University of Baghdad
College of Engineering
JE
JOURNAL OF ENGINEERING

Journal of Engineering

journal homepage: www.joe.uobaghdad.edu.iq

Volume 29 Number 9 September 2023

Dual Stages of Speech Enhancement Algorithm Based on Super Gaussian Speech Models

Humam Awad Hussein[1] , Shams Moaied Hameed[2], Basheera M. Mahmmod[3], Sadiq H. Abdulhussain[4,*], Abir Jaafar hussain[5]

[1,2,3,4] Department of Computer Engineering, College of Engineering, Baghdad, Iraq
[5] School of Computer Science and Mathematics, Liverpool John Moores University, Liverpool, UK
h.hussain1805@coeng.uobaghdad.edu.iq[1], s.hamid1805@coeng.uobaghdad.edu.iq[2], basheera.m@coeng.uobaghdad.edu.iq[3], sadiqhabeeb@coeng.uobaghdad.edu.iq[4], a.hussain@ljmu.ac.uk[5]

ABSTRACT

Various speech enhancement Algorithms (SEA) have been developed in the last few decades. Each algorithm has its advantages and disadvantages because the speech signal is affected by environmental situations. Distortion of speech results in the loss of important features that make this signal challenging to understand. SEA aims to improve the intelligibility and quality of speech that different types of noise have degraded. In most applications, quality improvement is highly desirable as it can reduce listener fatigue, especially when the listener is exposed to high noise levels for extended periods (e.g., manufacturing). SEA reduces or suppresses the background noise to some degree, sometimes called noise suppression algorithms. In this research, the design of SEA based on different speech models (Laplacian model or Gaussian model) has been implemented using two types of discrete transforms, which are Discrete Tchebichef Transform and Discrete Tchebichef-Krawtchouk Transforms. The proposed estimator consists of dual stages of a wiener filter that can effectively estimate the clean speech signal. The evaluation measures' results show the proposed SEA's ability to enhance the noisy speech signal based on a comparison with other types of speech models and a self-comparison based on different types and levels of noise. The presented algorithm's improvements ratio regarding the average SNRseq are 1.96, 2.12, and 2.03 for Buccaneer, White, and Pink noise, respectively.

Keywords: Speech Enhancement Algorithms (SEA), Gaussian speech model, Laplacian speech model, Discrete Tchebichef Transform (DTT), Discrete Tchebichef-Krawtchouk Transform (DTKT).

*Corresponding author
Peer review under the responsibility of University of Baghdad.
https://doi.org/10.31026/j.eng.2023.09.01

Article received: 28/03/2023
Article accepted: 18/07/2023
Article published: 01/09/2023

خوارزمية تحسين الكلام ثنائية المراحل بالاعتماد على تمثيل الكلام بتوزيع كاوسين الفائق

همام عواد حسين[1]، *، شمس مؤيد حميد[2]، بشيرة محمد رضا محمود[3]، صادق حبيب عبدالحسين[4]*، عبير جعفر حسين[5]

[1،2،3،4]قسم هندسة الحاسبات، كلية هندسة ، جامعة بغداد، بغداد، العراق

[5]علوم الحاسبات والرياضيات، جامعة ليفربول جون مورس، ليفربول، المملكة المتحدة

الخلاصة

مختلف خوارزميات تحسين الكلام (SEA) تم تطويرها في العقود القليلة الماضية. كل خوارزمية لها مزاياها وعيوبها وذلك كون الكلام يتأثر بالظروف البيئية بطرق مختلفة. يؤدي تشويه الكلام إلى فقدان ميزات مهمة تجعل من الصعب فهم هذه الإشارة. تهدف خوارزمية تحسين الكلام (SEA) إلى تحسين وضوح وجودة الكلام التي تتاثر بانواع مختلفة من الضوضاء. في معظم التطبيقات، يكون تحسين جودة الكلام أمرًا مرغوبا للغاية لأنه يقلل من جهد المستمع لسماع الصوت بوضوح، لا سيما في المواقف التي يتعرض فيها المستمع لمستويات عالية من الضوضاء ولفترات طويلة من الزمن (مثل المصانع). يقلل SEA أو يمنع الضوضاء الخلفية إلى حد ما ويطلق عليه أحيانًا خوارزميات قمع الضوضاء. في هذا البحث، تم تصميم خوارزمية لتحسين الكلام استنادا إلى نماذج تمثيل مختلفة وهي (نموذج لابلاس او نموذج كاوسين) وبناءً على نوعين من متعددات الحدود المتعامدة (OP) وهما تحويل Discrete Tchebichef Transform (DTT) وتحويل Discrete Tchebichef-Krawtchouk (DTKT). علاوة على ذلك، فان العمل المقترح يتكون من مرحلتين من مرشح ونر الذي له قابلية كبيرة في تخمين اشارة الكلام الاصلية. تظهر نتائج التقييم قدرة النظام المقترح على تحسين إشارة الكلام الصاخبة بناءً على المقارنة مع أنواع أخرى من نماذج الكلام بالإضافة إلى المقارنة الذاتية التي أجريت لإظهار قدرة كل خوارزمية في تحسين الإشارة الكلام الصاخبة والذي تم بالاعتماد على أنواع ومستويات ضوضاء مختلفة. على سبيل المثال ، معدل التحسن في SNRseq بالنسبة لضوضاء ال Buccaneer، white، وpink هي 1.96 ,2.12 2.03 على التوالي.

