The Effect of Speech Enhancement Techniques on the Quality of Noisy Speech Signals

Ahmed H. Y. Al-Noori[(⊠)] 🆔, Atheel N. AlKhayyat 🆔,
and Ahmed A. Al-Hammad 🆔

Al-Nahrain University, Baghdad Al Jadriya 10072, Iraq
ahmed.alnoori@gmail.com
atheelnowfal964@gmail.com
ahmed.a.al-hammad@nahrainuniv.edu.iq

Abstract. Recently, the human voice has been used widely in different modern applications. However, environmental noise represents one of the significant challenges for these types of speech applications since it affects the quality of the speech signal caused by decreasing the Signal to Noise Ratio (SNR). Speech enhancement techniques represent one of the remedies for this challenge. This paper presents a study of speech cleaning techniques to improve the quality of noisy speech signals in different environmental noise and SNR based on human perception. The study is based on the human auditory system for 50 volunteers. Three types of speech enhancement algorithms are utilised in this study. Despite the results, these enhancement techniques' effects are invariant based on the kind of noise and the value of the Signal to Noise Ratio. Still, the subspace approach performs better than the other two approaches, especially with the high SNR.

Keywords: Speech enhancement · Wiener filter · Spectral subtraction · Subspace approaches · Environmental noise

1 Introduction

One of the most natural types of human-to-human and human-to-machine communication is the human speech signal. It has been utilised in various applications recently, including Automatic Speech Recognition (ASR), Speaker Recognition (voice biometric), speech coding systems, mobile communication, and intelligent virtual assistant. Due to numerous ambient noises, the performance of these speech application systems are severely decarded; hence, the reception task becomes difficult for a direct listener and causes inaccurate transfer of information. Noise suppression and, in turn, enhancement of speech is the main motive of many researchers in the fields of speech signal processing over the decades [1, 2]. Speech enhancement algorithms are designed to increase one or more perceptual characteristics of noisy speech, most notably quality and intelligibility.

In specific applications, the principal objective of speech enhancement algorithms is to increase speech quality while retaining, at the very least, speech intelligibility [2, 3]. Hence, the focus of most speech enhancement algorithms is to improve the quality of speech. Speech enhancement methods seek to make the corrupted noisy speech

signal more pleasant to the listener. Furthermore, they are beneficial in other applications such as automatic speech recognition [4, 5].

Improving quality, however, does not always imply that intelligibility would increase. The major cause for this is the distortion imparted on the cleaned speech signal as a result of severe acoustic noise suppression. Speech enhancement algorithms create two forms of distortion: those that impact the speech signal itself (called speech distortion) and those that influence the background noise (called background noise distortion). Listeners appear to be the more impacted by speech distortion when making overall quality judgments of the two types of distortion. Unfortunately, no objective metrics currently exist that correlate high with either distortion or the overall quality of speech enhanced by noise suppression algorithms [6]. Hence, the fundamental challenge in developing practical speech enhancement approaches is to suppress noise without avoiding distortion of speech signal.

Several techniques have been proposed. These techniques can be categorised into three main approaches [7, 8]: Firstly, the spectral subtraction approaches [3, 9–11], which depends on anticipating and updating the spectrum of the noise when there is silence pause in the speech signal, then subtracting the outcome from a noisy speech signal.

Secondly, Statistical model-based techniques, in these techniqes the cleaning speech problems are represented in a statistical prediction framework. These approaches are use a set of measurements, such as the Fourier transform coefficients of the noisy speech signal, to obtain a linear (or nonlinear) estimator of the parameter of interest, referred to as the transform coefficients of the speech signal [7]. Examples of these types of techniques, the Wiener filter [12–14], Minimum mean square error (MMSE) algorithms [15–17], and the maximum-likelihood approach for predicting the spectrum of the clean signal [18–20] and a slew of additional techniques falls under this set. Finally, linear algebra-based algorithms known as Subspace Algorithms: these types are based on linear algebra. The basic notion underlying these algorithms is that the clean signal might be contained within a subspace of the noisy Euclidean Space. Hence, dividing the vector space of a noisy speech signal into a clean signal subspace, which is mostly filled by clean speech, and a noise subspace, which is primarily occupied by noise (Loizou 2013). These algorithms were developed firstly by (Dendrinos et al. 1991)and (Ephraim and Van Trees 1995). This paper aims to study the impact of speech enhancement techniques on improving the quality of speech signal contaminated with different environmental noise based on the auditory system of 50 volunteers. Various Signal to Noise ratios SNRs is used in this investigation. The speech samples (which obtained from SALU-AC speech database) are corrupted with various types of environmental noise (Cafeteria Babble, Construction, and Street Noise).

2   Environmental Noise

Understanding the properties of background noises and the distinctions between noise sources in terms of temporal and spectral characteristics is critical for designing algorithms for suppressing additive noise. The noise signal is known as any unwanted sound signal that you do not need or want to hear. The long-term average spectrum of the five categories of environmental noise is demonstrated in Fig. 1. (Inside the car

noise, Cafeteria speech Babble, inside the train trailer, street noise, and white noise signal).

