Speech Enhancement Algorithm Based on a Hybrid Estimator

Basheera M. Mahmmod¹, Sadiq H. Abdulhussain¹, Marwah A. Naser², Muntadher Alsabah³, Jamila Mustafina⁴

¹Department of Computer Engineering, University of Baghdad, Baghdad, Iraq
²University of Baghdad, Baghdad, Iraq
³Department of Electronic and Electrical Engineering, University of Sheffield, Sheffield, UK
⁴Kazan Federal University, Kazan, Republic of Tatarstan, Russia

*Corresponding author’s e-mail address: basheera.m@coeng.uobaghdad.edu.iq

Abstract. Speech is the essential way to interact between humans or between human and machine. However, it is always contaminated with different types of environment noise. Therefore, speech enhancement algorithms (SEA) have appeared as a significant approach in speech processing filed to suppress background noise and return back the original speech signal. In this paper, a new efficient two-stage SEA with low distortion is proposed based on minimum mean square error sense. The estimation of clean signal is performed by taking the advantages of Laplacian speech and noise modeling based on orthogonal transform (Discrete Krawtchouk-Tchebichef transform) coefficients distribution. The Discrete Krawtchouk-Tchebichef transform (DKTT) has a high energy compaction and provides a high matching between Laplacian density and its coefficients distribution that affects positively on reducing residual noise without sacrificing speech components. Moreover, a cascade combination of hybrid speech estimator is proposed by using two stages filters (non-linear and linear) based on DKTT domain to lessen the residual noise effectively without distorting the speech signal. The linear estimator is considered as a post processing filter that reinforces the suppression of noise by regenerate speech components. To this end, the output results have been compared with existing work in terms of different quality and intelligibility measures. The comparative evaluation confirms the superior achievements of the proposed SEA in various noisy environments. The improvement ratio of the presented algorithm in terms of PESQ measure are 5.8% and 1.8% for white and babble noise environments, respectively. In addition, the improvement ratio of the presented algorithm in terms of OVL measure are 15.7% and 9.8% for white and babble noise environments, respectively.

1. Introduction

Different types of background noise in our environment always interfere with clean speech signal and cause degraded signal for human. In noisy environment, quality and intelligibility attributes of speech signal must be unchanged. These attributes are used to ensure human satisfaction. As a result, a robust speech enhancement method is demanded. Moreover, a noise reduction that has minimum levels of distortion and residual noise is required in the recent SEAs. Nowadays, SEA takes an important role in different applications such as teleconferencing systems, military applications, and speech recognition. SEA enhances the noisy signal that is corrupted by background noise and thus improves the perception of human [1]. The goals of enhancing the noisy signal is to regain the clean speech signal that is corrupted by several types of background noise, improving the perceptual acoustic attributes, reducing listener fatigue, and solving the problem of noise pollution [2,3].

Over the recent years, different approaches and tools for enhancing speech signal have been proposed and developed. These methods/algorithms, can be generally classified into four classes [4]; spectral subtraction algorithms [5], Wiener filtering [6,7], subspace approaches [8], and statistical model-based algorithms [1]. In addition, other classification of SEAs depends on the domain used to mathematically analyze the noisy speech signal. In this classification, SEAs are classified into time domain-based SEAs
and transform domain-based SEAs [9]. In general, transform domain allows viewing the signals (speech, image, and video) in different domains and makes the processing and feature extraction relatively easier [10–12]. Many researchers and developers prefer working in transform domain because the energy of noise can be differentiated and isolated from the speech signal energy easily without much effort [13]. Various transforms techniques have been utilized to transform the noisy signal to uncorrelated domain, for example, Discrete Cosine Transform (DCT) [13], Discrete Tehebichef transform (DTT), Discrete Krawtchouk Transform (DKT) [14], Discrete Fourier Transform (DFT) [15,16], and Discrete Krawtchouk-Tchebichef transform (DKTT) [6]. Transform domain is very effective in the enhancement process, due to data compression property which is an ultimate goal in SEA, effective data processing, and feature extraction. Moreover, the filtering process in transform domain has less computational complexity. Real transforms are known widely for their good spectral resolution and energy compaction which positively influence the process of noise suppression [13,17]. The computation of real transform coefficients is easier to be calculated as there is no need to perform phase correction. Therefore, less severe consequences will be expected [13,18–20].

