Speech enhancement by using novel multiband spectral subtraction method along with a reduction of the cross spectral component

Jagadish S. Jakati[1], Ramesh B. Koti[2], Sidramayya Matad[3], Jagannath Jadhav[4], Shrishail Basvant Mule[5], Sanmati Bedakihale[4], Vireshkumar G. Mathad[6], Amar R. Bandekar[6]
[1]School of Computer Science Engineering and Application, D. Y. Patil International University Akurdi, Pune, India
[2]Department of Electronics and Communication Engineering, Gogte Institute of Technology, Belagavi, India
[3]Department of Electronics and Communication Engineering, S G Balekundri Institute of Technology Shivbasavnagar, Belagavi, India
[4]Department of Electronics and Communication Engineering, KLECET, Chikodi, India
[5]Department of Electronics and Telecommunication Engineering, Sinhgad College of Engineering, Pune, India
[6]Department of Electrical and Computer Engineering, Dr. A D Shinde Collage of Engineering Badagaon Kolapur, Gadhinglaj, India

Article Info

Article history:

Received Dec 22, 2023
Revised Jul 31, 2024
Accepted Aug 5, 2024

Keywords:

Cross spectral components
FFT
Multi-band spectral subtraction
SNR value
Spectral subtraction

ABSTRACT

It is essential to enhance the speech signal's clarity and quality in order to maintain the message's content. By boosting the noisy voice signal, the speech signal quality can be raised. Two techniques are presented in this study to significantly minimize the additive background noise. In order to minimize non-stationary additive noise concerning the speech signal, the first approach employs modified multiband spectral subtraction. With this technique, spectral subtraction is carried out based on the signal to noise ratio (SNR) values in various noisy speech frames. When the noisy signal and noise signal are somewhat correlated, a second method is used to minimize the cross spectral components. These techniques are used to get over the drawbacks of the fundamental spectrum subtraction method. To improve the noisy speech signal, both techniques are combined.

Corresponding Author:

Jagadish S. Jakati
School of Computer Science Engineering and Application, D. Y. Patil International University Akurdi
Pune, Maharashtra 411044, India
Email: jagadishjs30@gmail.com

1. INTRODUCTION

The evolution of speech has made it the main form of communication. Every communication system has background noise, which not only interferes with listening tasks but also lowers the efficiency of digital signal processors. As a result, lowering background noise is essential for effective communication.

Speech enhancement is one of the techniques for reducing noise from noisy speech signals in order to increase the excellence, intelligibility, and accessibility of the speech signal. Applications for speech augmentation include teleconferencing, hearing aids, recording systems, and mobile and remote communication. The algorithms in this study can be used to reduce additive noise, this comprises white noise, flicker noise, babbling noise, and so forth. Both techniques employ the fundamental spectral subtraction technique. By discusses various methods to lessen additive noise in [1]. Among this technique is the fundamental spectral subtraction approach, which includes subtracting the predicted noise's power spectrum magnitude from the noisy speech signal in order to acquire the speech signal. In the process, two presumptions are made. The first is the assumption that speech signals are stationary for brief periods of time.

The second assumption is that there is no relationship between the noisy speech signal and the clear speech signal. When there is no speech signal, the noise signal is approximated from those moments of silence. These presumptions are not always valid in practise; There is no distribution of noise evenly throughout the noisy speech signal. Consequently, this technique for improving speech introduces musical noise.

2. LITERATURE SURVEY

Zheng *et al.* [1] has explained the method to enhance the noise speech signal, which is corrupted by the acoustic noise. The approach adopted is the spectral subtraction method. The spectrum of magnitude of the clear speech signal is estimated by subtracting the estimated noise speech signal from the noisy speech signal. By analyzing the signal during non-speech signal activity, an estimate of the noise's magnitude spectrum is made. Other techniques are employed to eliminate any remaining noise caused by the spectral subtraction method. The methods are half wave, magnitude averaging and residual noise reduction. Limitations: assumption made that the background noise environment remains stationary which is not true, and it generates musical noise.

Zheng *et al.* [2] outlines the methods for getting over the drawbacks of the conventional spectrum subtraction approach. The assumption expressed in the spectral subtraction that the noise is stationary is not true in real time applications, hence it introduces musical noise. To avoid this here author has modified the basic spectral subtraction into two factors those are, subtraction factor or weighting factor which is used to eradicate most of the broad band noise by removing wide peaks. Another factor used is the spectral factor which helps in eliminating the musical noise perceived during the process. Merits: the proposed method removes added broad band noise as well as the musical noise signal generated by the basic spectral subtraction approach. Spectral subtraction factor mainly depends on the signal to noise ratio (SNR) value; hence the method is adaptive in nature. Limitations: the efficiency of the recommended method is dependent on the decision of the subtraction factor and the floor factor.