The first kind of monitoring is concerned with a lack of regularity in the spectrum, which gives a unique identity for each type of noise [8]. Noise can be generally classified as stationary noises (also known as Wide Sense Stationary WSS), such as the fan noise coming from PCs, which not change over time.
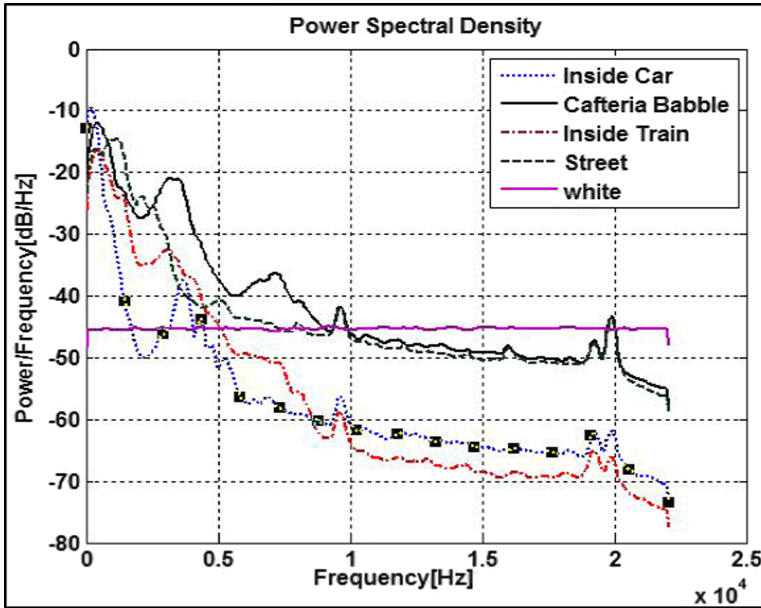

Fig. 1. Power spectral density of different types of noise [8].

Non-stationary noise has spectral characteristics continuously changing over time, such as in cafeteria Babble noise (Fig. 1), making suppression or removing of this type of noise more complicated than suppressing stationary noise. In Cafeteria babble noise, for example, maybe one of the most difficult forms of noise to handle in voice applications since several people chat in the background, which is sometimes mixed with noise from the kitchen. The spectral (and temporal) features of cafeteria noise are continuously changing as customers speak at neighbouring tables and servers engage and converse with them. The Signal to Noise Ratio (also known as speech to noise ratio) (SNR) is defined as the power level of disparity between speech and additive noise. SNR is typically expressed in decibels (dB), such that SNR = 0dB if the speech signal ratio is equal to the additive noise ratio.

In addition, the noise can also be classified into Continuous Noise (engine noise), Intermittent Noise (aircraft flying above your house), Impulsive Noise (explosions or shotgun), and Low-Frequency Noise (air movement machinery including wind turbines). In this paper, three types of non-stationary noise have been used **Cafeteria Speech Babble Noise, Construction Noise, and street noise.**

3   Speech Enhancement Techniques

Speech enhancement techniques, as previously discussed, is concerned with enhancing the perception of the speech signal that has been distorted by ambient noise. In most applications, these techniques' key aim is to increase the quality and intelligibility of the speech signal that is contaminated with environmental noise. In general, The enhancement in quality is more desirable since the technique can decrease listener fatigue, specifically in situations where the listener is exposed to high noise levels for a long time [2]. Since these techniques are applied to reduce or suppress background noise, speech enhancement is also known as the noise suppression algorithms (or speech cleaning) [2]. Various methods for cleaning speech signals and decreasing additive noise levels to increase speech efficiency have been improved in the literature. As stated previously, these strategies can be divided into three categories:

3.1   Spectral Subtraction Approaches

These approaches depend on the consideration that a noisy signal is a combination of both noise and clean speech signals. Consequently, the noise spectrum is calculated during speech pauses. Then the noise spectrum is subtracted from the original signal (noisy signal) to get clean speech [21]. These approaches were first suggested by Weiss et al. [22] and [23]. Consider a noisy signal $y(n)$ which consists of the clean speech $s(n)$ degraded by statistically independent additive noise $d(n)$ as follows:

$$y(n) = s(n) + d(n) \tag{1}$$

It is assumed that additive noise is zero mean and uncorrelated with clean speech. Because the speech signal is non-stationery and time-variant, The speech signal is supposed to be uncorrelated with the background noise. The representation in the Fourier transform domain is given by [24]:

$$Y(\omega) = S(\omega) + D(\omega) \tag{2}$$

The speech can be estimated by subtracting a noise estimate from the received signal.

$$\hat{S}(\omega) = |Y(\omega)| - |\hat{D}(\omega)|e^{j\theta_y(\omega)} \tag{3}$$

Where $|Y(\omega)|$ is the magnitude spectrum, $\theta_y(\omega)$ is the phase (spectrum) of the contaminated noisy signal, $\widehat{S}(\omega)$ the estimated clean speech signal.

The estimated speech waveform is recovered in the time domain by inverse Fourier transform $S(\omega)$ using an overlap and add approach. The drawback of this technique is the residual noise.

$$s(n) = IDTFT\{(|Y(\omega)| - |D(\omega)|e^{j\theta(\omega)}\} \tag{4}$$

where $s(n)$ is recovered speech signal.