الكلمات المفتاحية: خوارزمية تحسين الكلام (SEA)، نموذج كاوسين للكلام، نموذج لابلاسين للكلام ، تحويل Tchebichef المنفصل (DTT) ، تحويل Tchebichef-Krawtchouk المنفصل (DTKT).

1. INTRODUCTION

The advancement in signal processing and its applications increases scientific development and get human satisfaction and comfortability **(Mahmood et al., 2013).** Development in speech processing algorithms is essential because speech expresses thoughts and feelings by articulate sounds. Two principal criteria are used to measure the goodness of speech signals: quality and intelligibility **(Kolbæk et al., 2016; Jerjees, 2023)**. The quality of a speech signal deals with its clarity, nature of distortion, and amount of background noise, while intelligibility deals with the percentage of words that can be clearly understood. During capture, storage, transmission, processing, or conversion, the speech signal may be exposed to external noise called additive noise, such as White noise, Pink noise, Buccaneer noise, Babble noise, and Gaussian noise **(Elert, 2016; Jerjees, 2023)**. SEA enhances the signal corrupted by background noise and thus improves the perception of humans **(Win and Khine, 2019; Jerjees, 2023)**. Speech signals are random signals; therefore, they can be

represented by different models. Gaussian and Laplacian probability density functions are the most well-known functions used for this purpose. SEA is performed based on other speech models to serve better by reducing residual noise and speech distortion levels as much as possible **(Soon et al., 1998; Hasan and Hasan, 2010; Mahmmod et al., 2021)**. Different approaches for enhancing speech signals have been proposed. They include the spectral subtraction method, Wiener and Kalman filtering **(Upadhyay and Jaiswal, 2016; Nabi et al., 2016)**, MMSE estimation **(Scalart et al., 1996; Shi et al., 2023)**, Comb filter **(Zhang et al., 2023)**, subspace methods **(Ghorpade et al., 2023)**, and phase spectrum compensation **(Aghajan et al., 2009)**. Transform domain is very important in the speech enhancement process due to: data compression property, effective data processing, and feature extraction **(Abou-Loukh and Abdul-Razzaq, 2023; Kotz et al., 2001; Weisstein, 2002; Abood, 2023)**. There are different types of transforms have been used to transform the speech signal from the time domain to the transform domain, such as Discrete Krawtchouk Transform **(Feinsilver and Kocik, 2005)**, Discrete Squared Krawtchouk-Tchebichef Transform (SKTT) **(Abdulhussain et al., 2019)**, Discrete Tchebichef Transform (DTT) **(Abdulhussain et al., 2022)**, Discrete Krawtchouk-Tchebichef Transform (DKTT) **(Mahmmod et al., 2018)**, and Discrete Hahn Transform **(Mahmmod et al., 2022)**.
The main aim of the proposed work is to find a new estimator that can reduce different types of background noise without effects the quality and intelligibility of speech signals in different situations of life environments. Therefore, the focus is on proposing dual stages of linear filters based on different types of actual transforms to get the best trade-off between noise reduction and speech distortion problems. Two types of statistical modeling of speech and noise signals are used in the dual stages to improve the intelligibility and quality of speech signals. Notably, most of the existing SEA have failed to handle the conflict between noise reduction and speech distortion, including musical noise.
Therefore, in this work, two linear estimators are connected in cascade based on new actual transforms and super-Gaussian modeling for speech and noise transform components to develop the superiority of the proposed algorithm and tackle the tradeoffs by minimizing the background noise as much as possible that are often existing in the enhanced signal. Moreover, the literature review found that nearly most orthodox SEA used an attenuating filter for noise suppression. Nevertheless, in real-life situations, when both signal and noise exist in the same transform coefficient, the observed noisy coefficient magnitude may not always be greater than the clean signal, even though the noise was originally additive. This issue is not thoroughly investigated and is ignored by most methods leading to a lack of fairness. Therefore, the trend of this research has deliberately considered the constructive and destructive interference between the signal and noise to settle this problem in the cascade linear estimators.

