

Speech Enhancement based on Wiener Filter and Compressive Sensing

Amart Sulong^{*1}, Teddy Surya Gunawan^{*2}, Othman O. Khalifa³, Mira Kartiwi⁴,
Eliathamby Ambikairajah⁵

^{1,2,3}Department of Electrical and Computer Engineering, International Islamic University Malaysia (IIUM),
Malaysia

²Visiting Fellow, School of Electrical Engineering and Telecommunications, University of New South Wales
(UNSW), Australia

⁴Department of Information Systems, International Islamic University Malaysia (IIUM), Malaysia

⁵School of Electrical Engineering and Telecommunications, University of New South Wales (UNSW),
Australia

*Corresponding authors, e-mail: amartuia@gmail.com, tsgunawan@iium.edu.my

Abstract

In the last few decades, many advanced technologies have been proposed, in which communications played a great role as well as telecommunications applications. The noise elimination in various environments became the most concerned as it greatly hindered the speech communication applications. The improvement of noisy speech in terms of quality and intelligibility are taken into account without introducing any additional noise. Many speech enhancement algorithms have been proposed. Wiener filter is one of the classical algorithms that improve the noisy speech by reducing its noise components through selectively chosen Wiener gain. In this paper, compressive sensing method by randomize measurement matrix is combined with the Wiener filter to reduce the noisy speech signal to produce high signal to noise ratio. The PESQ is used to measure the quality of the proposed algorithm design. Experimental results show the effectiveness of our proposed algorithm to enhance noisy signals corrupted by various noises compared to other traditional algorithms, in which high PESQ scores were achieved across various noises and different SNRs.

Keywords: speech enhancement, Wiener filter, compressive sensing (CS), perceptual evaluation of speech quality (PESQ)

Copyright © 2016 Institute of Advanced Engineering and Science. All rights reserved.

1. Introduction

1.1. Speech Enhancement Algorithms

In advanced of today technologies enable to direct communication in a large distance, broader audiences, and more challenging circumstances. These fundamental principles lead to more crucial and provide a great interest to the scientists in getting to the field of speech enhancement [1]. Such as, the initial motivation of the interest area is to develop noise reduction algorithms that can be used to help hearing-impaired listeners (cochlear implant listeners) better communicate in noisy environments. It is motivated by improving perceptual aspects of speech that have been degraded by additive noise that corrupted speech [2]. However, there is always tradeoff between noise reduction and signal distortion – better noise reduction is always accompanied by larger signal distortion [3]. Hence, the main challenge in speech enhancements is to design effective algorithm to suppress the noise without introducing any perceptible distortion in the signal. The speech enhancement algorithms broadly introduced two types of speech distortion: the distortions that affect the speech signal itself called speech distortion and the distortions that affect the background noise called noise distortion [4-7]. Class of speech enhancement algorithms can be represented into three different speech enhancement methods used to date [2, 6-8], as will be explained in the following sections.

1.1.1. Spectral-Subtractive Algorithms

Spectral-Subtractive algorithms were proposed by Weiss et al. [2, 9] in the correlation domain and later by Boll [2, 10] in the Fourier transform domain. This noise estimation will be

evaluated during speech pause that normally happens in a normal conversation. It is widely known to suffer from perceptible artifacts by introducing musical noise. However, this method is the simplest enhancement algorithms to implement. The idea behind this basic principle is based on additive noise which can be estimated from the noisy spectrum when speech is not present and subtracts it from the noisy signal. The short-term spectral amplitude (STSA) has been exploited successfully in the development. These subtractive-type algorithms used STSA on the noisy speech input and recover an estimate of the clean STSA by removing the part contributed by the additive noise. The unprocessed phase of the noisy input signal is used to synthesize the enhanced speech signal under assumption that the human ear is not able to perceive the distortions in the phase of the speech signal [11]. Its enhanced signal is obtained by computing the inverse discrete Fourier transform of the estimated signal spectrum using the phase of the noise signal. In other words, the noise is assumed to be uncorrelated and additive to the speech signal. Its estimate of the noise signal is measured during silence or non-speech activity in the signal.

While the spectral subtraction method [11] can be easily implemented and effectively reduces the noise present in the corrupted signal, there exist some glaring shortcoming as the drawback of this algorithm. Its residual noise or musical noise is obvious that the effectiveness of the noise removal process is dependent on obtaining an accurate spectral estimate of the noise signal. The better the noise estimation, the lesser the residual noise content in the modified spectrum. However, since noise spectrum cannot be directly obtained. The noise removal process is forced to use an average estimate of the noise. Hence, there are some significant variations between the estimated noise spectrum and the actual noise content present in the instantaneous speech spectrum. The subtraction of these quantities results in the presence of isolated residual noise levels of large variance. This residual spectral contents manifest themselves in the reconstructed time signal as varying tonal sounds resulting in a musical disturbance of an unnatural quality. This musical noise can be even more disturbing and annoying to the listener than the original noise content. Several residual noise reduction algorithms have been proposed to overcome this problem. However, due to the limitations of the single-channel enhancement methods, it is not possible to remove this noise completely, without compromising the quality of the enhanced speech. Hence there is a tradeoff between the amount of noise reduction and speech distortion due to the underlying processing.

In addition, the distortion is also due to half/full wave rectification in the modified speech spectrum. It may contain some negative values due to the errors in estimated noise spectrum. These values are rectified using half-wave rectification (set to zero) or full-wave rectification (set to its absolute value). This can also lead to further distortions in the resulting time signal. Beside of that, the roughening of the speech due to the noisy phase affected the speech signal. The phase of the noise-corrupted signal is not enhanced before being combined with the modified spectrum to generate the enhanced time signal [12]. This is due to the fact that the presence of noise in the phase information does not contribute immensely to the degradation of the speech quality. This is especially true at high SNRs (>5dB). However, at the lower SNRs (<0dB), the noisy phase can lead to a perceivable roughness in the speech signal contributing to the reduction in speech quality. Estimating the phase of the clean speech is rather difficult and will greatly increase the complexity of the method. Moreover, the distortion due to noisy phase information is not very significant compared to that of the magnitude spectrum, especially for high SNRs. Hence the use of the noisy phase information is considered to be an acceptable practice in the reconstruction of the enhanced speech signal.