3.2   Approaches Based on Statistical-Models

These approaches modelled the cleaning speech problem by using a statistical esti-mating framework. This is based on a set of observations, such as the noisy speech signal's Fourier transform coefficients, to obtain a linear (or nonlinear) estimate of the parameter of interest, known as the transform coefficients of the speech signal [2]. The Wiener filter [25], the maximum likelihood estimator [12], and minimum mean square error (MMSE) algorithms [15] are only a few examples of these sorts of approaches. This paper adopted the Wiener filter as a statistical approach since it represents the most commonly used approach in speech enhancement.

The Wiener filter is one of the most popular noise reduction techniques, and it has been described in a variety of ways and used in various applications. It is based on decreasing the Mean Square Error (MSE) between the estimated signal magnitude spectrum $\hat{S}(\omega)$ and real signal S(ω). The following is the formula for the best wiener filter [12, 26]:

$$H(\omega) = \frac{S_{s(\omega)}}{S_{s(\omega)} + S_{n(\omega)}}$$

(5)

where $S_s(\omega)$ and $S_n(\omega)$ represent the estimated power spectra of the noise-free speech signal and the additive noise, which are assumed uncorrelated and stationary. After measuring the transfer function of the Wiener filter, the speech signal is recovered through [12]:

$$\hat{S}(\omega) = X(\omega) \cdot H(\omega)$$

(6)

In a modified form of the Wiener filter, an adjustable parameter α has been used [12].

$$H(\omega) = \left(\frac{S_{s(\omega)}}{S_{s(\omega)} + \beta S_{n(\omega)}} \right)^{\alpha}$$

(7)

where β is noise suppression factor.

3.3   Subspace Approaches

These approaches are primarily linear algebra based. The core principle of these methods is that the clean signal may be contained within a subspace of the noisy Euclidean Space. As a result of dividing the vector space of a noisy speech signal into a clean signal subspace, which is mostly filled by the clean speech, and a noise subspace, which is primarily occupied by the noise signal [7, 8]. Then, nullifying the noisy vector variable in the noise subspace to produces the cleaning voice signal. These approaches were suggested by [27, 28]. The signal subspace is plagued by unwanted residual noise. The unwanted noise is supposed to be uncorrelated with the speech signal so that the noisy signal covariance matrix can be written as follows:

$$R_x = R_s + R_w \tag{8}$$

Where R_x is the signal covariance matrix, R_s is the clean speech covariance matrix and R_w is the noise vector with covariance matrix. With these assumptions, the following linear subspace filter is developed to estimate the desired speech vector from the noisy observation:

$$\hat{S} = Hx = Hs + Hw \tag{9}$$

Where Hs and Hw is the filter output and the desired speech after applying filter respectively, the residual error is defined as follows:

$$R = (H - I)s + Hw \tag{10}$$

where r is the residual error. The aim here is to decrease the signal distortion subject to keeping every spectral component of the residual noise in the signal subspace as little as possible.

4   Experimental Setup

Based on the perception of the human auditory system, this paper investigates the impact of speech enhancement approaches for improving the quality of noisy speech signals under varied ambient noise and varying SNR. As previously stated, three types of enhancement are adopted in this paper. The speech signals are adopted in this study were corrupted by three kinds of noise (Cafeteria babble noise, Construction noise, and Street noise) at SNRs (15 dB, 10 dB, and 0 dB). The processed speech signal was presented to regular hearing listeners to evaluate its quality. The results investigated the effect of these filters are invariant based on the type of noise and the value of the signal to noise ratio. Figure 2 show the block diagram of the methodology of this study:
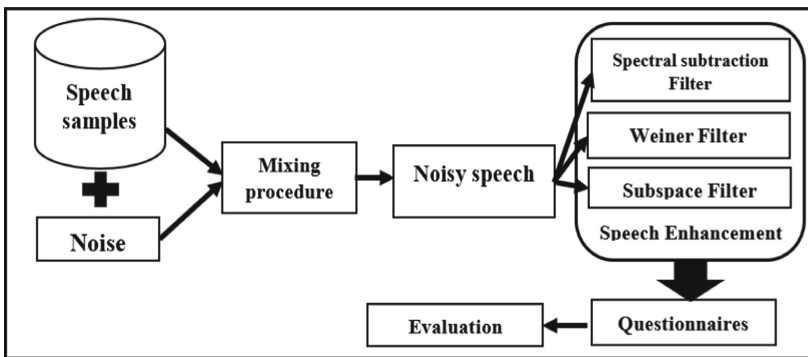

Fig. 2. The block diagram of the suggested study.

The experimental setup of this study can be summarised as follows:

1. Providing speech samples contaminated with different environmental noise and controlled SNR using the mixing procedure. This procedure is described in the next section.
2. Applying Speech Enhancement algorithms on noisy speech sample to produce the filtered speech samples.
3. Evaluating the performance of each speech enchantment algorithms based on the perception of human ears to filtered speech signals. Noisy speech samples and mixture procedure.

