

# New Speech Enhancement based on Discrete Orthonormal Stockwell Transform

Safa SAOUD

Physics Department  
Faculty of Science of Tunis  
Tunis, Tunisia

Mohamed BEN NASER

Physics Department  
Faculty of Science of Tunis  
Tunis, Tunisia

Souha BOUSSELMI

Physics Department  
Faculty of Science of Tunis  
Tunis, Tunisia

Adnane CHERIF

Physics Department  
Faculty of Science of Tunis  
Tunis, Tunisia

**Abstract**—S-transform is an effective time-frequency representation which gives simultaneous frequency and time distribution information alike the wavelet transforms (WT). However, the ST redundantly doubles the dimension of the original data set and the Discrete Orthonormal S-Transform (DOST) can decrease the redundancy of S-transform farther. So, this paper aims to propose a new method to remove additive background noise from noisy speech signal using DOST which supplies a multi-resolution analysis (MRA) spatial-frequency representation of image processing and signal analysis. Hence, the performances of the applied speech enhancement technique have been evaluated objectively and subjectively in comparison with respect to many other methods in four background noises at different SNR levels.

**Keywords**—MRA; Stockwell Transform; DOST; DWT; speech enhancement

## I. INTRODUCTION

The distortion of signals by noise is a ubiquitous problem. In fact, the background noise deteriorates the intelligibility and quality of the speech signals resulting in a harsh drop in performance of speech applications such as sound recording, telecommunications and teleconferencing. These applications need noise reduction and recover the clean signal from noisy signal. Speech enhancement is the most important technique in speech signal processing domain. It eliminates noise and ameliorates the quality and intelligibility of speech communication.

Over the last decades, noise suppression from speech signals is a very interesting area of researchers during speech processing.

The literature is enriched by many works which treat several methods for speech enhancement has been developed and investigated such as Discrete Fourier transformer (DFT), Discrete Cosine Transformer (DCT), Karhunen-Loeve transformer (KLT), Wiener filtering [2,3], Spectral Subtraction [1], Wavelet Transform (WT) [5-8, 22-26], etc. All the methods have their advantages and inconveniences. Particularly, although the Spectral Subtraction [1]-[12] provides a tradeoff between speech distortion and residual

noise, it suffers from a musical noise artifact that is perceptually annoying. Also, the Wiener estimator has a moderate computation load, but it offers no mechanism to control tradeoff between speech distortion and residual noise. Thus, the one major problem of wiener filter based methods [2]-[3] is the requirement of obtaining clean speech statistics necessary for their implementation. Among the methods using time-frequency analyses, an approach of reducing different types of noise that corrupt the clean speech is the use of Discrete Wavelet Transform (DWT) [5]-[9], which is a superior alternative to the analyses based on Short Time Fourier Transform (STFT).

Even though the wavelet transform (WT) has dominated signal denoising for years, and become a powerful tool of signal analysis and is widely used in many applications which comprise image processing and signal analysis. However, in wavelet transform, only the scale information is supplied, so the applications using the wavelet transform may be limited when the absolutely-referenced frequency and phase information are required [10].

The Stockwell Transform (ST) proposed by R. G. Stockwell in 1996 [11], is a time-frequency analysis method. The ST improves the time-frequency resolution of Short Time Fourier Transform (STFT), and can be regarded as an extension or special case of wavelet transform (WT) in the multi-resolution analysis domain. The use of S-transform can get more precise relationship between the distribution of frequency and time of the signal. Thus, the Stockwell Transforms [11] is a hybrid of the STFT and the WT. It provides a time-frequency representation of a signal with a frequency-dependent resolution and shows a great promise in various applications. However, the ST redundantly doubles the dimension of the original data set. Due to this redundancy; use of the ST is computationally expensive and even infeasible on some large size data sets.

Thus, to improve its computational efficiency, R. G. Stockwell proposed the Discrete Orthonormal S-Transform (DOST) in 2007 [11] which reduce the redundancy of S-transform further and makes S-transform practical in real life

and much more convenient. The DOST is based on a set of orthonormal basis functions that localize the Fourier spectrum of the signal. It samples the time-frequency representation given by the ST with zero information redundancy and retains the advantageous phase properties of the ST. Despite The DOST is fairly young compared to other transform, it has been demonstrated to be useful in some fields Such as in image compression [10][13][14][15], image restoration[12] , and image texture analysis[16] .It has also been successfully applied in signal analysis to channel instantaneous frequency analysis.