2. THE PROPOSED SPEECH ENHANCEMENT ALGORITHM

The proposed speech enhancement algorithm is based on dual stages of estimators that consist of two low distortion wiener filters (LDWF) connected in series. The first estimator is termed (LDWF1) which estimates the first enhanced speech signal. Then, the output of the first stage will be the input to the second estimator (LDWF2). LDWF2 enhances the properties of the first enhanced signal from LDWF1 by reducing the distortion in quality and intelligibility and providing a perspicuity-enhanced signal. Two types of speech models have been implemented, Laplacian PDF and Gaussian PDF, to get the best one through a self-comparison. Two speech models are used to get the best results based on quality and intelligibility measurement. The proposed estimator is designed to enhance the noisy signal and reduce the amount of noise while keeping the quality and intelligibility as high as possible. The validity and success of any speech enhancement process are based on maintaining high speech properties **(Win and Khine, 2019)**. The following sub-sections will give the details and mathematical descriptions of the proposed estimator and the used discrete transforms.

2.1 Discrete Tchebichef Transform (DTT)

In the proposed estimator, two types of transforms have been used. In the first filter (LDWF1), DTT is performed, while in the second filter (LDWF2), DTKT is performed. In this section, the mathematical equations of DTT will be presented. DTT transforms the noisy signal from the time domain to the transform domain. Its advantages reduce computational complexity so that the computation will be faster and more efficient than other existing transforms, such as Discrete Cosine Transform (DCT) **(Mahmmod et al., 2018)**. Besides, DTT has higher energy compaction than DCT, with similar properties to DCT, such as separability, symmetry, and orthogonality. This transform is built from discrete Tchebichef polynomials (DTP), where DTP is a class of hypergeometric orthogonal polynomials. The n^{th} order scaled of DTP is defined as follows **(Abdulhussain et al., 2022)**:

$$\hat{t}n(x) = \frac{tn(x)}{\sqrt{\rho(n,N)}} \tag{1}$$

where $tn(x)$ is the original DTP formula, and it is defined as:

$$t2n(x) = n! \sum_{k=0}^{n} \quad (-1)^{n-k} \begin{pmatrix} N-1 & -k \\ n & -k \end{pmatrix} \binom{n+k}{n} \binom{x}{k} \tag{2}$$

and $\rho(n,N)$ is the squared-norm of $\hat{t}n(x)$ Which can be written **(Xiao et al., 2016)**:

$$\rho(n,N) = \frac{N(N^2-1)(N^2-2^2)\ldots(N^2-n^n)}{2n+1} = (2n)! \binom{N+n}{2n+1} \tag{3}$$

where n represents the order of the polynomial, and N denotes the size of the polynomial.

2.2 Discrete Tchebichef-Krawtchouk Transform (DTKT)

The second transform used in the second filter (LDWF2) is DTKT. A new set of orthogonal functions was proposed based on a combination of Tchebichef and Krawtchouk transforms. This transform has a spatial localization property and provides better signal feature extraction than other transforms **(Mahmmod et al., 2018; Idan et al., 2021)**. However, the localization property of DTKT is presented for $p = 0.5$ and requires a special type of window for signal framing **(Mahmmod et al., 2018)**. DTKT is defined for $1D$ signal $f(x)$ with a length equal to N as:

$$\Phi_n = \sum_{\mathrm{x}=0}^{\mathrm{N}-1} R_n(x;\ p, N-1) f(x) \qquad n = 0,1,\ldots,N-1 \tag{4}$$

where, $R_n(x;\ p, N-1)$ is the n^{th} order Tchebichef-Krawtchouk orthogonal polynomial. It represents the polynomial derived by multiplying two orthogonal polynomials, namely Tchebichef and Krawtchouk polynomials. The mathematical idea is that the polynomials produced by multiplying any two orthogonal polynomials are also orthogonal **(Jassim and Raveendran, 2012)**.

2.3 Laplacian Distribution

The Laplacian PDF is the first model used to represent the speech signal in this work. It is a continuous probability distribution. It is also called the double exponential distribution because it can be thought of as two exponential distributions (with an additional location parameter) spliced back-to-back. The general form of Laplace distribution is defined as follows:

$$f(\,x \mid \mu, b\,) = \frac{1}{2b} e^{\left(-\frac{|x-\mu|}{b}\right)} \tag{5}$$

where μ represents the location parameter, and $b > 0$ represents the scale parameter **(Kumar and Chari, 2019).** This type of distribution is used to represent the random speech signal used in the Wiener filter's formula. The general mathematical derivation of the Wiener filter equation starts from the MSE formula, that is **(Mahmmod et al., 2017)**:

$$e_k = E\{(WY_k - X_k)^2\} \tag{6}$$

where, e_k is the error formula in the index k, W is the gain of the Wiener filter, Y_k is the noisy signal in the k index. X_k is the clean speech signal in the k index.