4.1 Speech Samples

The experiments were conducted on two speech samples' sets. These speech samples are collected from the University of Salford Anechoic Chamber Database (SALU-AC database) (Fig. 3). One of the database's most distinguishing characteristics is that it includes English speech sample spoken by native and non-native speakers and the recording environment that collected on it since data was gathered in the Anechoic Chamber. The principal purpose of this database was to offer clean speech samples, which make them more efficient while dealing with one adverse condition (such as noise) in isolation from other adverse conditions [29].
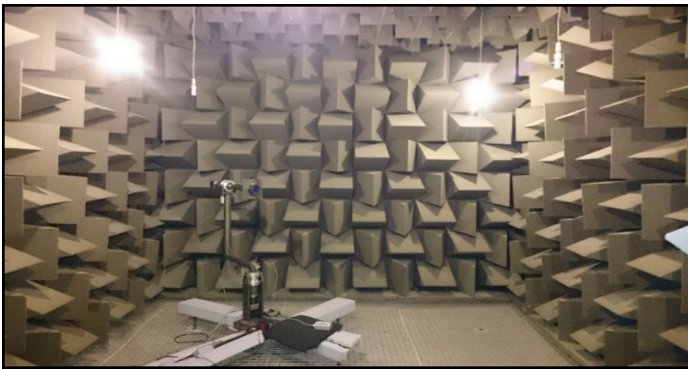

Fig. 3. Anechoic Chamber at University of Salford [8].

4.2 Noise Samples

As previously indicated, the speech datasets were generally collected in quiet environments (Anechoic Chamber) with no ambient noise influencing the organised speech signals. Consequently, the noisy speech samples were created by combining speech samples with the aforementioned sources of noise, each with a distinct regulated signal to noise ratio (SNR) (15, 10, and 0 dB).

The following is a summary of the mixing technique[8]:

1. To match the duration of target speech utterances, the noise signal was shortened. The main goal of this phase was to ensure that noise was evenly distributed across the speech signal.
2. Controlling the ratio at which the speech signal and noise were combined by specifying the SNR (in dB). As previously stated, 15dB, 10dB, and 0dB were chosen as mixing ratios because SNR 15 dB is near to clean (i.e., the ratio of speech is high relative to the noise, which is too low) and SNR 0 dB is hardly recognised by the human ear.
3. Normalising the speech and noise signals (this normalisation was done by using the root mean square RMS).
4. Finally, before mixed with the speech signal, the noise signal was scaled to achieve the appropriate SNR. Figure 4 shows a male voice signal that has been mixed with noise at various SNRs.
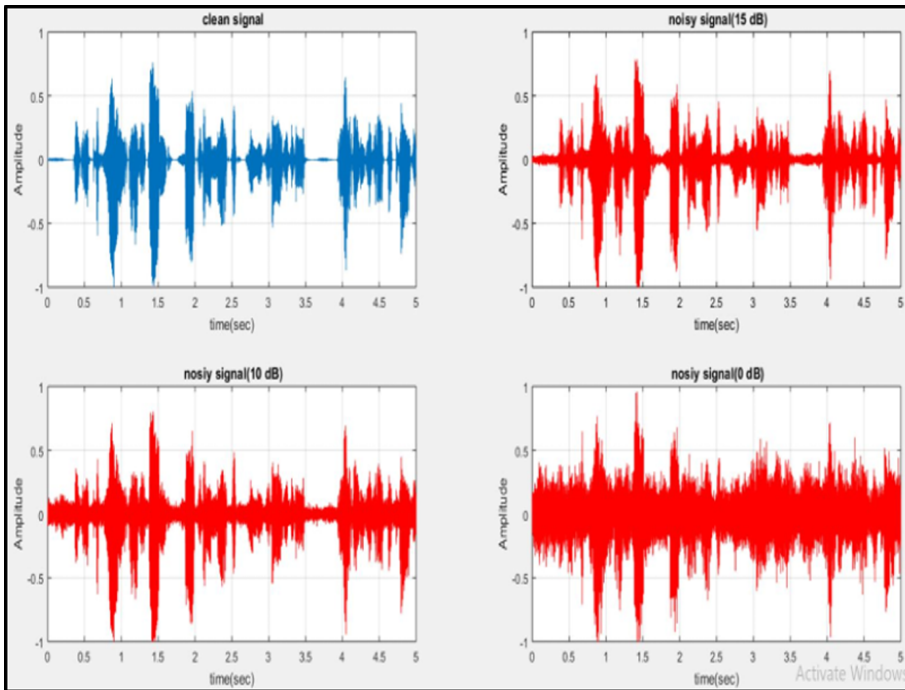

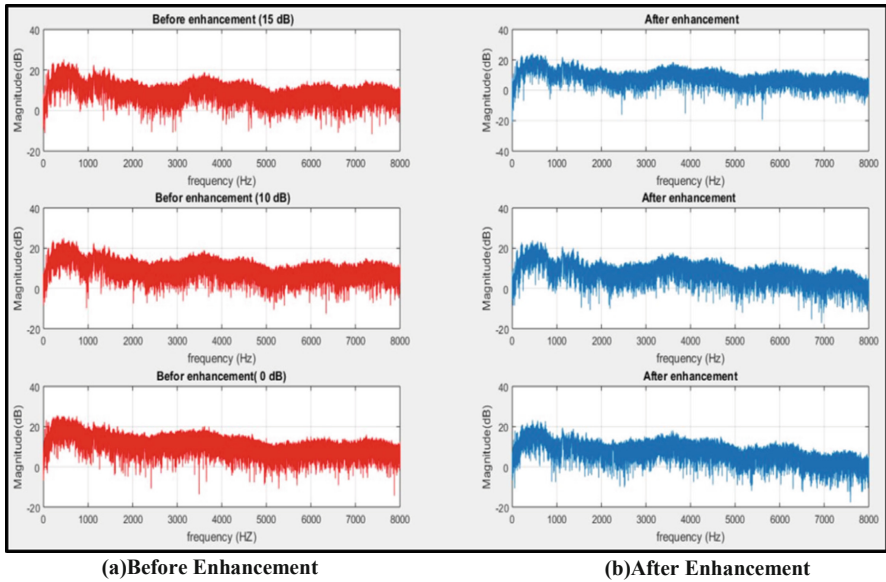**Fig. 4.** Speech Sample contaminated by environmental noise with different SNRs.