Toyin *et al.* [3] explains the method used to enhance the noisy speech signal, short duration non-stationary signal, corrupted by the broad band noise proposed approach is “adaptive wiener filter” where spectral estimate of the previous frames are used to find the spectral estimate of the current frame and spectral estimate of the current frame is used to update the coefficients of the wiener filter on previous frame. This method is followed by spectral smoothing, which is performed based on the spectral change. Slight temporal smoothing is applied for the fast spectral changes and slight temporal, when the spectral change is fast, amount of the smoothing is increased when spectral change is very slow. Merits: the adaptive nature of the proposed method has been increased by using spectral changes to update coefficients of the wiener filter. Limitations: this method introduces musical noise as well as the enhance signal is slight dull.

Zheng *et al.* [4] aims at reducing the musical noise generated by spectral subtraction method. Proposed method involves “wiener filtering” and “wavelet packet decomposition”. Enhanced signal is divided based on wavelet packet decomposition method. Power spectral density of the wavelet packet coefficients is filtered by wiener filtering method. It is observed that musical noise is placed in a more detailed manner as coefficients. So, the spectrum subtraction method's production of musical noise is decreases. Merits: the quality of musical noise is drastically reduced without affecting the intelligibility of the speech. Limitations: quality of clean speech is dependant the on number of decomposition stages, for improve the quality we have to increase number of decomposition stages.

Das *et al.* [5] presented the necessary correction to the fundamental spectral subtraction approach, the predicated on the concept that the clear speech signal and the noise signal don't correlate. Therefore, noise corrupted speech signal that has been tainted by noise that is linked with the speech signal can be enhanced using the proposed approach. We estimate the clean speech magnitude spectrum by subtracting an estimate of the cross-correlation between the clean speech signal and the noise signal after first eliminating the estimated noise spectrum from noisy speech. It is necessary to find the cross-correlation between noise and clean speech but clean speech is not accessed hence a correlation is computed between the noisy speech signal and noise. This method is followed by the “perceptual weighting function” to reduce noise as well as to improve the speech quality. Where weighting function is computed based on the psychoacoustics model. Merits: this technique gets over the drawbacks of the spectrum subtraction method., where it can be used for real time applications ex., speech corrupted by the coloured noise. Limitations: the computation of the correlation between noise and speech signal is a lengthy and complex method, which introduces distortion. Hence this method only focuses on the noise reduction.

Roy and Paliwal [6] addresses problem of the basic spectral subtraction approach, the following presupposes that the noise signal and the precise speech signal are both inherently stationary. But this method fails when there is non-stationarity present in the any of the both signals. To overcome this problem the author modelled the non-stationary signal into the sum of sinusoids known as “tones”, and stated that window length is selected based on the tone duration. In proposed work multiple spectral subtraction stages are used with different window length. Sum of enhanced signal from each stage are summed to get the output.

Boll [7] explained the technique for minimizing the auditory noise produced by the spectral subtraction method. Proposed method helps to enhance the clean speech signal corrupted by the stationary noise. This method estimates the spectrum of the noise better than the basic averaging method. In this method fist silence frame and speech frame are separated by using basic analysis frame. Further analysis frame length is increased till it covers all the silent frames. Limitations: proposed method can be used only when noise added is constant hence cannot be used for coloured noise.

In any typical speech enhancement techniques mentioned in the above papers describe the speech signal's small-time magnitude spectrum, keeping phase spectrum unchanged. But Chen *et al.* [8] has combined both changed magnitude and phase spectrum after the enhancement to form the modified complex signal spectrum. This method is applicable to the conditions where speech noise energy is less compared to the speech energy. Limitations: this method fails when noise is non-white noise as, babble noise, coloured noise, as it introduces distortion as well as residual noise.

Bharti *et al.* [9] has stated problem of the one having sensory-neural hearing impairment, facing problem in speech perception due to the increased intra-speech spectral masking. Proposed method involves two methods, first method is spectral subtraction method to suppress the external noise. Second method is multi-band frequency compression to reduce intra-speech masking. The speech signal's spectrum in the multi-band frequency compression technique is segmented into bands and spectral components are compressed at the centre of the frame and concentrate the speech energy into narrow bands to reduce masking by adjacent spectral components.