1.1.2. Statistical-Model-Based Methods and Wiener Filtering

It is a new speech enhancement method known as speech boosting. The method increases the relative power of the speech thus acting as a speech booster, instead of focusing on suppressing the noise. These speech enhancement algorithms [2, 6, 7] are posed in a statistical estimation framework. To find a linear (or nonlinear) estimator of the parameter of interest, namely the transform coefficients of the clean signal by given a set of measurements corresponding to the Fourier transform coefficients of the noisy signal. The Wiener filter and minimum mean-square error (MMSE) algorithms, among others, fall in this category. The area of this work was initiated by McAulay and Malpass [13], who proposed a maximum-likelihood approach for estimating the Fourier transform coefficients (spectrum) of the clean signal, and was followed by Ephraim and Malah [14], who proposed an MMSE estimator of the magnitude

spectrum. In addition, much work with the Wiener filter algorithm was initiated in the speech enhancement field by Lim and Oppenheim [15, 16]. Loizou [2] mention that the statistical-model focuses on nonlinear estimators of the magnitude (i.e. the modulus of the DFT coefficients) rather than the complex spectrum of the signal as done by the Wiener filter, using various statistical models and optimization criteria. These nonlinear estimators take the probability density function (PDF) of the noise and the speech DFT coefficients explicitly into account and use, in some cases, non-Gaussian prior distributions. These estimators are often combined with soft-decision gain modification that takes the probability of speech present into account.

A parameter of a statistical estimation framework in nonlinear estimator of interest depend on measurements correspond to the set of DFT coefficients of the noisy signal (i.e. the noisy spectrum) and the parameter of interest are the set of DFT coefficients of the clean signal (i.e. the clean signal spectrum). Various techniques exist in the estimation theory literature for deriving these nonlinear estimators and include the maximum-likelihood estimators. These estimators differ primarily in the assumptions made about the parameter of interest (e.g. deterministic but unknown, random) and the form of optimization criteria used. In [2], Loizou has mentioned the following algorithms: the maximum-likelihood estimator, an MMSE magnitude estimator, and a log-MMSE estimator. Bayesian estimators of the magnitude spectrum based on perceptually motivated distortion measure were also described. MAP estimators of the magnitude and phase spectra were presented. Several methods of incorporating speech-presence uncertainty in the proceeding estimators also discussed. These methods, when combined with the statistical estimators, substantially reduced the residual noise.

Furthermore, Yang [16] referred that speech enhancement in Wiener filter is also based on the Short Time Fourier Transform (STFT) technique, and used the same basic estimation principle as the spectral subtraction methods. The Wiener filter method can effectively reduce Gaussian noise. It is also used STFT in the Minimum Mean Square Estimation-Short Time Spectral Amplitude (MMSE-STSA) method. The method assumes that the noisy speech STFT coefficients for continuous frames are independent Gaussian variable, which can be statistically modeled to estimate the clean speech spectrum.

1.1.3. Subspace Algorithms

Unlike the preceding algorithms, the subspace algorithms are rooted primarily from linear algebra theory. In addition, vector subspace technique used STFT-based techniques for speech enhancement method [17, 18]. A vector subspace technique usually has the following measurement step to improve of speech signal. At first, the noisy speech is decomposed into a vector space. Then the noisy speech vector space is divided into a signal subspace and noise subspace. Finally, the noise subspace is removed and speech signal is reconstructed from the signal subspace. There are several transformation techniques using for vector subspace to speech enhancement.

Most of researches commonly used the Karhunen-Loeve Transform (KLT) and the discrete cosine transform (DCT) for noisy speech decomposition. KLT is an optimal Eigen decomposition technique, but DCT is more computationally efficient. In general, the vector subspace [17, 18] usually uses a Laplace model or Gaussian model to describe the signal subspace, and uses a Gaussian model to describe the noise subspace. In addition, the speech signal [17] degraded by uncorrelated additive noise based on the vector subspace of the noisy signal that can be decomposed into a signal plus noise subspace and the orthogonal noise subspace. Decomposition of the vector space of the noisy signal is performed by applying an eigenvalue or singular value decomposition or by applying the Karhunen-Loeve transform (KLT). The processing is only performed on the vectors in the single subspace while the noise subspace is removed first. The idea of this approach is that noisy speech frames are classified into speech-dominated frames. In speech dominated frames, the signal Karhunen-Loeve transform (KLT) matrix is used, and in the noise-dominated frames, the noise KLT matrix is used.

1.2. Compressive Sensing

Compressive sensing (CS) is a fundamentally new approach to data acquisition approach and a new type of sampling theory which predicts that sparse signals can be reconstructed from what previously believed to be incomplete information [19]. The theory asserts that one can recover certain signal from far fewer samples or measurements than

traditional methods use [20]. This CS theory relies on the empirical observation that many type of signals can be well approximated by sparse expansion in terms of suitable basis. The traditional approach of reconstructing signals from measured data follows the well-known Shannon sampling theorem [21]. Many solutions to sparse approximation have been proposed, such as matching pursuit (MP), least absolute shrinkage and selection operator (LASSO), basis pursuit (BP), and gradient pursuit (GP), in which of its performance show some interdependence between the number of measurement noise, signal sparsity and the reconstruction algorithm [23].

The CS can be explained by consider a real-valued, finite-length, one-dimensional, discrete-time signal x which can viewed as an $N \times 1$ column vector in $x \in R^N$ with elements $x[n], n = 1, 2, 3, \dots, N$ and treat it to higher dimensional data by vectorizing it into a long one-dimensional vector. Any signal x can be represented in terms of a basis of $N \times 1$ vectors $\{\psi_i\}_{i=1}^N$. For simplicity, assume that the basis matrix $\psi = [\psi_1 \psi_2, \dots, \psi_N]$ is the certain domain of the transform matrix with the vectors $\{\psi_i\}$ as columns and generally view as transform domain, i.e. Wavelet transform (WT), discrete cosine transform (DCT) and discrete Fourier transform (DFT). A signal x can be expressed as

$$x = \sum_{i=1}^N s_i \psi_i \quad \text{or} \quad x = \psi s \quad (1)$$

where s is the $N \times 1$ column vector of weighting coefficients $s_i = \langle x, \psi_i \rangle = \psi_i^T x$ and $(\cdot)^T$ denotes transposition. Clearly, x is in the ψ domain.