To represent speech and noise coefficients, different types of probability density functions (PDF) are used, for example, Gaussian, Laplacian and Gamma functions. Gaussian prior is used widely to model the distribution of speech and noise transform in the uncorrelated domain based on central limit theorem [5]. On the other hand, super-Gaussian functions have been adopted by many techniques for modeling speech and noisy signals [16,21,22]. The adoption of super Gaussian functions is adequate for signal representation because they have spikier peaks and longer tails. Besides, the Gaussian PDF is valid when frames duration is longer than the signal span correlation [18,23]. Generally, Gaussian prior is utilized to model noise, whilst the speech signal is assumed to be Gamma or Laplacian [1].

Super-Gaussian have been assumed by many SEAs such as [16,21] and an improvement has been obtained; however, the computational complexity is high and the residual noise has been lessen on the account of speech distortion. In addition, attenuation filter has been used by most of SEAs, where this fact is not permanently valid because two types of interference exist between the signal and noise (the constructive and destructive events) as suggested in [17,24]. Therefore, if the amplitude of noisy signal is increased, an attenuation filter is desirable; and if the amplitude is lessened, an amplifying filter is required [24]. Motivated by these points and as an extension to our work in [1], this paper introduces a new SEA using orthogonal transform in low distortion approach based on a hybrid cascade combination of linear bilateral Laplacian estimator (LBLE) and non-linear bilateral Laplacian estimator (NBLE). The proposed SEA aims to reduce speech distortion, minimize the residual noise, and gain an accurate enhanced speech signal.

This paper is organized as follows: Section two gives the basic theory and algorithms of the orthogonal transform (DKTT), LBLE and NBLE estimators, and then the proposed SEA is described. In Section three, the evaluation comparison with existing algorithms is performed to confirm the proposed SEA performance. In section four, the conclusion of this work is presented.

2. The Proposed Speech Enhancement Algorithm

The proposed estimator is combined using two estimators; linear and non-linear. The first estimator is the non-linear estimator termed as NBLE and the second estimator is the linear estimator termed as LBLE. In the proposed estimator, the first stage is based on the non-linear modeling technique, NBLE, which gives a good performance in SE process. While the second stage is the linear modeling technique, LBLE, that supports the work of the first stage. In other words, NBLE estimates the clean signal effectively based on its non-linear characteristics while, LBLE works as post processing filter that regenerates the distorted components of speech signals coming out from NBLE output. The proposed estimator is designed to reduce the output noise level as much as possible of the noisy signal without providing a noticeable distortion on the enhanced speech signal. The proposed estimator depends on stochastic models then the process of estimation clean signal was derived via optimizing MSE
mathematically. The system is worked in the uncorrelated domain of DKTT to provide an effective speech extraction with residual noise and less speech distortion. The mathematical definitions of DKTT, LBLE, and NBLE will be presented in the following sections.

2.1. Discrete Krawtchouk-Tchebichef Transform (DKTT)
DKTT is used specifically in the proposed work to transform the noisy signal into uncorrelated domain because of energy compaction property. With this property, better performance of SEA is obtained. Moreover, DKTT is a real transform, which means it will handle the problem of phase that exists in complex transform. Note that, the phase of DKTT is a binary phase and depends completely on the sign of its coefficients. Therefore, the mathematical computational complexity of the noisy signal analysis and clean signal synthesis are reduced. It is important to mention that the coefficients distribution of DKTT provides better fitting with the Laplacian distribution. DKTT [6] was constituted of two orthogonal polynomials which are Tchebichef polynomial [25] and Krawtchouk polynomial [26].

The mathematical expression of the \( n \)th order for the Krawtchouk-Tchebichef functions, \( R_n(x) \), is:

\[
R_n(x; p, N) = \sum_{i=0}^{N-1} K_i(n; p, N - 1)t_i(x) \\
n, x = 0, 1, ..., N - 1, N > 0, p \in (0,1)
\]

where \( t_i(x) \) represents the \( i \)th order of the normalized and weighted Tchebichef polynomial [25]. On the other hand, \( k_i(n; p, N) \) represents the weighted Krawtchouk polynomial of the \( i \)th order [26].