Two events represent the occurrence of noisy signals: constructive and destructive. Constructive event happens when the speech and noise signals have the same sign. While a destructive event happens when the speech and noise signals have different signs. The derivation set of the Wiener filter based on these two events will provide a pair of conditional MSEs based on the Laplacian model that is given as follows **(Mahmmod et al., 2021)**:

The gain of the Wiener filter for the constructive event is:

$$G_{W+} = \frac{\xi_k + \left(\frac{1}{2}\right)\sqrt{\xi_k}}{\xi_k + 1 + \sqrt{\xi_k}} \tag{7}$$

while the gain for the destructive event is:

$$G_{W-} = \frac{\xi_k - \left(\frac{1}{2}\right)\sqrt{\xi_k}}{\xi_k + 1 - \sqrt{\xi_k}} \tag{8}$$

where ξ_k is the posterior SNR in the k^{th} index.

2.4 Gaussian Distribution

The second type of speech model used to represent speech model is Gaussian PDF. Two types of speech models have been used for the best performance by making a self-comparison. Gaussian distribution is a continuous probability distribution for a real-valued random variable. The general form of Gaussian distribution is:

$$G = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{9}$$

The parameter μ is the mean or expectation of the distribution. While the parameter σ is its standard deviation. The variance of the distribution is σ^2[8]. As mentioned before, there are two events to represent the occurring of noisy signal: constructive and destructive. The derivation set of the Wiener filter based on these two events will provide a pair of conditional MSEs based on the Gaussian model that is given by **(Mahmmod et al., 2019)**:

The gain of the Wiener filter for the constructive event and Gaussian pdf is:

$$Gw_+ = \frac{\xi_k + (2/\pi)\sqrt{\xi_k}}{\xi_k + 1 + (4/\pi)\sqrt{\xi_k}} \tag{10}$$

while the gain for the destructive event and Gaussian pdf is:

$$Gw_- = \frac{\xi_k - \left(\frac{2}{\pi}\right)\sqrt{\xi_k}}{\xi_k + 1 - \left(\frac{4}{\pi}\right)\sqrt{\xi_k}} \tag{11}$$

The gains (Gw_+ *and* Gw_-) are used for constructive and destructive interferences, respectively. The final estimated speech signal can be found as follows:

$$\hat{X}_k^{speech\ signal} = [f_k Gw_+ + (1 - f_k) Gw_-] * Y_k \tag{12}$$

Ideally, the polarity estimator parameter is $f_k = 1$ for constructive events and $f_k = 0$ for the destructive event, which connects the two mutually exclusive events for the enhanced signal.

2.5 Speech Enhancement Algorithm

Two stages of the Wiener filter have been implemented in the proposed SEA: (LDWF1) based DTT and (LDWF2) based DTKT. Two types of distributions have been used to module the speech signal samples (Laplacian PDF and Gaussian PDF). These two types have been evaluated to get the best noise reduction in the noisy signal and get the higher intelligibility and quality of the speech signal. Wiener filter (WF) is also used to estimate the clean signal, which is an optimal linear estimator. It can be considered one of the most fundamental noise reduction approaches**.** It enhances the speech signal corrupted with random noise **(Xia and Bao, 2014).** The goal of the WF is to compute a statistical estimate of an unknown speech signal using a related signal as an input and filter that known signal to produce the estimate as an output **(Mahmmod et al., 2018)**. WF minimizes the mean square error between the estimated random and desired processes and gets less MSE **(Omatu and Seinfeld, 1981; Kotz et al., 2001; Loizou, 2013)**. The steps of implementing the proposed SEA of this research are shown in **Fig. 1**, where the block diagram of the speech enhancement method is presented. The noisy signal is the input to LDWF1, where DTT is used firstly to transform the noisy signal to transform the domain and get a better noise reducing process. Then, the output of LDWF1 is considered the input to the LDWF2 based on the DTKT transform. Then, after these two filters, the final enhanced speech signal is obtained. The use of these transform is based on their powerful properties in energy compaction and localization properties that provide the best performance in the enhancement process. The estimation of the clean signal is performed based on the above-proposed design.

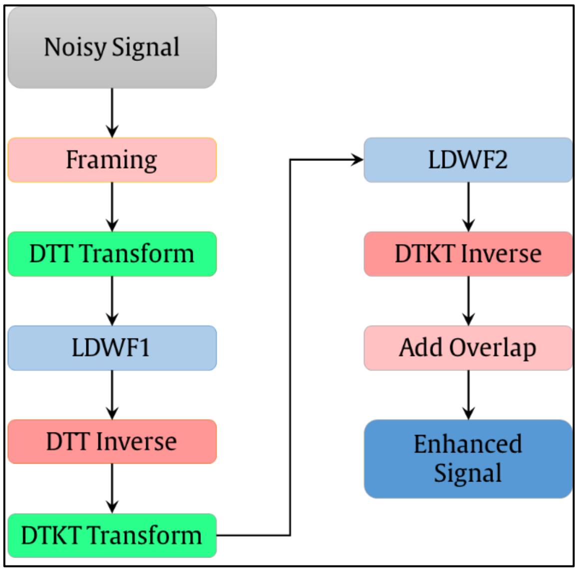

Figure 1. The block diagram of the proposed SEA.