In the CS method [22], the K -sparse signal represents the foundation forms of the transform coding that can compress signals which approximated well in data acquisition systems. This transform coding plays a central role to N sample of the data signal x . This CS approach addresses the inefficiencies of classical approach that introduced by Shannon-Nyquist theorem by directly acquiring a compressed signal representation without going through the intermediate state of acquiring N sample. Consider a general linear measurement process that computes $M \ll N$ inner products between x and a collection of vectors $\{\psi_i\}_{i=1}^M$ as in $y_j = \langle x, \phi_j \rangle$. Arrange the measurements y_j in an $M \times 1$ vector y and measurement vector ϕ_j^T as rows in an $M \times N$ matrix ϕ . Then by substituting ψ from the (1), y can be written as

$$y = \phi x = \phi \psi s = \Theta s \quad (2)$$

where $\Theta = \phi \psi$ is $M \times N$ matrix of random linear which represent the measurement process and typically $M > \text{const} \times K \log \left(\frac{N}{K} \right)$. The measurement process is not adaptive, meaning that ϕ is fixed and does not depend on signal x . The problem consists of designing a stable measurement matrix ϕ such that the salient in any K -sparse or compressible signal is not damaged by the dimensionality reduction from $x \in R^N$ and $y \in R^M$ and a reconstruction algorithm to recover x from only measurements y (or about as many measurements as the number of coefficients recorded by tradition transform coder (see Figure 1)).

In CS's sparsity of the desired signal with sparse representation in a known transform domain. Number of significant (strictly speaking nonzero) components is relatively small compared to signal length. The sparsity representation in the form of l_0 norm, l_1 norm and l_p norm $= \|x\|_p = (\sum_{i=1}^N |x|^p)^{\frac{1}{p}}$ that count the number of nonzero component of x . This CS can compressible signal down to a much smaller observation space by using appropriate observation matrix, then the non-linear reconstruction techniques developed for building sparse representations that can be used to decode the signal. Furthermore, it has been shown that both in terms of the number of samples and the number of bit required to encode the samples, compressive sensing can be almost as efficient as using a sparse transform domain representation with traditional sampling with low margin of the error for the reconstruction.

2. Proposed Speech Enhancement Algorithm

Various speech enhancement algorithms have been proposed to improve the performance of modern communication device in noisy environments. The background noise level and the characteristics are constantly changing in a real environment. The elusion of the noisy signal that is reliable and fair comparison between algorithms have been emerged. There are several researches show that the fatigue and exhaustion of the signal depends on the lack of common speech database for evaluation of new algorithms, differences in the types of noise useand differences in testing methodology. Furthermore, understanding the speech characteristics and a common speech database will help in designing speech enhancement algorithms to access to nearly possible for researchers to compare at very least the objective performance of their algorithms with that of others.

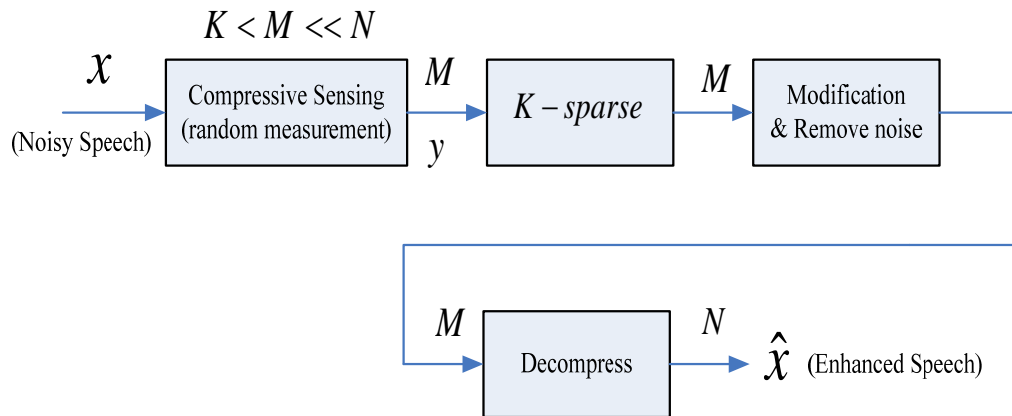


Figure 1. The compressive sensing approach for sensing the measurement matrix.

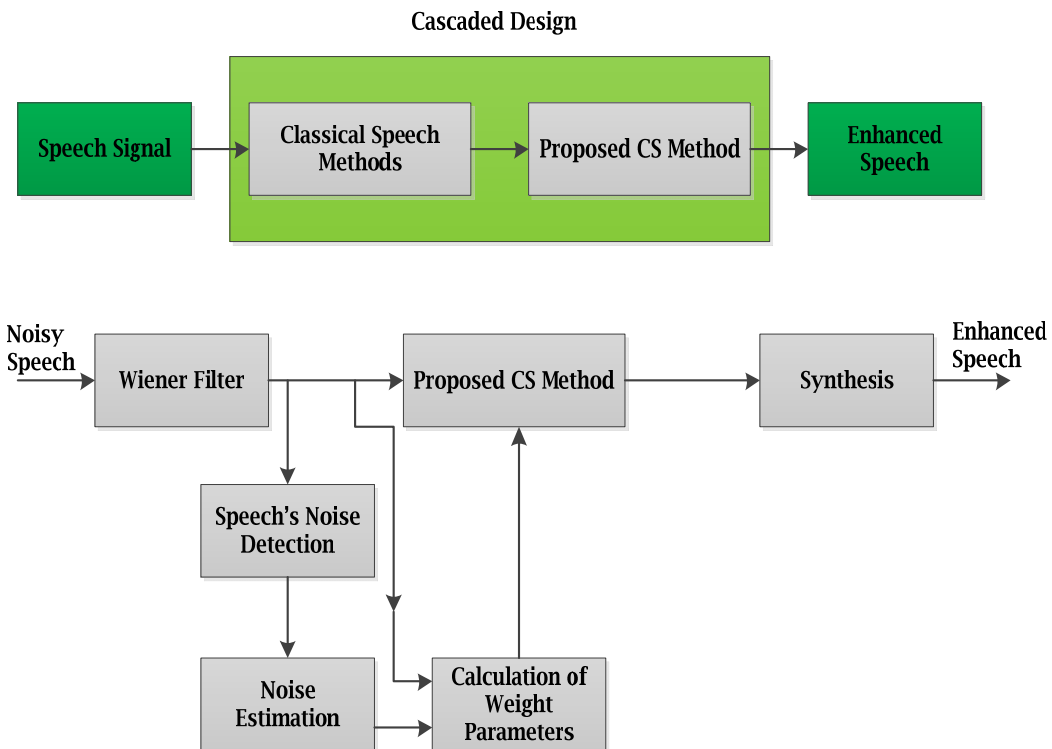


Figure 2. The proposed speech enhancement based on Wiener filter and compressive sensing