The DKTT moments (\( \psi_n \)) of speech signal \( f(x) \) with a frame size of \( N \) is computed by:

\[
\psi_n = \sum_{i=0}^{N-1} R_n(x)f(x) \quad ; n = 0, 1, ..., N - 1
\]

DKTT has been chosen to transform the observed signal to uncorrelated domain based on its good energy compaction and localization properties [6]. Since, energy compaction is the main advantage to improve the performance of noise suppression [27]. DKTT makes speech signal energy distributed into lesser coefficients in the uncorrelated domain. Consequently, DKTT separates the clean signal from the noise more accurately. Moreover, DKTT has an accurate localization property. This property determines the location of the signal frequencies based on the localization parameter, \( p \). To return the original signal back, the inverse of DKTT is carried out as:

\[
f(x) = \sum_{n=0}^{N-1} \psi_n R_n(x) \quad ; n = 0, 1, ..., N - 1
\]

2.2. The Linear Bilateral Laplacian Estimator (LBLE)
Two estimators are used in the proposed SEA based on our previous work [1]. LBLE represents a post processing filter to the enhanced signal that outcomes from NBLE. The derivation of LBLE formula depends basically on the notion of Wiener filter since the linear MMSE estimator gives Wiener estimator [28]. It is noteworthy that, LBLE is working on the principles of low distortion approach. LBLE enhanced speech signal based on two gain functions: constructive and destructive events [28]. The constructive gain is used to attenuate the noisy, and the destructive gain is used to amplify the noisy signal. The two gain functions are derived as shown in the following paragraphs. In additive noise model,
$x(n)$, $y(n)$, and $d(n)$ represent the clean speech, noisy speech, and noise signals, respectively. Where $y(n)$ is given by:

$$y(n) = x(n) + d(n)$$  (4)

Using DKTT, the noisy signal is transformed into DKTT domain to obtain $X_l(k)$, $Y_l(k)$ and $D_l(k)$ as follows:

$$Y_l(k) = X_l(k) + D_l(k)$$  (5)

where $l$ is the frame number and $k$ is the frequency index. The Laplacian distribution of speech signal is defined as follows:

$$p_x(x_k) = \frac{1}{2b_x} e^{-\frac{|x_k|}{b_x}}$$  (6)

While, the Laplacian distribution for the noise signal is defined as:

$$p_d(d_k) = \frac{1}{2b_d} e^{-\frac{|d_k|}{b_d}}$$  (7)

The speech signal and noise signal are independent and they have zero mean. To derive the gain functions of the estimator, the expectation value of speech and noise signals must be calculated. The expectation value of $|X_k|$ is:

$$E[|X_k|] = \int_{0}^{\infty} X_k p_X(|X_k|) dX = b_x$$  (8)

While, the expectation value for the noise signal is:

$$E[|X_k|] = b_d$$  (9)

Then, the mathematical formula of the cross term in constructive event is calculated to be:

$$E[|X_k||D_k|] = b_x b_d$$  (10)

For the constructive event, the following condition $X_l(k)D_l(k) \geq 0$ is satisfied. In destructive event, the mathematical formula of cross term is:

$$E[|X_k||D_k|] = -b_x b_d$$  (11)

And the following condition $X_l(k)D_l(k) \leq 0$ is satisfied. The optimal estimator coefficient of the linear multiplicative gain estimators is computed via minimizing MSE between the clean signal $X_l(k)$ and the estimated signal $\hat{X}_l(k)$. The resulted formulas are [1]:

$$G_{l+} = \frac{\xi_k + 1/2\sqrt{\xi_k}}{\xi_k + 1 + \sqrt{\xi_k}}$$  (12)

$$G_{l-} = \frac{\xi_k - 1/2\sqrt{\xi_k}}{\xi_k + 1 - \sqrt{\xi_k}}$$  (13)

where $G_{l+}$ is the LBLE gain for constructive event and $G_{l-}$ is the LBLE gain for destructive event, $\xi_l(k)$ is a priori Signal to Noise ratio, that is:
\[ \xi_i(k) = \frac{E[X_i^2(k)]}{E[D_i^2(k)]} = \frac{\lambda_{X_i}(k)}{\lambda_{D_i}(k)} \]  