3. RESULTS AND DISCUSSION

In this work, Dual cascade stages of WF are performed using Laplacian pdf and Gaussian PDF to improve the intelligibility and quality of noisy speech degraded by different types of additive noise. The results of speech evaluation measures show the robustness of the proposed SEA based on these other speech models in enhancing the noisy speech signal for different situations where these two distributions have been used individually to get the best one. MATLAB program has been implemented to perform the speech enhancement process based on discrete moments. Speech signals were taken from a well-known database called TIMIT **(Garofolo, 1993)**. This database was designed to be task and speaker-independent and is suitable for general acoustic-phonetic research. It contains broadband recordings of 630 speakers of eight significant dialects of American English, each reading ten phonetically rich sentences. The TIMIT corpus includes a 16-bit, 16kHz speech waveform file for each utterance. Three noises were used in this paper from the NOISEX database **(Varga et al., 1993)**, which consisted of pink, white, and buccaneer.

The sampling rate is 16 kHz. The Hamming function is applied to each frame individually. The windowed speech frame is then transformed to the transform domain using DTT and DTKT with $p = 0.2$ for further processing. The noise is extracted from the noisy speech using WF based on Laplacian PDF or Gaussian PDF. Then, the enhanced frames resulting from inverse transformation were combined using the add-overlap technique to get back and reconstruct the speech signal. Three types of noise were used to corrupt the speech signal: Pink, Buccaneer, and White. The obtained improvement of noisy signals with two stages of Laplacian PDF or Gaussian PDF is shown below in the following Tables. In **Table 1**, the improvement results for Buccaneer noise are obtained using four quality and intelligibility measurements: SNRseg, SNR, PESQ, and OVL measurements.

It can be seen that, for the three sizes, SNRseg, SNR, and PESQ, higher improvement is obtained for the cases of the proposed estimator with equal rates for both types of distributions (Laplacian or Gaussian) at LDWF1 and LDWF2. The best improvement has been obtained for Gaussian distribution for OVL measurement that measures the overall quality of the enhanced signal. Objectively, OVL scores are defined in the composite method based on combining four objective quality measures. It is used for overall quality formed by linearly combining three powerful measures.

Table 1. Results of the four measurements in the Buccaneer noise condition.

Method \ Level (dB)	-5	0	5	10
SNRseg comparison between proposed methods and other existing works.				
Laplace	-4.8247	-2.28	0.5841	4.3073
Gaussian	-4.868	-2.1254	0.6171	4.4137
Laplace + Laplace	-2.8411	-0.7357	2.6772	6.7703
Gaussian + Gaussian	-2.8411	-0.7357	2.6772	6.77031
SNR comparison between proposed methods and other existing works.				
Laplace	1.3053	4.8703	9.0802	13.9154
Gaussian	1.1926	5.0921	9.096	14.0113
Laplace + Laplace	4.3523	6.7756	11.714	16.7876
Gaussian + Gaussian	4.3523	6.7756	11.714	16.7876
PESQ comparison between proposed methods and other existing works.				
Laplace	1.0258	1.042	1.042	1.3071
Gaussian	1.0263	1.0428	1.0428	1.3156
Laplace + Laplace	1.0346	1.0661	1.2201	1.5343
Gaussian + Gaussian	1.0346	1.0661	1.2201	1.5343
OVL comparison between proposed methods and other existing works.				
Laplace	1	1	1.0955	1.5923
Gaussian	1	1	1.1217	1.6278
Laplace + Laplace	1	1	1.3899	1.3899
Gaussian + Gaussian	1	1	1.5343	1.9649

Table 2 presents the results for White noise using SNRseg, SNR, PESQ, and OVL measurements. These three measures can separately measure speech distortion, noise reduction, and enhanced speech quality. For the measurements SNR and PESQ, higher improvement is obtained for the cases of the proposed estimator with equal rates for both distributions (Laplacian or Gaussian) at LDWF1 and LDWF2. But better improvement has been obtained for Gaussian distribution based on SNRseg and OVL measurements that measure the overall quality of the enhanced signal. Therefore, the proposed SEA ensures a good enhancing process without distortion of the reconstructed speech signal using Gaussian models. The results for Pink noise are presented in **Table 3.** The results are calculated using the four performed measurements. They are very successful objective measures, and they are selected carefully because of their higher correlation with subjective testing.