Figure 1 shows the CS modification for the speech signal to eliminate the noise. Figure 2 shows the proposed algorithm that applied Wiener Filter and compressive sensing for the speech enhancement process. In the proposed algorithm, as shown in Figure 2, it is started with initial state by acquiring noisy speech then measure the noisy signal in Wiener filter to obtain noise estimation and estimate speech and calculate the gain parameters of the speech and noise. The compressive sensing (CS) by using Gradient projection for sparse reconstruction (GPSR) algorithm [24] will measure the value of the noise reduction and producing the estimation of speech signal. Then synthesis block will produce the enhanced speech signal. This speech enhancement signal later will be evaluated using PESQ to measure the quality of the enhanced speech.

In Figure 1, the processing of the CS will use GPSR by measuring the signal following equation (3) to estimate the clean speech signal.

$$\min_x \frac{1}{2} \|y - Ax\|_2^2 + \tau \|x\|_1 \quad (3)$$

where $x \in R^N$, $y \in R^M$, and A is $k \times n$ matrix. The τ is a nonnegative parameter, $\|v\|_1$ refer to the l_1 norm of v , and $\|v\|_2$ refers the Euclidean norm of v . Equation (3) is related to the following convex constrained optimization problems

$$\min_x \|x\|_1 \quad \text{subject to} \quad \|y - Ax\|_2^2 \leq \varepsilon \quad (4)$$

and

$$\min_x \|y - Ax\|_2^2 \quad \text{subject to} \quad \|x\|_1 \leq t \quad (5)$$

where ε and t are nonnegative real parameters. It was utilized due to it reconstruction quality to trade with available processing power at inverse transform domain and then synthesi back to gain the enhancement of the speech signal. At the end of the process, the measurement of the quality of speech signal also proposed by using the perceptual evaluation of speech quality (PESQ) score [2].

3. Results and Discussion

The proposed algorithm and other algorithm were utilized its performance levels using objective measure of PESQ score of ITU-T P.862 to achieve the main objective of the enhanced speech signal [25]. Its objective PESQ correlation with subjective test is 93.5% compare with other objective test [2]. The PESQ objective assessment tests was evaluated at four different type's noise, i.e. babble, car, exhibition, restaurant noise respectively, under 0, 5, 10, and 15 dB SNR. New speech quality assessment test is introduced in [7], in terms of percentage PESQ improvement (δ) and can be expressed as follows

$$\delta = \frac{PESQ_{proc} - PESQ_{ref}}{PESQ_{ref}} \times 100\% \quad (6)$$

where $PESQ_{proc}$ is defined as PESQ score of the enhanced speech. $PESQ_{ref}$ is defined as the PESQ score of the clean speech as the reference speech respectively. Its improvement δ is also evaluated based on noise corrupted to the speech signal within various environments and SNRs. Its objective measures used the noisy speech corpus (NOIZEUS) of IEEE subcommittee 1996 standard [2]. Other traditional algorithms are original Wiener filter algorithm [26], spectral subtraction (specsub) [27], ss_rdc [28], logmmse_SPU [29], and klt [30].

Figure 3 shows the comparison of the enhanced speech signal of the proposed algorithm and other traditional methods. At various enviroments of noise attack to the speech signal, the proposed algorithms produced the best result than traditional methods in term of speech wave form while klt and ss_rdc algorithm are highly distorted the speech signal. Figure 4 clearly presents the worse case scenario for klt because it suppressed most identity of speech

signal and also in Figure 5. The overall proposed algorithm in Figure 3, Figure 4, and Figure 5 performed the best improvement among other algorithms. In other words, logmmse_SPU was observed with acceptable result.