3. OUTCOME OF LITERATURE SURVEY

We are going to summarize the existing speech enhancement methods. This also gives the limitations or problems faced during the processing as well as drawbacks of those methods. From the literature survey we can observe that the method basically used to suppress the noise present in the corrupted speech is spectral subtraction method, where by subtracting the magnitude spectrum of the predicted noise from that of the noisy speech signal, we are going to be able to estimate the clean speech signal. The main drawbacks of the method were the assumptions and the method work only when the speech signal is stationary in nature and noise is uncorrelated to the speech signal. Some other methods of speech enhancement discussed in the literature survey are wiener filtering method, wavelet packet decomposition method, Frequency compression method, multi-band spectral subtraction method. Every method involves limitations some of them are mentioned below,

- Problems faced by all of the aforementioned techniques were the generation of musical noise, presence of residual noise after processing and distortion caused due to the processing.
- Wavelet packet decomposition method reduces musical noise, but the effectiveness of this approach is dependent on the number of decomposition stages, hence better quality can be obtained only with the increased complexity.
- Signal subspace approach can be used to avoid the non-stationarity problem as well as it also reduces the residual noise but causes some distortion in the processed speech.

Hence some of the methods were successful in reducing the musical noise but were failed when it come for residual or distortion. Proposed work is designed to overcome these problems.

By enhancing the speech signal's perception and comprehensibility, the speech enhancement approach is used to increase the speech signal's intelligence. It is also used in many real time applications as in mobile communication, and hearing aids. Among the existing methods of speech enhancement is the basic spectral subtraction method, that enhances the noise corrupted speech signal which, is considered to be non-stationary and where, noise is uncorrelated to the speech signal. These assumptions not really exist in most of the real time applications. Therefore, an effective speech enhancement approach is required in order to get around the drawbacks of the spectral subtraction technique. Hence proposed the work “Speech enhancement using multi-band spectral subtraction using cross spectral subtraction”. Objectives of proposed work are i) to minimize the noise present in the corrupted speech signal; ii) to eliminate the musical noise generated by the spectral subtraction method; and iii) to eliminate the possible distortion that can cause during processing.

4. PROPOSED METHOD

The proposed work introduces two speech enhancement algorithms in order to perform the fundamental spectral subtraction approach and prevent the musical noise that can be produced as a result of this method. The first technique uses a modified form of multi-band spectral subtraction. This approach is used to handle noisy speech signals that have suffered from additive noise and are non-stationary [10]. The SNR of the current frame serves as the foundation for doing spectral subtraction. Section 1 of the document contains all of the method's specifics. In order to handle the noisy speech signal, that has been tainted by noise linked to the speech signal, the second technique comprises computing the relationship between the noise-free and clean signals of speech. In section 2, the computation's specifics are explained [11].

4.1. Modified multi-band spectral signal

Let:
w(n)-Enhancing the noisy speech signal.
a(n)-Additive noise.
s(n)- Noise free speech signal.
Thus, W(n) is provided by (1).

$$w(n) = s(n) + a(n) \tag{1}$$

Using the frequency domain conversion, let s(n) be expressed as (2).

$$W(f) = S(f) + A(f) \tag{2}$$

The noise-corrupted speech signal's power spectrum is shown as (3).

$$|W(f)|^2 = |S(f)|^2 + |A(f)|^2 + S(f)A(f)* + S(f)*A(f) \tag{3}$$

The fundamental spectral subtraction method requires that the noise signal along with the noise-corrupted signal are uncorrelated. $S(f).A(f)* + S(f)*.A(f)$ terms in (3) are neglected. Therefore, (4) may be used to produce the clean speech signal S(f).

$$|S(f)|^2 = |W(f)|^2 - |A(f)|^2 \tag{4}$$

Even so, this approach makes the unfeasible assumption that noise is equally distributed across the tainted speech signal, which is impractical. So, if we use the same technique, the noisy speech signal will be reduced by the same amount of anticipated noise [12]. A different approach of speech enhancement is necessary to prevent this, in which the amount of noise to be removed relies on the SNR in the relevant section of the signal W(n). The over subtraction factor is calculated using modified multi-band spectral subtraction, which is dependent on the SNR value. In order to calculate clean speech, over subtraction can be introduced into (4). The factor can then be obtained by (5).