(14)

where \( \lambda_{D_i}(k) \) is the noise variance and \( \lambda_{X_i}(k) \) is the speech signal variance. The resulted LBLE estimator formula is:

\[ \hat{X}_k^{LBLE} = [f_k G_{L+} + (1 - f_k) G_{L-}] Y_k \]  

(15)

where, \( \hat{X}_k^{LBLE} \) represents the estimated speech signal, and \( f_k \) represents the polarity estimator parameter and in this work, it is assumed to be ideal. This parameter determines the type of noise interference (constructive or destructive) therefore \( f_k \) is either zero or one. It is noteworthy to mention that the input of this estimator is the estimated speech signal that comes from the first stage estimator, NBLE. The second stage estimator, LBLE, is a multiplicative filter. It handles the problem of over attenuated that appears when attenuating the necessary components of original speech signal. LBLE regenerates the over-suppressed components of the original signal that are distorted by first stage estimator, NBLE.

2.3. The NBLE Estimator

The Laplacian prior in the MMSE sense is used to represent the distribution of DKTT coefficients for speech and noise signals [1]. The speech signal is estimated according to the noisy signal component. Hence, NBLE requires the speech and noise information [5]. As stated before, there are two events of noise interference; \( E_+ \) (constructive) and \( E_- \) (destructive). Thus, NBLE formula is based on the conditional expectation formulas for these two events, \( E[X_k|Y_k,E_+] \) and \( E[X_k|Y_k,E_-] \), as follows [1]:

\[ \hat{X}_k^{NBLE} = f_k E[X_k|Y_k,E_+] + (1 - f_k) E[X_k|Y_k,E_-] \]  

(16)

These expectation formulas are given in details as follows. For constructive events, NBLE gain equation is:

\[ G_{N_+} = \sqrt{\frac{\xi_k}{2\gamma_k (1 - \sqrt{\xi_k})}} - \frac{e^{\frac{\xi_k}{2\gamma_k}}}{e^{\sqrt{\xi_k}} - e^{\sqrt{2\gamma_k}}} \]  

(17)

Where \( \gamma_k = \frac{\sigma^2}{E[N_k^2]} \) is the posteriori SNR. For destructive event:

\[ G_{N_-} = \frac{e^{\frac{-\sqrt{2\gamma_k}}{\xi_k}}}{e^{\frac{-\sqrt{2\gamma_k}}{\xi_k}} + e^{-\sqrt{2\gamma_k}}} \cdot \sqrt{\frac{\xi_k}{2\gamma_k (1 + \sqrt{\xi_k})}} - \frac{e^{\frac{\sqrt{2\gamma_k}}{\xi_k}}}{e^{\sqrt{\xi_k}} - e^{-\sqrt{2\gamma_k}}} \]  

(18)

Finally, the mathematical formula of NBLE is given by:

\[ \hat{X}_k^{NBLE} = \left(f_k G_{N_+} + (1 - f_k) G_{N_-}\right) Y_k \]  

(19)
2.4. The Hybrid Speech Estimator

In the proposed SEA, two estimators named as NBLE (first stage) and LBLE (second stage) are combined in cascade to gain the estimated speech signal effectively. NBLE and LBLE give good results in estimation of clean signal as compared to other existing algorithms [1]. However, some residual noise and some harmonic distortion appeared in the enhanced signal because of a little error between the assumed Laplacian prior and the real distribution of DCT coefficients. These issues are handled through using DKTT, and the combination filter. In the proposed SEA, NBLE will be connected as a first estimator to get the estimated speech signal ($\hat{X}_{k}^{NBLE}$), then this signal will be more processed by LBLE. NBLE is a non-linear dual gains filter based on low distortion approach to handle the two types of noise interference, constructive and destructive. In constructive event, NBLE will attenuate the noisy components to reverse the additive noise process. While in destructive event, NBLE amplifies the noisy components to reduce the reducing in speech amplitude or reverse in polarity that resulted from noise affect. Moreover, LBLE is linear dual gains filter based on low distortion approach that works in the same procedure of NBLE. LBLE is connected as a post processing filter to regenerate and reprocess the estimated speech signal, ($\hat{X}_{k}^{LBLE}$). Moreover, it further reduces the residual noise, including musical noise, despite the fact that residual noise level revealed from NBLE output is already low. LBLE has the ability to attenuate the attenuated noisy components in constructive case and amplifies the amplified noisy components coming from NBLE output based on its low distortion approach and DKTT domain. DKTT has a high energy compaction and good localization property [6], besides DKTT coefficients distribution has a good matching with Laplacian prior. These properties result in enhancing the performance of noise suppression with minimal speech distortion and residual noise. The proposed SEA is shown in Figure 1. It can be seen that, the final estimated signal ($\hat{X}_{k}^{LBLE}$) is obtained using the combination of NBLE and LBLE estimators.