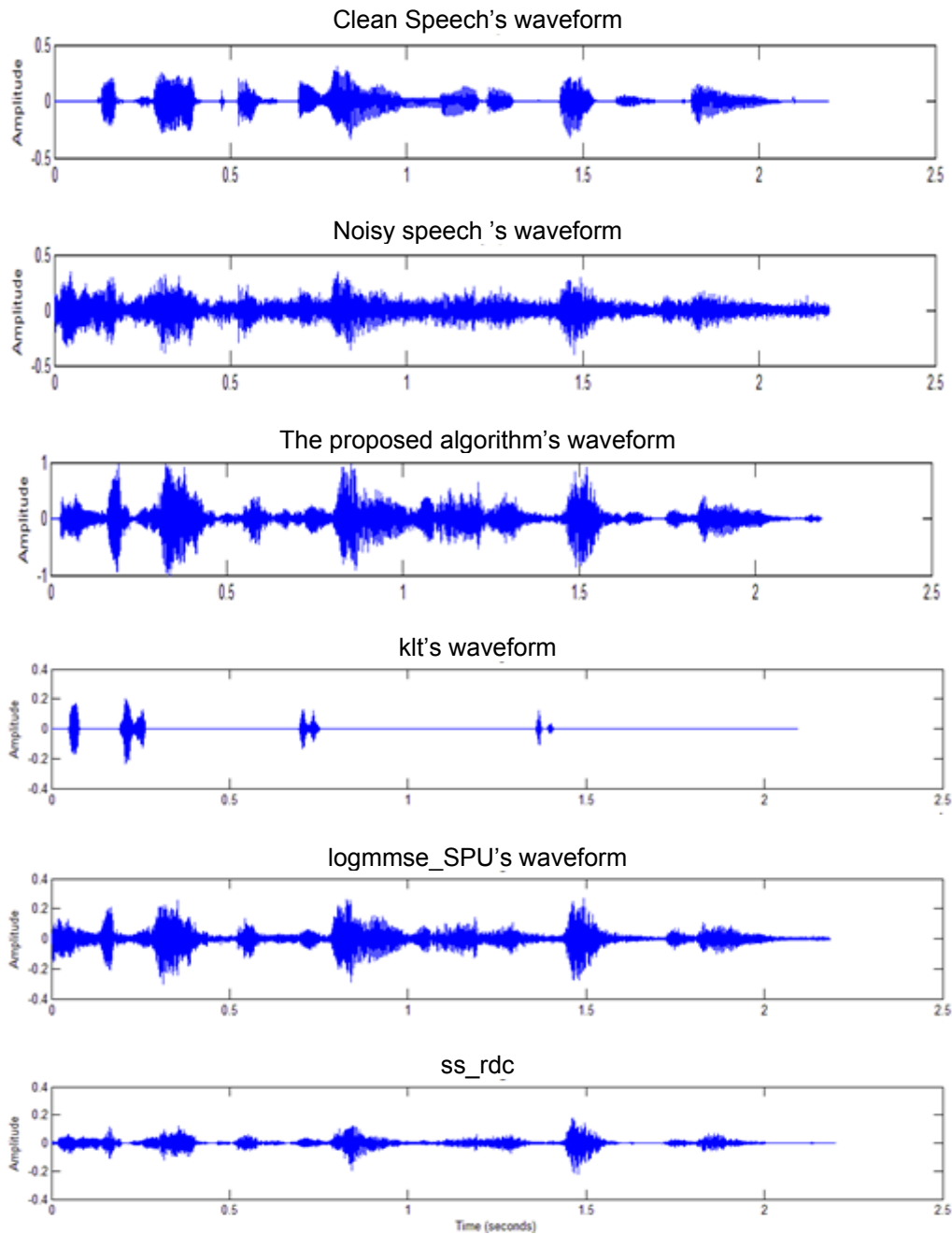


Figure 3. Comparison of the enhanced speech waveform of the proposed algorithm with other algorithm of the babble noise “sp9.wav” at 0 dB SNR

Figure 5 represents comparison of the PESQ score of the proposed algorithm with traditional methods at various noise condition, i.e. restaurant, exhibition, car, babble noise of 0, 5, 10, and 15 dB SNR. The PESQ score in restaurant and exhibition noise of the proposed algorithm outperforms than traditional method. Particularly, the enhance speech of restaurant noise at 0 dB SNR produced lower PESQ score comparing to PESQ score of noisy. However,

when the dB SNR were increased to 5, 10, 15 dB SNR, the PESQ score results with better performance level in term of speech quality especially in the proposed algorithm. Most of traditional methods in 5 and 10 dB SNR show the PESQ scores close to the PESQ of the Noisy Speech except in klt and Wiener algorithm.

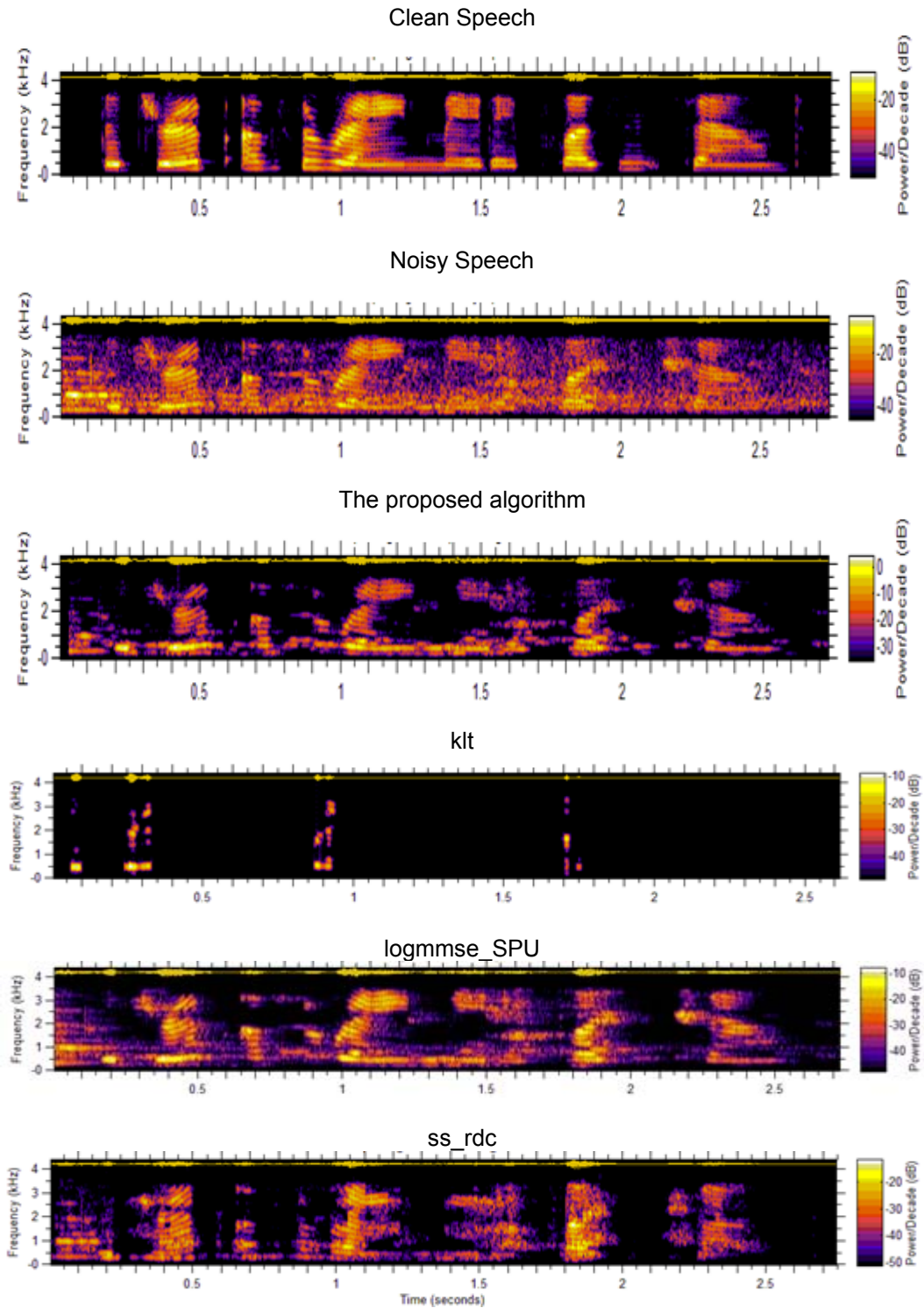


Figure 4. The spectrograms of the proposed algorithm compare with other algorithm of the babble noise "sp9.wav" at 0 dB SNR