$$|S(f)|^2 = |W(f)|^2 - \alpha|A(f)|^2 \tag{5}$$

In the connection between and SNR in this paper explained [13]. As for the relationship:

$$\alpha = \begin{cases} 5 & SNR < 5 \\ 4 - 0.15\,(SNR) & -5 < SNR < 20 \\ 1 & SNR > 20 \end{cases} \tag{6}$$

4.2. Cross-correlation technique

In (3), $S(f).A(f)* + S(f)*.\ A(f)$ is regarded as the cross-correlation term that is ignored in the spectral subtraction method. Nonetheless, there is some association between speech signal and noise in real-time applications [14]. Thus, it is essential to identify these correlation words, x_{dc} and $x_{cd.}$ Since we are unale to acess to precise speech, we can determine a correlation between the noise signal and the corrupted speech signal. For example, x_{yd} where x_{sd} is yields,

$$x_{sd} = x_{cd} + x_{dd} \tag{7}$$

$$|W(f)^2| = \begin{cases} |y(f)|^2 - \alpha|N(f)^2| - \delta|y(f)| * |N(f)| & if\,|y(f)|^2 > \alpha|N(f)^2| \\ \beta|N(f)^2| & else \end{cases} \tag{8}$$

where is the calculation of the over spectral subtraction factor, using (6) and it is the reported value of 0.002 for the spectral floor factor from the publication [15]. Is a correlation factor that estimates how closely the noisy speech signal and estimated noisy speech signal are correlated. Equation to determine is:

$$\delta = \left|\frac{\chi_{yn} - \mu_y \mu_n}{\sigma_y \sigma_n}\right| \tag{9}$$

where,

$$\chi_{yn} = \frac{1}{N/2} \sum_k |y(f)| * |N(f)|$$

$$\mu_y = \frac{1}{N/2} \sum_k |y(f)|$$

$$\mu_n = \frac{1}{N/2} \sum_k |N(f)|$$

where, μ_y, μ_n are the means values of the noise corrupted speech signal and noise signal. Where $0<n<N/2$, N is the lenght of fast fourier transform. σ_y^2, σ_n^2 are the variances of the noise corrupted speech signal and enhanced noisy signal.

A noise corrupted speech signal is initially separated into 20-ms frames (160 samples/frame). For this, Hamming Window is employed (with window size 160). Techniques for windowing may cause spectral leakage near the edges of the window, which may result in information loss; hence, 50% overlapping is performed before signal processing to prevent the same [16]. Windowed noisy voice signal is denoted by (10).

$$Ww(n) = W(n) * w(n) \tag{10}$$

From (1),

$$\begin{aligned} Ww(n) &= [s(n) + a(n)] * w(n) \\ &= sw(n) + aw(n) \end{aligned} \tag{11}$$

After computing, fast fourier transfrom (FFT) of the noise corrupted speech sample, the magnitude spectrum is calculated using (3) and (2). The noisy speech signal's frame magnitude spectrum is separated into frames with 40 samples each in modified multiband spectral subtraction [17]. Based on the SNR values of these frames, spectral subtraction is carried out independently for each frame using (5) and (6) by calculating the over subtraction factor.

Finally, the correlation factor is determined using (8) and (9) and (7). It is possible to retrieve the precise speech signal magnitude. Complex spectrum is created by incorporating the predicted non noisy speech signal magnitude spectrum with the original speech signal unaltered phase spectrum [18]. To translate a complex spectrum into a time domain signal, inverse frequency fourier transform is used. In the same way that 50% overlapping is used for framing [19]. The clear speech signal is achieved via 50% overlap addition. The Figure 1 illustrates the various processes that go into putting the suggested strategy into practice [20].

5. RESULTS AND DISCUSSION

In this proposed method, speech quality is evaluated using a subjective listening test and spectrogram analysis. These analysis techniques are used to compare the performance of the stated approach to those of the current speech enhancement method. In a subjective listening test, listeners are used to compare processed speech to an unprocessed speech signal. The speech quality can be rated by listeners using a predetermined scale [21]. A spectrogram is a representation of a speech signal's time and frequency, with the frequency changing as the time changes. The spectrogram's colour corresponds to the speech's energy at that frequency. Dark colour indicates a high-energy speaking signal [22].