Fig. 5. The wiener filter for speech contaminated with babble noise (a) before enhancement (b) after enhancement with different SNRs.

4.3 Applied Speech Enhancement Algorithms

As mentioned earlier, three speech enhancement algorithms are adopted in this work: Spectral subtraction, Wiener filter, and Sub-space algorithms. Each of these filters is applied to two speech signals. These signals consist of one male speech signal and one female speech signal in order to study the effect of these algorithms of both gender signals. Each speech signal is mixing with different SNRs (15, 10, and 0 dB) for a specific type of noise that discussed earlier. Therefore, in total, we have 18 filtered speech samples, six filtered samples for each enchantment algorithms. Figure 5(a), (b) shows the spectrum of enhanced signals for male signal contaminated with cafeteria babble filtered by Wiener filter algorithm before and after the filtering process.

5 Questionnaires and Evaluations

The last level in this work is to evaluate each enhancement algorithm's performance for each speech signal contaminated with environmental noise mentioned before with each SNR based on the perception of the human auditory system. Fifty volunteers have been chosen (25 males and 25 females) for this purpose. The volunteer's ages are between 20–40 years old. Each one has been checked that has not any hearing issues. The evaluation is conducted in the Multimedia Laboratory at the College of Engineering in Al-Nahrain University, as seen in Fig. 6.

First, each listener has been instructed to listen to the clean signal without any additive noise (original signal). Then, he/she listen to the three filtered speech signals

filtered by the three speech enhancement algorithms (Spectral subtraction, Winer filter, and Subspace filter) for each SNR mentioned before. Finally, the volunteer chose the suitable filtered signal that think it clearly close to the original signal. The experiment, then, repeat for each type of noise (Cafeteria Babble, Street, and construction noise) and for each male and female signal.

Fig. 6. Multimedia lab for listening the enhanced speech signal.

5.1   Experimental Results

As mentioned before, this work evaluates the impact of different type of speech enhancement algorithms on speech signals contaminated with different environmental noises and with various SNR depends on the perception hearing of different volunteers. Figure 7 illustrates the bar chart of the evaluation percentages for the three-speech enhancement algorithms in the case of the male speech signal contaminated with cafeteria babble noise. The x-axis represents SNR in 15 dB, 10 dB and 0 dB, respectively, while the y-axis represents the percentage of the evaluation for each filter. Obviously, the subspace filter in the 15 dB and 10 dB have the highest impact (with 37.5% and 54.10% respectively) if compared with the effects of the other two filters (Wiener, and Spectral subtraction filters). In contrary, the Spectral subtraction filter shows the highest impact on the contaminated signal at 0 dB compared with the two other filters with 54.10%. In contrast, the subspace filter are degraded to 12.5% only at the same SNR.
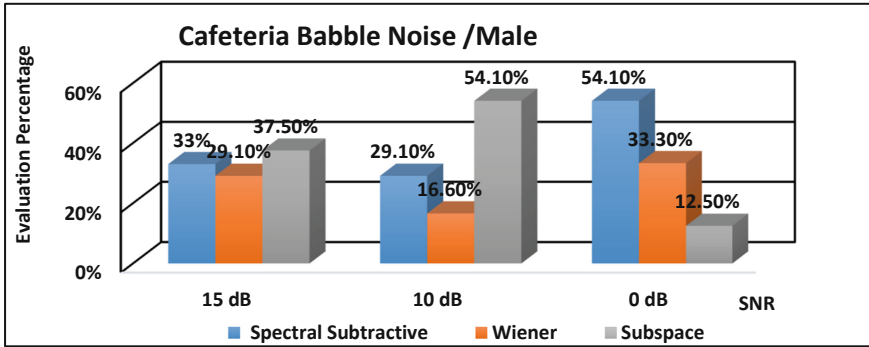
Fig. 7. Bar chart for the effectiveness of speech enhancement Algorithms on male speech signal contaminated with Cafeteria Babble noise.
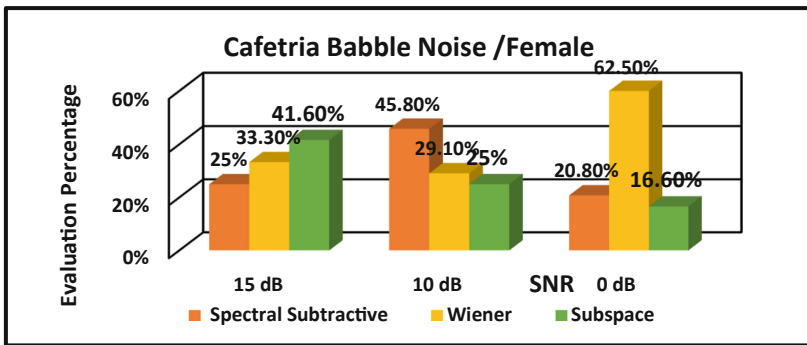

Fig. 8. Bar chart for the effectiveness of speech enhancement Algorithms on female speech signal contaminated with Cafeteria Babble noise.