![Figure 1. The block diagram of the proposed SEA.](image-url)
3. Results and discussion

In this research, the output of NBLE estimator is used as an input of LBLE estimator. Hence, the comparative evaluation results are obtained for each estimator individually with the combination of them to demonstrate the improvement brought by the using of the post processing filter as a harmonic regeneration process. Note that, the results of NBLE and NBLE are implemented in DCT domain based on our previous work [1] which termed as DCT-NBLE and DCT-LBLE, respectively. Then, to show the effect of implemented the proposed SEA in DKTT domain, the combination of these two filters are implemented in DCT and DKTT domains which termed as DCT-Two stages and DKTT-Two stages, respectively. The proposed SEA was tested using ten speech files. The utterances are taken randomly from TIMIT dataset [29] with sampling frequency equal to 16 KHz. The speech signals were corrupted by white noise and babble noise taken from NOISEX92 database [30]. Five levels of input SNR from -10 to 10 dB were performed for performance evaluation in different situations.

For the framing process, the length of frame is taken to be 25 ms. Furthermore, Hamming window is considered with 75% overlap. DCT and DKTT transforms are utilized to transform noisy signal to uncorrelated domain. Then the enhancement process is carried out. To reconstruct the enhanced signal, the inverse of DCT and DKTT are employed individually.
To calculate the estimated $\xi_k$, decision-directed approach [15] is employed as shown below:

$$\hat{\xi}(k) = a \frac{\hat{X}_{i-1}(k)}{\lambda_{D,T-1}(k)} + (1 - a) \max (\hat{\gamma}(k)) - 1. \quad (20)$$

where, $a=0.98$. Furthermore, noise power is calculated from the initial noisy signal frames. Besides, the localization parameters of DKTT have been set to 0.2 for the purpose of ensuring a good enhancing process without distortion the reconstructed speech signal. Various powerful measurements are performed to obtain reliable ratings of speech quality. They are very successful as objective measures and they are selected carefully because of their higher correlation with subjective testing. The segmental signal-to-noise ratio (segSNR) [5] and perceptual evaluation of speech quality (PESQ) [31] were used. These objective measures have a significant correlation with the subjective quality measures [32]. The objective measurement PESQ is an international standard measure. Moreover, to measure the levels of noise distortion and speech distortion in the proposed SEA, the following measures are utilized: mean opinion score of overall speech quality (OVL), signal distortion (SIG), background intrusiveness (BAK), in addition to LogLikelihood Ratio (LLR) [32]. As showing in Figure 2, the results have been reported in terms of SegSNR, PESQ, LLR, SIG, BAK, and OVL. The reported comparison values are the averages over ten input signals. The best values have been obtained by the proposed SEA (DKTT-Two stage) which have produced superior output signals compared to the DCT-NBLE, DCT-LBLE, and DCT-Two stage. On the other hand, Figure 3 presents the case of Babble noise. Babble noise is a highly non-stationary type of noise and the enhancing process in its environment is a challenging process. However, it can be seen that the proposed SEA in most testing measures and different conditions provides the best results. It can be seen that for LLR, PESQ, and OVL, the proposed SEA presents a comparable result with the other estimators for some levels of SNR only, but for the other levels of SNR and generally it provides the best results for most cases as can be seen in Figure 3.
4. Conclusion
In this paper, new estimator based on a combination of LBLG and NBLG has been proposed. The focus of this paper is to reduce the residual noise, including musical noise, without scarifying the speech signal and to regenerate the distorted speech components. Moreover, DKTT transform has a high energy compaction and localization properties that contribute positively to the enhancement performance. In addition, the distribution of DKTT coefficients has an identical matching with Laplacian speech and noise models. This matching makes the suppression of noise perfect against other existing algorithms. Distinct to the other speech enhancement approaches, we minimize the distortion based on the low-distortion approach in different conditions without compromising the process of noise reduction; therefore, LBLE made a good regeneration of speech components. Also, the proposed SEA handles the problem of polarity reversal. Analogous to previous works, the proposed estimator is formulated. However, the proposed SEA is hybrid estimator that consists of two cascade stages based on a high properties orthogonal transform that provide effective enhancement on the noisy signals. The comparative evaluation of the proposed SEA presents its effectiveness and abilities to lessen noise in terms of LLR, SIG, OVL, BAK, PESG, and segSNR measurements. Simulation of different noisy environments makes it evident that the proposed SEA reduces the unwanted noise in a superior way when compared to existing algorithms. For future work, the proposed work will be continued to calculate an optimum estimation for speech signal in different practical cases.