The spectrogram analysis of noisy speech signals perverted by babble noise at 0 dB and 15 dB is shown in Figures 2 and 3, respectively [23]. Also, the spectrogram analysis of the improved speech sounds is shown in Figures 4 and 5. It has been noticed that utilizing the suggested strategy has improved speech quality. The average rating for the modified multiband spectral subtraction subjective listening test for signals with 0 dB SNR and 15 dB SNR is 2.7 and 2.6 (low), respectively [24]. The suggested approach has an average rating of 3.7 and 3.6 for signals with 0 dB SNR and 15 dB SNR, respectively (greater than that of previous method) [25], [26].

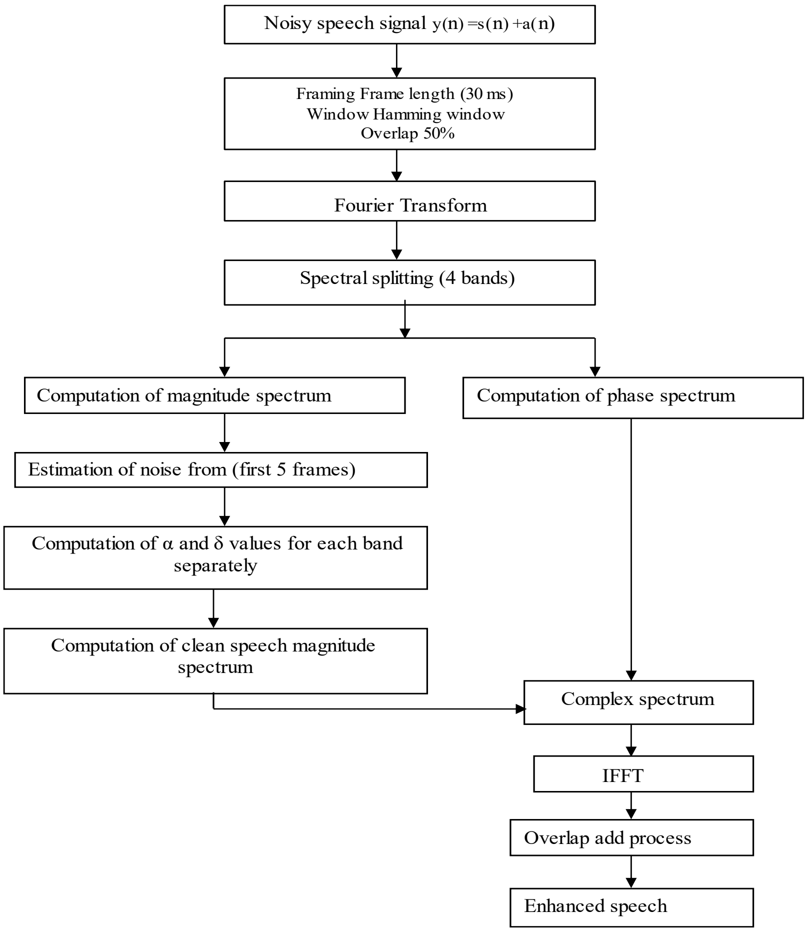

Figure 1. Flow chart for the proposed method

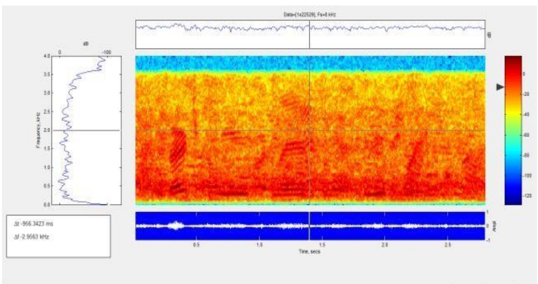

Figure 2. Signal 0 dB SNR babble noise speech signal

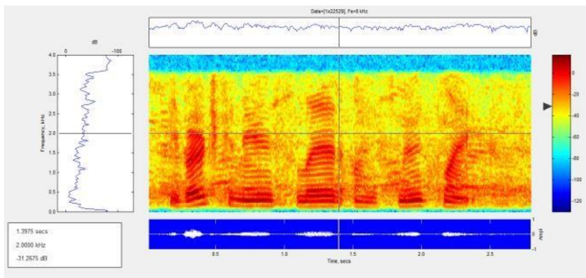

Figure 3. Signal 15 dB SNR babble noise speech signal

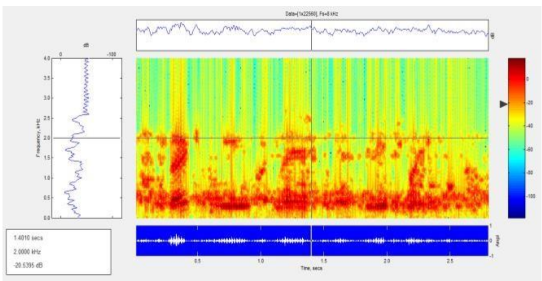

Figure 4. Signal enhanced by the proposed method

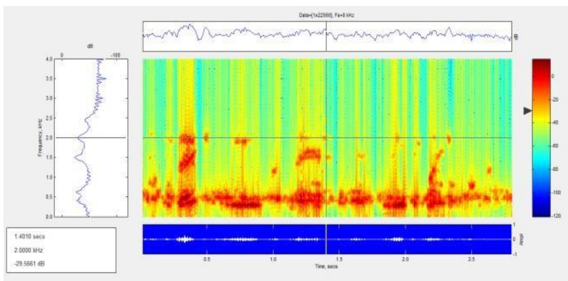

Figure 5. Signal enhanced by multiband spectral subtraction

6. CONCLUSION

In this study, the issues and restrictions of the base spectral subtraction method are discussed. By calculating the value of the over subtraction factor, we performed multiband spectral subtraction in this research. Cross-correlation approach was used to construct additional cross spectral components. According to the results and discussions, the suggested method performs the spectrum subtraction method in terms of the excellence of speech signal.

REFERENCES

[1] C. Zheng, Y. Ke, X. Luo, and X. Li, "Convolutional neural network-based models for speech denoising and dereverberation: algorithms and applications," in *IoT-enabled Convolutional Neural Networks: Techniques and Applications, New York: River Publishers*, 2023, pp. 65–95.

[2] C. Zheng, W. Liu, A. Li, Y. Ke, and X. Li, "Low-latency monaural speech enhancement with deep filter-bank equalizer," *The Journal of the Acoustical Society of America*, vol. 151, no. 5, pp. 3291–3304, May 2022, doi: 10.1121/10.0011396.