On the other hand, and as seen in Fig. 8, which represents the effect of the same filters on female speech signal contaminated with Cafeteria babble noise, the sub-space filter still has the higher impact among the other filters at 15 dB with 41.6%. But, at 10 dB, we noticed that the spectral subtraction takes the large effect among filters with 45.8%, while the Wiener filter takes the best evaluation at 0 dB with 62.5%.
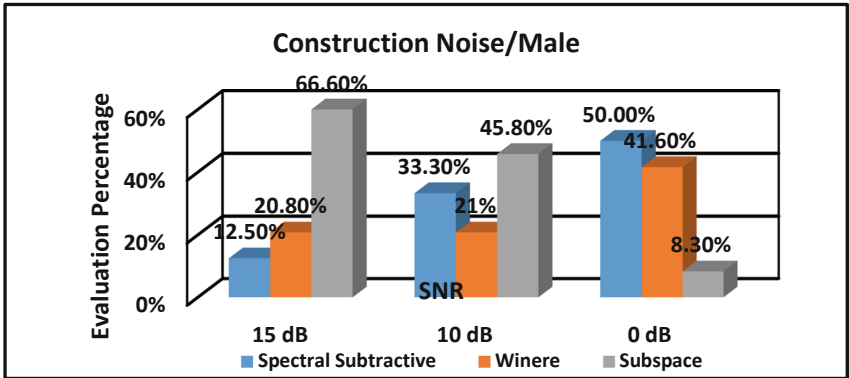
Fig. 9. Bar chart for the effectiveness of speech enhancement Algorithms on male speech signal contaminated with Construction noise.

Figure 9 illustrates the evaluation of the three filters when applied to male speech signal that contaminated with construction noise. It is also clear that the subspace filter still has the best evaluation at 15 and 10 dB compared with 66.6% and 45.88, respectively. On the other hand, Spectral subtraction and wiener filters get the better evaluation compared with sub-space at 0dB with 50% and 41.6%, respectively.

We can notice the same thing in Fig. 10, which illustrates the evaluation of the three filters on the female signal contaminated with the same noise. Subspace algorithm has the highest evaluation at 15 and 10 dB with 50%, 41.6%, respectively.
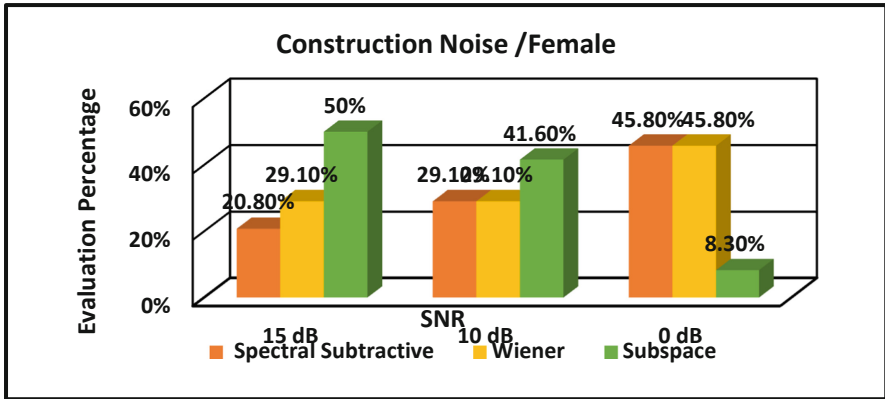

Fig. 10. Bar chart for the effectiveness of speech enhancement Algorithms on female speech signal contaminated with Construction noise.

On the contrary, Spectral subtraction and Wiener filter have the same evaluation at 0dB with 45.8%, while sup-space get only 8.3% at the same SNR. Figure 11 represents the bar chart of the evaluation of the three filters when applied on the male signal when

it is contaminated with Street noise. It is clear that the Sub-space filter has the highest effectiveness for cleaning the signal at 15dB with 45.8%. On the other hand, the Wiener filter and Spectral subtraction show the highest effect at 10 dB and 0 dB, respectively.
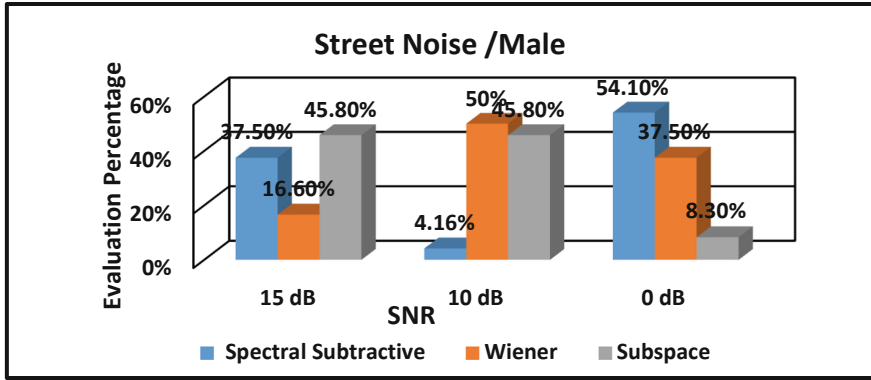

Fig. 11. Bar chart for the effectiveness of speech enhancement Algorithms on male speech signal contaminated with Street noise.