It is noticed that, for the whole four types of measurements, higher improvement is obtained for the cases of the proposed estimators with equal rates for both types of distributions (Laplacian or Gaussian). It is good to mention that the objective measurement PESQ is an international standard measure. Moreover, the following measures are utilized to measure the noise and speech distortion levels in the proposed SEA: mean opinion score of overall speech quality (OVL).

Table 2. Results of the four measurements in the White noise condition.

Method \ Level(dB)	-5	0	5	10
SNRseg comparison between proposed methods and other existing works.				
Laplace	-4.8247	-2.0293	1.1333	4.4777
Gaussian	-4.868	-4.868	1.0236	4.3619
Laplace + Laplace	-2.8411	0.0363	3.1256	6.9227
Gaussian + Gaussian	-2.6842	0.0363	3.1256	6.9227
SNR comparison between proposed methods and other existing works.				
Laplace	1.3053	5.4834	9.8773	14.1665
Gaussian	1.1926	5.2273	9.7034	14.0283
Laplace + Laplace	4.5851	8.4308	12.4704	17.0321
Gaussian + Gaussian	4.5851	8.4308	12.4704	17.0321
PESQ comparison between proposed methods and other existing works.				
Laplace	1.0258	1.0351	1.0828	1.2357
Gaussian	1.0263	1.0354	1.0799	1.2199
Laplace + Laplace	1.0355	1.0662	1.1906	1.447
Gaussian + Gaussian	1.0355	1.0662	1.1906	1.447
OVL comparison between proposed methods and other existing works.				
Laplace	1	1	1.0073	1.4368
Gaussian	1	1	1.0222	1.4305
Laplace + Laplace	1	1	1.2805	1.7621
Gaussian + Gaussian	1	1	1.3899	1.9649

In the distorted signal, residual noise and speech distortion cannot be quantified easily. Therefore, four types of measurements for quality and intelligibility (SNR, SNR seg, OVL, and PESQ) have been used. High improvement was obtained when the speech signal was corrupted by Buccaneer or White noise using two stages of Gaussian PDF. While for Pink noise, it is shown that Laplacian PDF and Gaussian PDF give good results. The improvement of enhancing noisy signal is realized by calculating the average rate improvement. For instance, the average SNRseq improvements for Buccaneer, white, and pink noise are 1.96, 2.12, and 2.03, respectively. Overall, it is evident that the proposed dual-stage estimator improves all measurements over existing SEA methods that use a single-stage Wiener filter with Gaussian or Laplacian distributions.

Table 3. Results of the four measurements in the pink noise condition.

Method \ Level(dB)	-5	0	5	10
SNRseg comparison between proposed methods and other existing works.				
Laplace	-4.8247	-1.8932	1.0723	4.5746
Gaussian	13.88	-2.093	0.8973	4.5902
Laplace + Laplace	-2.8411	0.1185	3.5557	6.8174
Gaussian + Gaussian	-2.8411	0.1185	3.5557	6.8174
SNR comparison between proposed methods and other existing works.				
Laplace	1.3053	5.2792	9.5811	13.913
Gaussian	1.1926	4.9504	9.3088	13.88
Laplace + Laplace	4.3523	7.8037	12.6343	16.2129
Gaussian + Gaussian	4.3523	7.8037	12.6343	16.2129
PESQ comparison between proposed methods and other existing works.				
Laplace	1	1	1.1974	1.6347
Gaussian	1	1	1.2161	1.6838
Laplace + Laplace	1	1.0046	1.5737	2.029
Gaussian + Gaussian	1	1.0046	1.5737	2.029
OVL comparison between proposed methods and other existing works.				
Laplace	1.0258	1.0438	1.1974	1.6347
Gaussian	1.0263	1.0405	1.0405	1.3188
Laplace + Laplace	1.0346	1.0862	1.295	1.5666
Gaussian + Gaussian	1.0346	1.0862	1.295	1.5666

4. CONCLUSIONS

This work addresses the substantial problem of SEA in estimating a speech signal under various background noise conditions. A dual-stage of linear filters has been connected in cascade in this work to get a robustness estimator that can reduce the noise with less speech distortion. The clean speech signal is estimated as the output of the proposed estimator by taking advantage of different speech models and noise modeling based on orthogonal transform coefficient distribution. The outcomes of the proposed estimator demonstrate the effectiveness and the capability to reduce unwanted background noise in terms of different speech quality and intelligibility measures. Based on the comparison results with existing works for different types of noise and noise levels, it can be seen that the corruption has been reduced effectively but with different improvements. The noise components have been canceled effectively using a wiener filter based on Gaussian pdf to provide a clean speech signal with high quality and intelligibility. Therefore, the noise has been reduced with less speech distortion. Other types of distributions can be used in the future to evaluate their effects in improving the quality and intelligibility of speech signals.