[3] O. E. Toyin, A. D. Oluwole, and A. Z. Kayode, "Speech enhancement in wireless communication system using hybrid spectral-Kalman filter," *International Journal of Electrical and Electronic Engineering and Telecommunications*, vol. 11, no. 5, pp. 363–372, 2022, doi: 10.18178/ijeetc.11.5.363-372.

[4] C. Zheng, X. Peng, Y. Zhang, S. Srinivasan, and Y. Lu, "Interactive speech and noise modeling for speech enhancement," in *Proceedings of the AAAI Conference on Artificial Intelligence*, May 2021, vol. 35, no. 16, doi: 10.1609/aaai.v35i16.17710.

[5] N. Das, S. Chakraborty, J. Chaki, N. Padhy, and N. Dey, "Fundamentals, present and future perspectives of speech enhancement," *International Journal of Speech Technology*, vol. 24, no. 4, pp. 883–901, Dec. 2021, doi: 10.1007/s10772-020-09674-2.

[6] S. K. Roy and K. K. Paliwal, "Robustness and sensitivity tuning of the Kalman filter for speech enhancement," *Signals*, vol. 2, no. 3, pp. 434–455, 2021, doi: 10.3390/signals2030027.

[7] S. Boll, "Suppression of acoustic noise in speech using spectral subtraction," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 27, no. 2, pp. 113–120, Apr. 1979, doi: 10.1109/TASSP.1979.1163209.

[8] Z. Chen, Y. Liu, G. Wang, S. Wang, and W. Geng, "Multiband spectral subtraction speech enhancement algorithm with phase spectrum compensation," in *2019 IEEE 4th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*, Dec. 2019, pp. 2681–2685, doi: 10.1109/IAEAC47372.2019.8997837.

[9] S. S. Bharti, M. Gupta, and S. Agarwal, "A new spectral subtraction method for speech enhancement using adaptive noise estimation," in *2016 3rd International Conference on Recent Advances in Information Technology (RAIT)*, Mar. 2016, pp. 128–132, doi: 10.1109/RAIT.2016.7507888.

[10] P. Sunitha and K. S. Prasad, "Multi band spectral subtraction for speech enhancement with different frequency spacing methods and their effect on objective quality measures," *International Journal of Image, Graphics and Signal Processing*, vol. 11, no. 5, pp. 54–62, May 2019, doi: 10.5815/ijigsp.2019.05.06.