Almost the same effect is noticed on the female signal that contaminated with the same environmental noise (street noise), as demonstrated in Fig. 12. Table 1 demonstrates the overall evaluation for the three filters to improve the quality of noisy signals with different environments. In this table, for both male and female, the Spectral subtractive gave virtually the best results in 0 dB where the noise level is high. In contrast, the Subspace stratifies the best result 15 dB for both male and female.
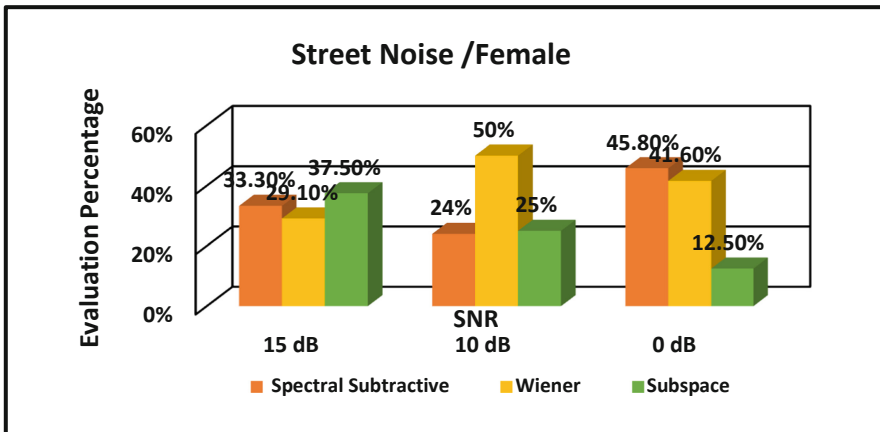

Fig. 12. Bar chart for the effectiveness of speech enhancement Algorithms on female speech signal contaminated with Street noise.

Table 1. Overall evaluation of questionnaires.

Filter	Babble noise					
	Male			Female		
	15 dB	10 dB	0 dB	15 dB	10 dB	0 dB
Spectral	33%	29.1%	54.1%	25%	45.8%	20.8%
Wiener	10%	16.6%	33.3%	33.3%	29.1%	62.5%
Subspace	37.5%	54.1%	12.5%	41.6%	25%	16.6%
Construction noise						
Spectral	12.5%	33.3%	50%	20.8%	29.1%	45.8%
Wiener	20.8%	21%	41.6%	29.1%	29.1%	45.8%
Subspace	66.6%	45.8%	8.3%	50%	41.6%	8.3%
Street noise						
Spectral	37.5%	4.16%	54.1%	33.3%	24%	45.8%
Wiener	16.6%	50%	37.5%	29.1%	50%	41.6%
Subspace	45.8%	45.8%	8.3%	37.5%	25%	12.5%

In summary, and based on human perception human auditory system, we can conclude the following:

1. Each noise has a different effect on the speech signal, making it challenging to select the best speech enhancement approach to improve speech signal quality. However, the Subspace filter shows the best quality improvement among the other filters at 15 dB and 10 dB. On the contrary, Spectral subtraction shows the best improvement for quality at 0 dB.
2. The impact of a speech enhancement algorithm for improving speech signal quality may vary from one signal to another at the same environmental noise at the same SNR. Mainly when the speech signals belong to different signal, as seen in the effect of Subspace filter on the signals contaminated with cafeteria babble noise at 10 dB.

6   Discussion and Conclusion

The essential goal of this work is to study the effects of speech cleaning algorithms for improving the quality of speech signals contaminated with different ambient noise with different signal to noise ratios SNRs depends on the human hear perception for these enhanced speech signals. The speech signals are used in this study are cleaned from various effects of the environment except for the environmental noise and collected from a different gender. Furthermore, the evaluation of the performance of these algorithms is achieved in a professional environment at the Multimedia lab. Three different types of noise are used in this experiment with controlled SNR. The results demonstrate that the Subspace algorithm performs better than the other two filters in terms of enhancing speech quality (Wiener, and Spectral subtraction) in most cases of 15 dB and 10 dB for different types of noise. The main reason that makes the Subspace

approach has the higher quality enhancement among the other two approaches is returned to it is natural, which is based on the linear algebra and the way with dealing with environmental noise. However, at 0 dB, the spectral subtraction algorithm shows the best performance for improving speech quality. Furthermore, the effect of these algorithms may vary according to the type of noise and type of speech signal that belongs to males and females. However, this study focuses mainly on a study the quality of cleaning speech but not the intelligibility. What is now needed is a study involving improving adaptive approaches to deal with quality and intelligibility at the same time.