5. References
[1] B. M. Mahmoud, A. R. Ramli, S. H. Abdulhussain, S. A. R. Al-Haddad, and W. A. Jassim 2017 Low-Distortion MMSE Speech Enhancement Estimator Based on Laplacian Prior IEEE Access, 5 1, pp. 9866–9881
[2] J. S. Lim and A. V. Oppenheim, Enhancement and bandwidth compression of noisy speech 1979 Proceedings of the IEEE, 67, 12, pp. 1586–1604
[3] X. Yousheng and H. Jianwen, 2014 Speech enhancement based on combination of wiener filter and subspace filter in Audio, Language and Image Processing (ICALIP) International Conference on. IEEE pp. 459–463.
[4] H. Liu, R. Zhang, Y. Zhou, X. Jing, and T.-K. Truong 2017 Speech denoising using transform domains in the presence of impulsive and gaussian noises IEEE Access 5, pp. 21193–21203
[5] P. C. Loizou, Speech enhancement: theory and practice 2007 CRC press
[6] B. M. Mahmoud, A. R. bin Ramli, S. H. Abdulhussain, S. A. R. Al-Haddad, and W. A. Jassim, 2018 Signal compression and enhancement using a new orthogonal-polynomial-based discrete transform,” IET Signal Processing 12 1, pp. 129–142
[7] B. Xia and C. Bao 2014 Wiener filtering based speech enhancement with weighted denoising auto-encoder and noise classification Speech Communication, 60 pp. 13–29
[8] Y. Ephraim and H. L. Van Trees 1995 A signal subspace approach for speech enhancement IEEE Transactions on speech and audio processing 3 4 pp. 251–266
[9] S. I. Yann 2003 Transform based speech enhancement techniques 2003 Soon Ing Yann, PhD Thesis
[10] S. H. Abdulhussain, S. A. R. Al-Haddad, M. I. Saripan, B. M. Mahmoud, and A. Hussien 2020 Fast Temporal Video Segmentation Based on Krawtchouk-Tchebichef Moments IEEE Access 8, pp. 72347–72359
[11] S. H. Abdulhussain, A. R. Ramli, A. J. Hussain, B. M. Mahmoud, and W. A. Jassim 2019 Orthogonal polynomial embedded image kernel in Proceedings of the International Conference on Information and Communication Technology - ICICT ’19, ser. ICICT ’19. New York, New York, USA: ACM Press pp. 215–221.
[12] A. M. Abdul-Hadi, S. H. Abdulhussain, and B. M. Mahmmod, 2020 On the computational aspects of Charlier polynomials Cogent Engineering 7 1
[13] Y. Soon, S. N. Koh, and C. K. Yeo, 1998 Noisy speech enhancement using discrete cosine transform Speech communication, 24 3, pp. 249–257
[14] W. Jassim, R. Paramesran, and M. Zilany, 2014 Enhancing noisy speech signals using orthogonal moments, ” IET Signal Processing 8 8 pp. 891–905
[15] Y. Ephraim and D. Malah, 1984 Speech enhancement using a minimum-mean square error short-time spectral amplitude estimator,” IEEE Transactions on acoustics, speech, and signal processing 32 6, pp. 1109–1121
[16] B. Chen and P. C. Loizou, 2007 A laplacian-based mmse estimator for speech enhancement Speech communication, vol. 49 2 pp. 134–143
[17] T. Hasan and M. K. Hasan, 2010 MMSE estimator for speech enhancement considering the constructive and destructive interference of noise IET Signal Processing 4 1 pp. 1–11