In Figure 5, the car and babble noise in the proposed algorithm given the best results when compare among other traditional methods. The the performed PESQ scores of traditional methods slightly can be competed when it compared with PESQ score of noisy. Only in Specsusb, logmmse, and klt at 15 dB SNR, the traditional methods produced better PESQ score than the proposed algorithm. In Babble noise at 0 dB SNR, most the traditional methods were lower than PESQ score of Noisy but the proposed algorithm outperforms than other methods. Particularly, the proposed algorithm in babble noise can clearly be observed its best performance score comparing with others methods and noisy.

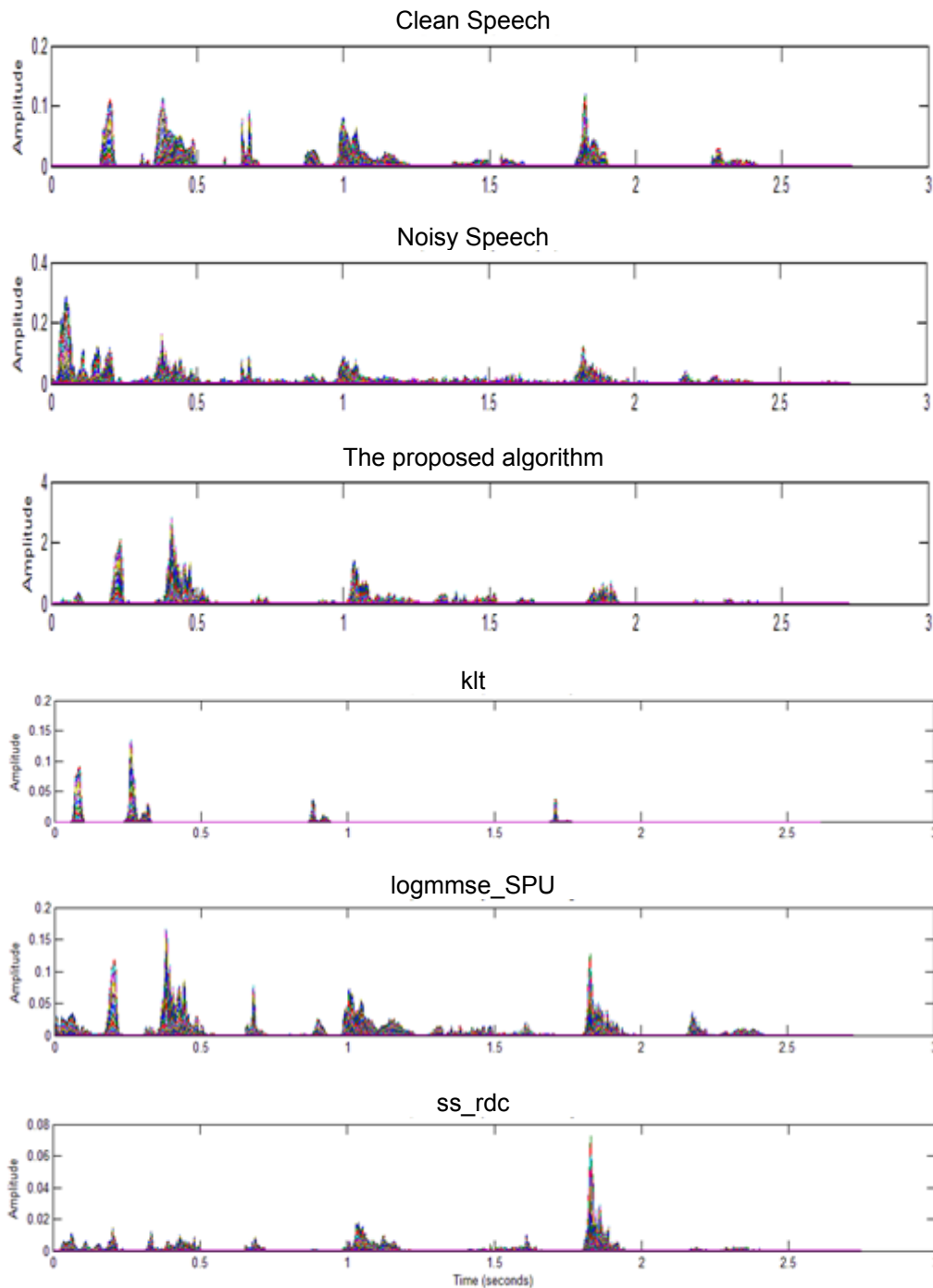


Figure 5. The spectrum density of the proposed algorithm compared to other algorithms of the babble noise “sp9.wav” at 0 dB SNR