References

1. Upadhyay, N., Karmakar, A.: The spectral subtractive-type algorithms for enhancing speech in noisy environments. In: 2012 1st International Conference on Recent Advances in Information Technology (RAIT) (2012)
2. Loizou, P.C.: Speech Enhancement: Theory and Practice. CRC Press, Boca Raton (2013)
3. Kamath, S., Loizou, P.: A multi-band spectral subtraction method for enhancing speech corrupted by colored noise. In: 2002 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) (2002)
4. Zhang, X., Wang, Z., Wang, D.: A speech enhancement algorithm by iterating single- and multi-microphone processing and its application to robust ASR. In: 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (2017)
5. Wang, Z.Q., Wang, P., Wang, D.: Complex spectral mapping for single- and multi-channel speech enhancement and robust ASR. IEEE/ACM Trans. Audio Speech Lang. Process. **28**, 1778–1787 (2020)
6. Soni, M.H., Patil, H.A.: Non-intrusive quality assessment of noise-suppressed speech using unsupervised deep features. Speech Commun. **130**, 27–44 (2021)
7. Loizou, P.C.: Speech enhancement based on perceptually motivated Bayesian estimators of the magnitude spectrum. Speech Audio Process. IEEE Trans. **13**(5), 857–869 (2005)
8. Al-Noori, A.: Robust speaker recognition in presence of non-trivial environmental noise (toward greater biometric security), in School of Computing, Science and Engineering. University of Salford (2017)
9. Weiss, M.R., Aschkenasy, E., Parsons, T.W.: Study and development of the INTEL technique for improving speech intelligibility. NICOLET Scientific Corp Northvale, NJ (1975)
10. Boll, S.: Suppression of acoustic noise in speech using spectral subtraction. IEEE Trans. Acoust. Speech Signal Process. **27**(2), 113–120 (1979)
11. Oh, I., Lee, I.: Speech enhancement system using the multi-band coherence function and spectral subtraction method (2019)
12. Lim, J.S., Oppenheim, A.V.: Enhancement and bandwidth compression of noisy speech. Proc. IEEE **67**(12), 1586–1604 (1979)
13. de la Hucha Arce, F., et al.: Adaptive quantization for multichannel Wiener filter-based speech enhancement in wireless acoustic sensor networks. Wirel. Commun. Mob. Comput. **2017**, 1-15 (2017)
14. Yanlei, Z., Shifeng, O., Ying, G.: Improved Wiener filter algorithm for speech enhancement. Autom. Control. Intell. Syst. **7**(3), 92 (2019)

15. Ephraim, Y., Malah, D.: Speech enhancement using a minimum-mean square error short-time spectral amplitude estimator. IEEE Trans. Acoust. Speech Signal Process. **32**(6), 1109–1121 (1984)
16. Mahmmod, B.M., et al.: Low-distortion MMSE speech enhancement estimator based on Laplacian prior. IEEE Access **5**, 9866–9881 (2017)
17. Wang, Y., Brookes, M.: Speech enhancement using an MMSE spectral amplitude estimator based on a modulation domain Kalman filter with a Gamma prior. In: 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (2016)
18. McAulay, R., Malpass, M.: Speech enhancement using a soft-decision noise suppression filter. IEEE Trans. Acoust. Speech Signal Process. **28**(2), 137–145 (1980)
19. Kjems, U., Jensen, J.: Maximum likelihood based noise covariance matrix estimation for multi-microphone speech enhancement. In: 2012 Proceedings of the 20th European Signal Processing Conference (EUSIPCO) (2012)
20. Takuya, Y., et al.: Maximum likelihood approach to speech enhancement for noisy reverberant signals. In: 2008 IEEE International Conference on Acoustics, Speech and Signal Processing (2008)
21. Aghazadeh, F., Tahan, A., Thomas, M.: Tool condition monitoring using spectral subtraction and convolutional neural networks in milling process. Int. J. Adv. Manuf. Technol. **98**(9–12), 3217–3227 (2018). https://doi.org/10.1007/s00170-018-2420-0
22. Weiss, M., Aschkenasy, E., Parsons, T.: Study and development of the INTEL technique for improving speech intelligibility. DTIC Document (1975)
23. Boll, S.: Suppression of acoustic noise in speech using spectral subtraction. Acoust. Speech Signal Process. IEEE Trans. **27**(2), 113–120 (1979)
24. Upadhyay, N., Karmakar, A.: Speech enhancement using spectral subtraction-type algorithms: a comparison and simulation study. Proc. Comput. Sci. **54**, 574–584 (2015)
25. Jae, L., Oppenheim, A.: All-pole modeling of degraded speech. IEEE Trans. Acoust. Speech Signal Process. **26**(3), 197–210 (1978)
26. Abd El-Fattah, M.A., et al.: Speech enhancement with an adaptive Wiener filter. Int. J. Speech Technol. **17**(1), 53–64 (2013). https://doi.org/10.1007/s10772-013-9205-5
27. Dendrinos, M., Bakamidis, S., Carayannis, G.: Speech enhancement from noise: a regenerative approach. Speech Commun. **10**(1), 45–57 (1991)
28. Ephraim, Y., Trees, H.L.V.: A signal subspace approach for speech enhancement. IEEE Trans. Speech Audio Process. **3**(4), 251–266 (1995)
29. Al-Noori, A.: Robust speaker recognition in presence of non-trivial environmental noise (toward greater biometric security). University of Salford (2017)