NOMENCLATURE

Abbrev.	Description	Abbrev.	Description
DDT	Discrete Tchebichef Transform.	PESQ	Perceptual Evaluation of Speech Quality.
DTKT	Discrete Tchebichef-Krawtchouk Transform.	SNR	Signal-to-Noise Ratio
LDWF	Low Distortion Wiener Filters.	SNR_{seg}	The average of SNR Values over segments ith. speech activity
OVL	Overall Speech Quality		

REFERENCES

Abdulhussain, S.H., Ramli, A.R., Mahmmod, B.M., Saripan, M.I., Al-Haddad, S.A.R., and Jassim, W.A., 2019. A New Hybrid form of Krawtchouk and Tchebichef Polynomials: Design and Application. *Journal of Mathematical Imaging and Vision*, 61(4), pp. 555–570. Doi:10.1007/s10851-018-0863-4.

Abdulhussain, S.H., Mahmmod, B.M., Baker, T., and Al-Jumeily, D., 2022. Fast and accurate computation of high-order Tchebichef polynomials. Concurrency and Computation. *Practice and Experience*, 34(27), P. 7311 Doi:10.1002/cpe.7311.

Abood, Z.I., 2023. Image Compression Using 3-D Two-Level Technique. *Journal of Engineering*, 19(11), pp. 1407–1424. Doi:10.31026/j.eng.2013.11.05.

Abou-Loukh, S.J., and Abdul-Razzaq, S. M., 2023. Isolated Word Speech Recognition Using Mixed Transform. *Journal of Engineering*, 19(10), pp. 1271–1286. Doi:10.31026/j.eng.2013.10.06.

Aghajan, H., Augusto, J.C., and Delgado, R.L.C., 2009. *Human-centric interfaces for ambient intelligence.* Academic Press.

Elert, G., 2016. *The nature of sound--the physics hypertextbook. physics*. info. Retrieved, pp. 6–20.

Feinsilver, P., and Kocik, J., 2005. Krawtchouk Polynomials and Krawtchouk Matrices. In: Recent Advances in Applied Probability. Boston: Kluwer Academic Publishers, pp. 115–141. Doi:10.1007/0-387-23394-6_5.

Garofolo, J.S., 1993. Timit acoustic phonetic continuous speech corpus. Linguistic Data Consortium, 1993. Doi:10.35111/17gk-bn40.

Ghorpade, K., and Khaparde, A., 2023. Single-Channel Speech Enhancement Using Single Dimension Change Accelerated Particle Swarm Optimization for Subspace Partitioning. *Circuits, Systems, and Signal Processing*, pp. 1-19, Doi:10.1007/s00034-023-02324-3.

Hasan, T., and Hasan, M.K., 2010. MMSE estimator for speech enhancement considering the constructive and destructive interference of noise. *IET Signal Processing*, 4(1), pp. 1–11. Doi:10.1049/iet-spr.2008.0114.

Idan, Z.N., Abdulhussain, S.H., Mahmmod, B.M., Al-Utaibi, K.A., Al-Hadad, S.A.R., and Sait, S.M., 2021. Fast shot boundary detection based on separable moments and support vector machine. *IEEE Access*, 9, pp. 106412-106427. Doi:10.1109/ACCESS.2021.3100139.

Jerjees, S. A., Mohammed, H. J., Radeaf, H. S., Mahmmod, B. M., and Abdulhussain, S. H., 2023. Deep Learning-Based Speech Enhancement Algorithm Using Charlier Transform. In 2023 15th

International Conference on Developments in eSystems Engineering (DeSE) pp. 100-105. IEEE. Doi:10.1109/DeSE58274.2023.10099854.

Jassim, W.A., and Raveendran, P., 2012. December. Face Recognition Using Discrete Tchebichef-Krawtchouk Transform. *2012 IEEE International Symposium on Multimedia.* pp. 120–127. Doi:10.1109/ISM.2012.31.

Kolbæk, M., Tan, Z.H., and Jensen, J., 2016. Speech intelligibility potential of general and specialized deep neural network based speech enhancement systems. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 25(1), pp. 153-167. Doi:10.1109/TASLP.2016.2628641.

Kotz, S., Kozubowski, T., and Podgórski, K., 2001. The Laplace distribution and generalizations: a revisit with applications to communications, economics, engineering, and finance. *Springer Science & Business Media*. Doi:10.1007/978-1-4612-0173-1.

Kumar, M.A., and Chari, K.M., 2019. Noise reduction using modified wiener filter in digital hearing aid for speech signal enhancement. *Journal of Intelligent Systems*, 29(1), pp. 1360–1378. Doi:10.1515/jisys-2017-0509.

Loizou, P.C., 2013. Speech enhancement: theory and practice. CRC press, Boca Raton, FL.