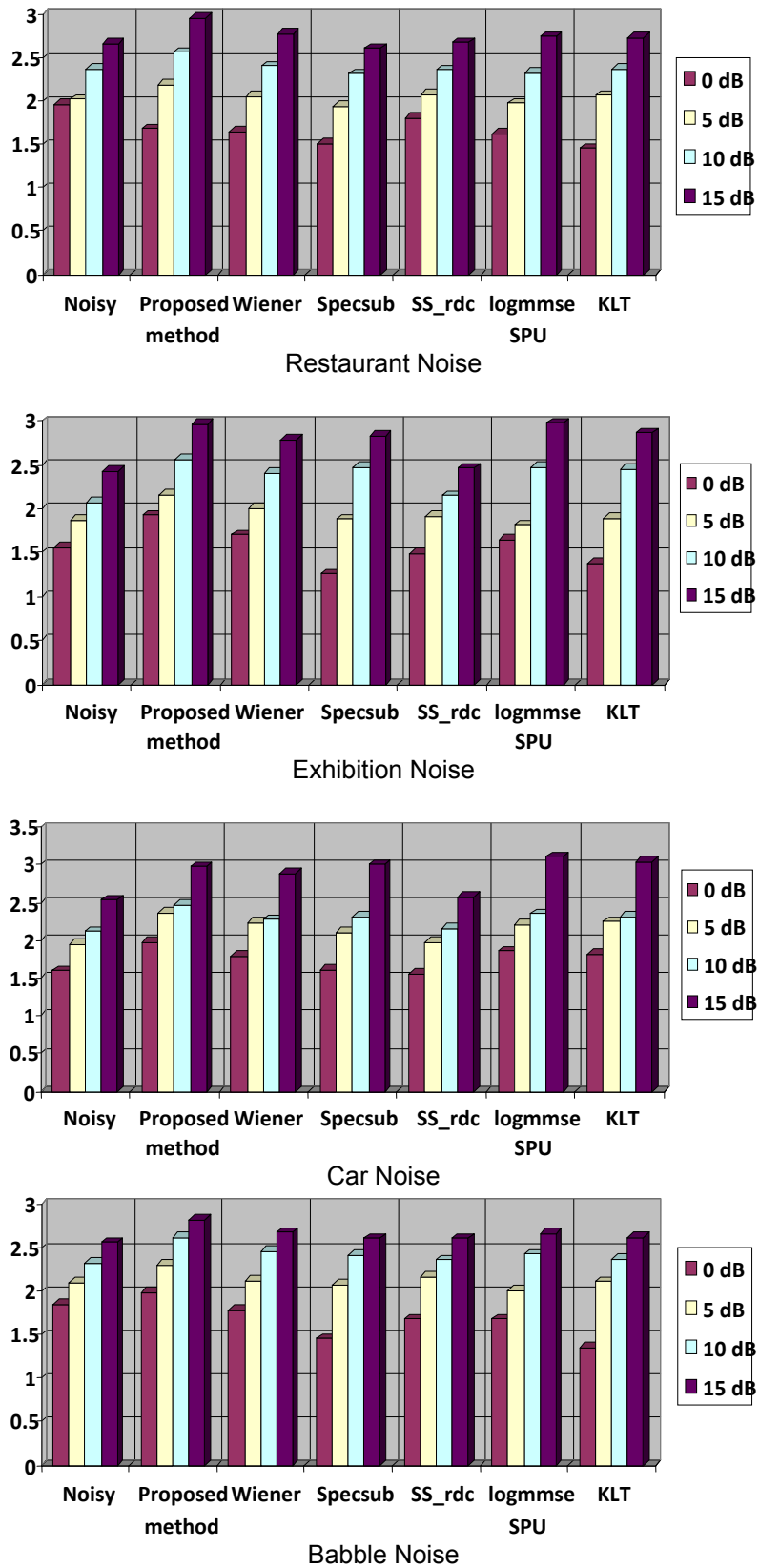


Figure 6. PESQ score comparison of the proposed algorithm compared to the other algorithms

Table 1 presents the percentage PESQ improvement (%) at various noise type of 0, 5, 10, 15 dB SNR. It shown that the proposed algorithm outperform than traditional algorithm when compared all types of noisy conditions, i.e. babble, car, exhibition, and restaurant noise of various SNRs. The worst case can be observed in babble and restaurant noise of 0 dB SNR that the most of methods did not improve percentage of the speech quality. However, when its noisy is increased to 5, 10, and 15 dB SNR the results of proposed and traditional algorithm produced better improve the speech quality. Especially, the percentage PESQ improvement of proposed method performs the best performance improvement. In babble noise gives highest improvement results of 6.75%, 9.75%, 13.21%, and 10.71% in respect to 0, 5, 10, 15 dB SNR. In the car noise at 0, 5, and 10 dB SNR produced 23.45%, 21.34%, and 16.42% respectively.

Table 1. The percentage PESQ improvement (%) at various noise type of 0, 5, 10, 15 dB SNR

Speech	Babble Noise	Wiener Filter	PESQ Improvement in Percentage (%)				Proposed Method
			specsub	ss_rdc	logmmse_SPU	klt	
Babble Noise	0 dB	-4.03	-21.60	-9.37	-9.30	-27.33	6.75
	5 dB	1.30	-0.97	3.15	-4.30	0.54	9.75
	10 dB	6.23	4.42	1.79	4.64	2.23	13.21
	15 dB	4.41	1.78	1.95	4.09	2.05	10.17

Speech	Car Noise	Wiener Filter	PESQ Improvement in Percentage (%)				Proposed Method
			specsub	ss_rdc	logmmse_SPU	klt	
Car Noise	0 dB	12.33	0.85	-2.11	16.23	14.20	23.45
	5 dB	14.07	7.88	0.84	13.03	14.79	21.34
	10 dB	7.56	9.63	1.89	11.19	9.25	16.42
	15 dB	13.77	18.35	1.24	22.39	19.95	17.58

Speech	Exhibition Noise	Wiener Filter	PESQ Improvement in Percentage (%)				Proposed Method
			specsub	ss_rdc	logmmse_SPU	klt	
exhibition Noise	0 dB	9.39	-19.27	-3.95	6.21	-11.39	23.80
	5 dB	6.99	0.32	2.50	-3.03	1.42	15.75
	10 dB	16.13	19.74	3.74	19.52	18.01	23.99
	15 dB	14.84	16.85	1.76	22.79	18.33	22.06

Speech	Restaurant Noise	Wiener Filter	PESQ Improvement in Percentage (%)				Proposed Method
			specsub	ss_rdc	logmmse_SPU	klt	
restaurant Noise	0 dB	-16.62	-23.03	-8.59	-17.79	-26.02	-14.62
	5 dB	1.79	-3.81	2.97	-2.22	2.28	8.63
	10 dB	1.02	-2.66	-0.59	-2.13	-0.36	7.86
	15 dB	4.12	-2.11	0.16	2.82	2.55	10.66

In exhibition noise given highest improvement at 0, 5, and 10 dB SNR which was produced 23.80%, 15.75%, and 23.99% respectively. However, in exhibition noise at 15 dB SNR produced 22.06% which was better among traditional method excepted at logmmse_SPU (22.79%). In restaurant noise at 5, 10, 15 dB SNR performed the best results excepted at 0 dB did not improve speech quality which is produce -14.62% but still better than traditional methods excepted at ss_rdc produced -8.59%.