[18] X. Zou and X. Zhang, 2007 Speech enhancement using an mmse short time dct coefficients estimator with supergaussian speech modeling Journal of Electronics (China) 24 3 pp. 332–337
[19] S. H. Abdulhussain, A. R. Ramli, B. M. Mahmmod, M. I. Saripan, S. A. R. Al-Haddad, and W. A. Jassim, 2007 A New Hybrid form of Krawtchouk and Tchebichef Polynomials: Design and Application,” Journal of Mathematical Imaging and Vision, 61 4 pp. 555–570
[20] H. S. Radeaf, B. M. Mahmmod, S. H. Abdulhussain, and D. Al-Jumaely, 2019 A steganography based on orthogonal moments in Proceedings of the International Conference on Information and Communication Technology - ICICT ’19, ser. New York, New York, USA: ACM Press pp. 147–153.
[21] R. Martin 2002 Speech enhancement using mmse short time spectral estimation with gamma distributed speech priors in Acoustics, Speech, and Signal Processing (ICASSP), 2002 IEEE International Conference on 1. IEEE, 2002, pp. I–253.
[22] G. Zhao, B. Zhou, X. Zhang, and S. Lu-ying, 2012 A new speech enhancement algorithm with generalized gamma speech model in Wireless Communications & Signal Processing (WCSP) International Conference on. IEEE pp. 1–5.
[23] H. R. Abutalebi and M. Rashidinejad, 2015 Speech enhancement based on β-order mmse estimation of short time spectral amplitude and laplacian speech modeling Speech communication 67 pp. 92–101
[24] I. Soon and S. Koh, 2000 Low distortion speech enhancement IEE Proceedings-Vision, Image and Signal Processing, 147 3 pp. 247–253
[25] S. H. Abdulhussain, A. R. Ramli, S. A. R. Al-Haddad, B. M. Mahmmod, and W. A. Jassim, 2017 On Computational Aspects of Tchebichef Polynomials for Higher Polynomial Order IEEE Access 5 1, pp. 2470–2478
[26] B. M. Mahmmod, A. M. Abdul-Hadi, S. H. Abdulhussain, and A. Hussien, 2020 On Computational Aspects of Krawtchouk Polynomials for High Orders Journal of Imaging 6 8 p. 81
[27] H. Ding, Y. Soon, and C. K. Yeo, “A dct-based speech enhancement system with pitch synchronous analysis 2011 IEEE Transactions on audio, speech, and language processing 19 8 pp. 2614–2623
[28] B. M. Mahmmod, A. R. Ramli, T. Baker, F. Al-Obeidat, S. H. Abdulhussain, and W. A. Jassim, 2019 Speech Enhancement Algorithm Based on Super-Gaussian Modeling and Orthogonal Polynomials IEEE Access 7 pp. 103485–103504
[29] J. S. Garofalo, L. F. Lamel, W. M. Fisher, J. G. Fiscus, D. S. Pallet, and N. L. Dahlgren, 1993 The darpa timit acoustic-phonetic continuous speech corpus cdrom, Linguistic Data Consortium
[30] A. Varga and H. J. Steeneken, 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

[31] A. W. Rix, J. G. Beerends, M. P. Hollier, and A. P. Hekstra, 2001 Perceptual evaluation of speech quality (pesq)a new method for speech quality assessment of telephone networks and codecs,” in *Acoustics, Speech, and Signal Processing, Proceedings.(ICASSP’01). 2001 IEEE International Conference on* 2 pp. 749–752.

[32] Y. Hu and P. C. Loizou, 2008 Evaluation of objective quality measures for speech enhancement *IEEE Transactions on audio, speech, and language processing* 16 1 pp. 229–238