Mahmmod, B.M., Ramli, A.R., Abdulhussian, S.H., Al-Haddad, S.A.R., and Jassim, W.A., 2017. Low-Distortion MMSE Speech Enhancement Estimator Based on Laplacian Prior. *IEEE Access*, 5(1), pp. 9866–9881. Doi:10.1109/ACCESS.2017.2699782.

Mahmmod, B.M., Ramli, A.R., Abdulhussian, S.H., Al-Haddad, S.A.R., and Jassim, W.A., 2018. Signal compression and enhancement using a new orthogonal-polynomial-based discrete transform. IET Signal Processing, 12(1), pp. 129–142. Doi:10.1049/iet-spr.2016.0449.

Mahmmod, B.M., Ramli, A.R., Baker, T., Al-Obeidat, F., Abdulhussian, and S.H., Jassim, W.A., 2019. Speech Enhancement Algorithm Based on Super-Gaussian Modeling and Orthogonal Polynomials. *IEEE Access*, 7, pp. 103485–103504. Doi:10.1109/ACCESS.2019.2929864.

Mahmmod, B.M., Abdulhussian, S.H., Naser, M.A., Alsabah, M., and Mustafina, J., 2021. Speech Enhancement Algorithm Based on a Hybrid Estimator. *In: IOP Conference Series: Materials Science and Engineering. Samawah, Iraq*: IOPscience, P. 012102. Doi:10.1088/1757-899X/1090/1/012102.

Mahmmod, B.M., Abdulhussain, S.H., Suk, T., and Hussain, A., 2022. Fast Computation of Hahn Polynomials for High Order Moments. *IEEE Access*, 10, pp. 48719–48732, Doi:10.1109/ACCESS.2022.3170893.

Mahmood, B.M.R., Younis, M.I., and Ali, H. M., 2013. Construction of a General-Purpose Infrastructure for Rfid – Based Applications. *Journal of Engineering*, 19(11), 1425–1441. Doi:10.31026/j.eng.2013.11.06.

Nabi, W., Aloui, N., and Cherif, A., 2016. Speech enhancement in dual-microphone mobile phones using Kalman filter. Applied Acoustics, 109, pp. 1-4. Doi:10.1016/j.apacoust.2016.02.009.

Omatu, S., and Seinfeld, J.H., 1981. Filtering and smoothing for linear discrete-time distributed parameter systems based on Wiener-Hopf theory with application to estimation of air pollution. *IEEE Transactions on Systems, Man, and Cybernetics*, 11(12), pp. 785–801. Doi:10.1109/TSMC.1981.4308618.

Scalart, P., Vieira Filho, J., and Chiquito, J.G., 1996. On speech enhancement algorithms based on MMSE estimation. In 1996 8th European Signal Processing Conference (EUSIPCO 1996) (pp. 1-4). IEEE.

Shi, S., Paliwal, K., and Busch, A., 2023. On DCT-based MMSE estimation of short time spectral amplitude for single-channel speech enhancement. *Applied Acoustics*, 202, P.109134, Doi:10.1016/j.apacoust.2022.109134.

Soon, Y., Koh, S.N., and Yeo, C.K., 1998. Noisy speech enhancement using discrete cosine transform. *Speech communication,* 24(3), pp. 249–257. Doi:10.1016/S0167-6393(98)00019-3.

Upadhyay, N., and Jaiswal, R.K., 2016. Single channel speech enhancement: using Wiener filtering with recursive noise estimation. *Procedia Computer Science*, 84, pp. 22-30, Doi:10.1016/j.procs.2016.04.061.

Varga, A., Steeneken, H.J.M., and others, 1993. Assessment for automatic speech recognition: II. NOISEX-92: A database and an experiment to study the effect of additive noise on speech recognition systems. *Speech communication,* 12(3), pp. 247–251. Doi:10.1016/0167-6393(93)90095-3.

Weisstein, E.W., 2002. Normal distribution. From MathWorld. A Wolfram Web Resource. https://mathworld.wolfram.com/NormalDistribution.htm.

Win, H.P.P., and Khine, P.T.T., 2019. Speech enhancement techniques for noisy speech in real world environments. MERAL Portal.

Xia, B., and Bao, C., 2014. Wiener filtering based speech enhancement with weighted denoising auto-encoder and noise classification. *Speech Communication*, 60, pp. 13–29, Doi:10.1016/j.specom.2014.02.001.

Xiao, B., Lu, G., Zhang, Y., Li, W., and Wang, G., 2016. Lossless image compression based on integer Discrete Tchebichef Transform. *Neurocomputing,* 214, pp. 587–593, Doi:10.1016/j.neucom.2016.06.050.

Zhang, T., Wang, H., Geng, Y., Zhao, X., and Kong, L., 2023. A speech separation algorithm based on the comb-filter effect. *Applied Acoustics*, 203, P.109197. Doi:10.1016/j.apacoust.2022.109197.