system with howling-control for in-car speech communication," *Frontiers in Signal Processing*, vol. 2, Mar. 2022, doi: 10.3389/frsip.2022.819113.

[11] P. Lei, M. Chen, and J. Wang, "Speech enhancement for in-vehicle voice control systems using wavelet analysis and blind source separation," *IET Intelligent Transport Systems*, vol. 13, no. 4, pp. 693–702, Apr. 2019, doi: 10.1049/iet-its.2018.5094.

[12] M. Wang, E. Zhang, and Z. Tang, "Speech enhancement based on NMF under electric vehicle noise condition," *IEEE Access*, vol. 6, pp. 9147–9159, 2018, doi: 10.1109/ACCESS.2018.2797165.

[13] T. Tao *et al.*, "Sound localization and speech enhancement algorithm based on dual-microphone," *Sensors*, vol. 22, no. 3, Jan. 2022, doi: 10.3390/s22030715.

[14] X. Li, M. Fan, L. Liu, and W. Li, "Distributed-microphones based in-vehicle speech enhancement via sparse and low-rank spectrogram decomposition," *Speech Communication*, vol. 98, pp. 51–62, Apr. 2018, doi: 10.1016/j.specom.2017.12.008.

[15] A. Panda, "Denoising algorithms using stacked RNN models for in-car speech recognition system," in *2018 4th International Conference for Convergence in Technology (I2CT)*, Oct. 2018, pp. 1–5. doi: 10.1109/I2CT42659.2018.9057894.

[16] L. Qian, F. Zheng, X. Guo, Y. Zuo, and W. Zhou, "Vehicle speech enhancement algorithm based on TanhDBN," in *2020 IEEE 3rd International Conference of Safe Production and Informatization (IICSPI)*, Nov. 2020, pp. 434–438. doi: 10.1109/IICSPI51290.2020.9332320.

[17] S. Krause, O. Otto, and F. Stolzenburg, "Fast classification learning with neural networks and conceptors for speech recognition and car driving maneuvers," in *Lecture Notes in Computer Science*, Springer International Publishing, 2021, pp. 45–57. doi: 10.1007/978-3-030-80253-0_5.

[18] N. Siegel, J. Rosen, and G. Brooker, "Faithful reconstruction of digital holograms captured by FINCH using a Hamming window

function in the Fresnel propagation," *Optics Letters*, vol. 38, no. 19, Oct. 2013, doi: 10.1364/OL.38.003922.
[19] J. Pathrose, M. M. Ismail, and V. Govindaraj, "MASTER: microphone array source time difference eco canceller via reconstructed spiking neural network," *IEEE Sensors Journal*, vol. 23, no. 11, pp. 12058–12065, Jun. 2023, doi: 10.1109/JSEN.2023.3266259.
[20] K. Rohith and R. Bhandarkar, "A comparative analysis of statistical model and spectral subtractive speech enhancement algorithms," in *Advances in VLSI, Signal Processing, Power Electronics, IoT, Communication and Embedded Systems*, Springer Singapore, 2021, pp. 397–416. doi: 10.1007/978-981-16-0443-0_32.
[21] M. Barhoush, A. Hallawa, A. Peine, L. Martin, and A. Schmeink, "Localization-driven speech enhancement in noisy multi-speaker hospital environments using deep learning and meta learning," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 31, pp. 670–683, 2023, doi: 10.1109/TASLP.2022.3231700.

BIOGRAPHIES OF AUTHORS

Jeyasingh Pathrose received his B.Sc. degree in physics from Madras University, Chennai, India, in 1986, his B.E. degree in electronics and communication engineering (ECE) from the Karunya Institute of Technology, Coimbatore, India in 1990, the M.E. degree in applied electronics from the Government College of Technology, Coimbatore, in 1994, and the EMBA degree from IIT Madras, Chennai, in 2019. He is pursuing his Ph.D. degree in signal processing-beamforming from B.S. Abdur Rahman Crescent Institute of Science and Technology (Deemed University), Chennai. He had worked as an Assistant Professor with the Department of Electronics and Communication Engineering, Mohamed Sathak Engineering College, Keelakarai. He had worked at the Indchem Research and Development Laboratories, Chennai, in the field of audio signal processing and at Bharat Dynamics Ltd., Hyderabad (Ministry of Defense, Government of India) as an assistant manager at Image Processing with the Research and Development Department. He is the Chief Technology Officer of Jasmin Infotech Pvt Ltd., Chennai. He has a multifaceted experience in audio signal processing, image processing, pattern recognition, and wavelets. He has authored or co-authored papers in different international conferences and journals. He can be contacted at gpjs@hotmail.com.

Mohamed Ismail received a bachelor's degree in engineering from Madurai Kamaraj University, Tamil Nadu, India, in 1989, a master's degree in electronics and control from the Birla Institute of Technology and Science, Hyderabad, India, in 1997, and a Ph.D. degree in very-large-scale integration (VLSI) from Jawaharlal Nehru Technological University, Hyderabad, in 2014. He is working as a professor and the Dean (Academic Affairs) of the B.S. Abdur Rahman Crescent Institute of Science and Technology, Chennai, India. He has multifaceted experience in the field of VLSI modeling, micro-processor and microcontroller designs, and digital and embedded systems applications. He has authored or co-authored papers in different international conferences and journals. He can be contacted at mmismail@crescent.education.