4. Conclusions

In this proposed algorithm using the cascaded design of Wiener filter and CS has met the goals of speech enhancement algorithm. This algorithm can significantly reduce the noise in noisy speech and produce the enhanced speech signal to the listener ends. The overall

improvement of the performance system can be clearly seen by comparison of the PESQ scores and its percentage improvement of the proposed algorithm to traditional algorithms. In other words, this speech enhancement quality provided the useful task to satisfy that any noise condition attacked to system are not much distort the speech signal. Once, the objective evaluation of the proposed algorithm produced high performance results without reducing listener fatigue at the receiving ends. Therefore, the main objective of the proposed algorithm is achieved when compare the objective tests with other traditional algorithms.

References

- [1] Kamath SD. A Multi-Band Spectral Subtraction Method for Speech Enhancement. Master of Science in Electrical Engineering, the University of Texas at Dallas. 2001.
- [2] Loizou P. Speech Enhancement: Theory and Practice. First Edition. New York: CRC Press. 2007: 1-10.
- [3] Sudirga R. A Speech Enhancement System based on Statistical and Acoustic-Phonetic Knowledge. Master of Science (Engineering), Queen's University, Kingston, Ontario, Canada.
- [4] Hu Y, Loizou P. Evaluation of Objective Quality Measures for Speech Enhancement. *IEEE Transactions on Audio, Speech, and Language Processing*. 2008; 16(1): 229-238.
- [5] Hu Y, Loizou P. *Subjective Comparison of Speech Enhancement Algorithms*. In Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. 2006; 1: 153-156.
- [6] Hu Y, Loizou P. Subjective Comparison and Evaluation of Speech Enhancement Algorithms. *Speech Commun.* 2007; 49: 588-601.
- [7] Gunawan TS. Audio Compression and Speech Enhancement Using Temporal Masking Models. Doctor of Philosophy, the University of New South Wales. 2007.
- [8] Chen RF, Chan CF, So H. Speech Enhancement in Car Noise Environment Based on Analysis-Synthesis Approach Using Harmonic Noise Model. *IEEE ICASSP*. 2009: 4413-4416.
- [9] Weiss M, Aschkenasy E, Parsons T. *Study and the Development of the INTEL Technique for Improving Speech Intelligibility*. Technical report NSC-FR/4023, Nicolet scientific corporation. 1974.
- [10] Boll SF. Suppression of Acoustic Noise in Speech Using Spectral Subtraction. *IEEE trans. Acoust. Speech Signal Process.* 1979; 27(2): 113-120.
- [11] Lu Y, Loizou P. A Geometric Approach to Spectral Subtraction. *Speech Communication, Science Direct*. 2008; 50: 1-15
- [12] Zoghalmi N, Lachiri Z, Ellouze N. Speech Enhancement Using Auditory Spectral Attenuation. 17th European Signal Processing Conference (EUSIPCO). 2009; 223-227.
- [13] McAulay RJ, Malpass ML. Speech Enhancement Using a Soft-Decision Noise Suppression Filter. *IEEE Trans. Acoust. Speech Signal Process.* 1980; 28: 137-145
- [14] Ephraim Y, Malah D. Speech Enhancement Using a Minimum Mean Square Error Short-Time Spectral Amplitude Estimator. *IEEE Trans. Acoust. Speech Signal Process.* 1984; 32(6): 1109-1121.
- [15] Lim J, Oppenheim AV. *Enhancement and Bandwidth Compression of Noisy Speech*. Proc. IEEE. 1979; 67(12): 1586-1604.
- [16] Yand X. Auditory Domain Speech Enhancement. Master of Science (Engineering). Queen's University, Kingston, Ontario, Canada. 2008.
- [17] Ephraim Y, Van Trees HL. A Signal Subspace approach for Speech Enhancement. *IEEE Transactions on Speech and Audio Processing*. 1995; 3(4): 251-266.
- [18] Gazor S, Wei Z. Speech Enhancement Employing Laplacian-Gaussian Mixture. *IEEE Transactions on Speech and Audio Processing*. 2005; 13(5): 896-904.
- [19] Fornasier M, Rauhut H. *Compressive Sensing*. In O. Scherzer (Ed.). Handbook of Mathematical Methods in Imaging. Springer Science Business Media LLC. 2011.
- [20] Candes EJ, Watkin MB. An Introduction to Compressive Sampling. *IEEE Signal Processing Magazine*. 2008; 21-30.
- [21] Unser M. *Sampling – 50 Years after Shannon*. Proc. IEEE. 2000; 88(4): 569-587.
- [22] Baraniuk RG. Compressive Sensing. *IEEE Signal Processing Magazine*. 2007; 118: 120-124.
- [23] Gunawan TS, Khalifa OO, Shafie AA. *Speech Compression Using Compressive sensing on a Multicore system*. Proceedings of the 4th International Conference on Mechatronics (ICOM). Kuala Lumpur. 2011.
- [24] Figueiredo MAT, Nowak RD, Wright SJ. Gradient Projection for Sparse Reconstruction: Application to Compressed Sensing and Other Inverse Problems. *IEEE Journal of Selected Topics in Signal Processing*. 2007; 1: 586-597.
- [25] ITU-T RP 862. Perceptual evaluation of speech quality (PESQ): an objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs. International Telecommunication Union-Telecommunication Standardisation Sector.
- [26] Berouti M, Schwartz M, Makhoul J. *Enhancement of speech corrupted by acoustic noise*. Proc. IEEE Int. Conf. Acoust. Speech Signal Process. 1979; 208-211.

- [27] Gustafsson H, Nordholm S, Claesson I. Spectral subtraction using reduced delay convolution and adaptive averaging. *IEEE Trans. Speech Audio Process.* 2001; 9(8): 799-807.
- [28] Cohen I. Optimal speech enhancement under signal presence uncertainty using log-spectra amplitude estimator. *IEEE Signal Process. Lett.* 2002; 9(4): 113-116.
- [29] Hu Y, Loizou P. A generalized subspace approach for enhancing speech corrupted by colored noise. *IEEE Trans. Speech Audio Process.* 2003; 11: 334-341.