[11] Q. Huang, C. Bao, X. Wang, and Y. Xiang, "DNN-based speech enhancement using MBE model," in *2018 16th International Workshop on Acoustic Signal Enhancement (IWAENC)*, Sep. 2018, pp. 196–200, doi: 10.1109/IWAENC.2018.8521278.

[12] J. S. Jakati and S. S. Kuntoji, "A noise reduction method based on modified LMS algorithm of real time speech signals," *WSEAS Transactions on Systems and Control*, vol. 16, pp. 162–170, Mar. 2021, doi: 10.37394/23203.2021.16.13.

[13] A. Hussain, K. Chellappan, and M. S. Zamratol, "Single channel speech enhancement using ideal binary mask technique based on computational auditory scene analysis," *Journal of Theoretical and Applied Information Technology*, vol. 91, no. 1, pp. 12–22, 2016.

[14] N. Tiwari, S. K. Waddi, and P. C. Pandey, "Speech enhancement and multi-band frequency compression for suppression of noise and intraspeech spectral masking in hearing aids," in *2013 Annual IEEE India Conference (INDICON)*, Dec. 2013, pp. 1–6, doi: 10.1109/INDCON.2013.6726008.

[15] J. S.Jakati, "Efficient speech de-noising algorithm using multi-level discrete wavelet transform and thresholding," *International Journal of Emerging Trends in Engineering Research*, vol. 8, no. 6, pp. 2472–2480, 2020, doi: 10.30534/ijeter/2020/43862020.

[16] M. Krawczyk and T. Gerkmann, "STFT phase reconstruction in voiced speech for an improved single-channel speech enhancement," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 22, no. 12, 2014, doi: 10.1109/TASLP.2014.2354236.

[17] Y. Wang, A. Narayanan, and Y. Wang, "On training targets for supervised speech separation," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 22, no. 12, pp. 1849–1858, Dec. 2014, doi: 10.1109/TASLP.2014.2352935.

[18] C. Li and W.-J. Liu, "A novel multi-band spectral subtraction method based on phase modification and magnitude compensation," in *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2011, pp. 4760–4763, doi: 10.1109/ICASSP.2011.5947419.

[19] R. Udrea and D. Vizireanu, "An improved multi-band speech enhancement method for colored noise estimation and reduction," *International Journal on Advance in Telecomunications*, vol. 3, no. 3 & 4, pp. 271–280, 2010.

[20] C. W. Li and S. F. Lei, "Signal subspace approach for speech enhancement in nonstationary noises," in *ISCIT 2007 - 2007 International Symposium on Communications and Information Technologies Proceedings*, May 2007, pp. 1580–1585, doi: 10.1109/ISCIT.2007.4392269.

[21] M. Okazaki, T. Kunimoto, and T. Kobayashi, "Multi-stage spectral subtraction for enhancement of audio signals," in *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, 2004, vol. 2, pp. ii-805–8, doi: 10.1109/icassp.2004.1326380.

[22] S. Ayat, M. T. Manzuri, R. Dianat, and J. Kabudian, "An improved spectral subtraction speech enhancement system by using an adaptive spectral estimator," in *Canadian Conference on Electrical and Computer Engineering*, 2005, vol. 2005, pp. 261–264, doi: 10.1109/CCECE.2005.1556923.

[23] K. Wójcicki, M. Milacic, A. Stark, J. Lyons, and K. Paliwal, "Exploiting conjugate symmetry of the short-time fourier spectrum for speech enhancement," *IEEE Signal Processing Letters*, vol. 15, pp. 461–464, 2008, doi: 10.1109/LSP.2008.923579.

[24] N. Saleem, M. Irfan, X. Chen, and M. Ali, "Deep neural network based supervised speech enhancement in speech-babble noise," in *2018 IEEE/ACIS 17th International Conference on Computer and Information Science (ICIS)*, Jun. 2018, pp. 871–874, doi: 10.1109/ICIS.2018.8466542.

[25] Y. Wang and D. L. Wang, "Towards scaling up classification-based speech separation," *IEEE Transactions on Audio, Speech and Language Processing*, vol. 21, no. 7, pp. 1381–1390, Jul. 2013, doi: 10.1109/TASL.2013.2250961.

[26] J. S. Jakati and S. S. Kuntoji, "A novel speech enhancement solution using hybrid wavelet transformation least means square method," *International Journal of Engineering Trends and Technology*, vol. 69, no. 7, pp. 233–243, 2021, doi: 10.14445/22315381/IJETT-V69I7P230.

BIOGRAPHIES OF AUTHORS

Dr. Jagadish S. Jakati received B.E. in Electronics and Communication Engineering in 2010, M. Tech in 2012 and Ph.D. in 2022 from VTU, Belagavi. Presently, he is working as Assistant Professor at School of computer Science Engineering and Application, D. Y. Patil International University Akurdi, Pune, Maharashtra 411044, India. He holds a Ph.D. degree in Electrical Engineering with specialization in speech processing. His research areas are image/signal processing, biometrics, medical image analysis, and pattern recognition. He has filed a number of patents and industrial designs on his innovative ideas and has been published with four Indian patents and more than 20 international journals. His research interests include image/signal processing, biometrics, medical image and analysis, and pattern recognition. He can be contacted at email: jagadishjs30@gmail.com.

Dr. Ramesh B. Koti received B.E. in Electronics and Communication Engineering in 2008, M. Tech in 2013 and Ph.D. in 2023 from VTU, Belagavi. Presently, he is working as Assistant Professor in Electronics and Communication Engineering Department at Gogte Institute of Technology Belagavi, 590008, Karnataka, India. He holds a Ph.D. degree in Electrical Engineering with Specialization in Vehicular Networks. His research areas are wireless sensor networks, computer networks, MANET, VANET, and IoT. He has filed a number of patents and industrial designs on his innovative ideas and has been published with one Indian patent and more than 11 international journals. He can be contacted at email: rameshkoti1984@gmail.com.

Dr. Sidramayya Matad is Associate Professor and Dean R&D, at ECE Department, S G Balekundri Institute of Technology, Belagavi, India. He holds a Ph.D. degree in Electrical Engineering with Specialization in Communication. His research areas are communication, image/signal processing. He has published more than 10 international journals. His research interests include Communication image/signal processing, biometrics, medical image and analysis, and pattern recognition. He can be contacted at email: siddu.matad@gmail.com.

Dr. Jagannath Jadhav received B.E. in Electronics and Communication Engineering, M. Tech in Digital Electronics and Communication and Ph.D. in 2020 from SSSUTMS sehore, post-doctoral fellowship from Lincoln University College, Jalan Kota Bharu-Pengkalan Kubor, 15050 Kota Bharu, Kelantan, Malaysia in 2021. Presently he is working as Professor at KLECET Chikodi-591201, affiliated to VTU Belagavi, Karnataka, India. He is the renowned researcher having 4 patents published in Indian patent journal. He has guiding many Ph.D. students and has authored more than 10 books and published more than 12 international journals. His research interests include image/signal processing, pattern recognition, and remote sensing applications. He can be contacted at email: jagannathjadhav3030@gmail.com.

Dr. Shrishail Basvant Mule Associate Professor, Department of Electronics and Telecommunication Engineering, Sinhgad College of Engineering, Pune-41. He has 18 publications in reputed journals and one Indian patents. His research interests include image/signal processing, biometrics, medical image and analysis. He can be contacted at email: mulesb1@gmail.com.

Mrs. Sanmati Bedakihale is Assistant Professor at ECE Department, S G Balekundri Institute of Technology, Belagavi. She holds M. Tech degree in Electronics and Communication with Specialization in Industrial Electronics. Her research interests include communication system, electronics and signal processing. She can be contacted at email: sanmatidupadal1001@gmail.com.

Dr. Vireshkumar G. Mathad received B.E. in Electrical and Electronics Engineering in 2010, and Ph.D. in 2022 from VTU, Belagavi, India. Presently he is working as Associate Professor and HOD in Electrical and Computer Engineering, Dr. A D Shinde Collage of Engineering Badagaon Kolapur. He has 18 publications in reputed journals and one Indian patents. His research interests include FACTS, image/signal processing, and power system enhancement. He can be contacted at email: vireshmatad@gmail.com.

Mr. Amar R. Bandekar received B.E. in Electrical and Electronics Engineering in 2014, and M.Tech. in Electrical Power Systems in 2019 from V.T.U. Belagavi, India. Presently he is working as Assistant Professor in Electrical and Computer Engineering, SGMCOE, Mahagaon. His research interests include FACTS. He can be contacted at email: amarbandekar3@gmail.com.