Therefore, in order to preserve useful information in a speech signal and eliminate as much noise as possible, we propose in this paper, a new method for speech enhancement using the Discrete Orthonormal Stockwell Transform (DOST). The main objective of the proposed method is to decrease the speech distortion and increase the speech intelligibility of degraded speech signals and reduce the listener’s fatigue. This method was compared to the Discrete Wavelet Transform and Spectral Subtraction by means of objective and subjective criteria. The obtained results indicate a good performance of the proposed method and show its high potential for speech enhancement. This paper is organized as follows: Section 2 depicts a brief introduction to the Stockwell transform, the DOST and wavelet theory. Section 3 attempts to explain the methodology for the proposed speech enhancement technique using DOST and DWT. In Section 4, we present the objective and subjective performance measurement parameters used for speech enhancement and we discuss the results. Finally, Section 5 concludes this paper with a discussion and the imagination of our future work.

## II. THEORY

### A. Stockwell Transform

The Stockwell Transforms proposed in 1996 [11, 17, 18, 19, 20], gives a full time-frequency decomposition of a signal. The Stockwell transforms (ST) of  $h(t)$  is defined as the Fourier transform (FT) of the product between a Gaussian window function, and  $h(t)$ .

$$s(\tau, f) = \int_{-\infty}^{+\infty} h(t) \frac{|f|}{\sqrt{2\pi}} e^{\frac{(\tau-t)^2 f^2}{2}} e^{-i2\pi f t} dt \quad (1)$$

Where  $f$  is the frequency,  $t$  and  $\tau$  are time variables, the Stockwell transform decomposes a signal into frequency ( $f$ ) and temporal ( $\tau$ ) components. The relation between  $S(\tau, f)$  and the Fourier transform of  $h(t)$  is expressed as:

$$\int_{-\infty}^{+\infty} s(\tau, f) d\tau = H(f) \quad (2)$$

Where  $H(f)$  is the Fourier Transform of  $h(t)$ . Therefore, we can get the original signal by using this relationship between FT and S-transform:

$$h(t) = \int_{-\infty}^{+\infty} \left\{ \int_{-\infty}^{+\infty} s(\tau, f) d\tau \right\} e^{i2\pi f t} df \quad (3)$$

### B. Discrete Orthonormal S-Transform

The DOST is introduced as an orthonormal version of ST.It can be defined as an inner product between a time series  $h[k]$  and the basis function  $d[k]$ . We use  $v$  to specify the center of each frequency band,  $\beta$  represents the bandwidth and  $\tau$

represents the location in time. Using these parameters, the  $k$  th basis vector is expressed as:

$$S_{(v,\beta,\tau)} = \langle d[k]_{[v,\beta,\tau]}, h[k] \rangle = \frac{1}{\sqrt{\beta}} \sum_{k=0}^{N-1} \sum_{f=v-\beta/2}^{v+\beta/2-1} e^{-i2\pi \frac{k}{N} f} e^{i2\pi \frac{\tau}{\beta} f} e^{-i\pi \tau} h[k] \quad (4)$$

For a signal of length  $N$ , the discrete ST generates  $N^2$  coefficients, while the DOST can represent the same signal with only  $N$  coefficients. For that, the DOST is a non-redundant version of the ST. In order to calculate the DOST much faster we introduce the FFT into DOST:

$$S_{(v,\beta,\tau)} = \langle d[k]_{[v,\beta,\tau]}, h[k] \rangle = \frac{1}{\sqrt{\beta}} \sum_{k=0}^{N-1} \sum_{f=v-\beta/2}^{v+\beta/2-1} \exp\left(-i2\pi \frac{k}{N} f\right) \exp\left(i2\pi \frac{\tau}{\beta} f\right) \exp(-i\pi \tau) H(f) \quad (5)$$

Using FT we can write:

$$S_{(v,\beta,\tau)} = \frac{1}{\sqrt{\beta}} \sum_{f=v-\beta/2}^{v+\beta/2-1} \exp(-i\pi \tau) \exp\left(i2\pi \frac{\tau}{\beta} f\right) H(f) \quad (6)$$

### C. Discrete Wavelet Transforms (DWT)

The Discrete Wavelet Transform (DWT) is a powerful tool of signal and image processing that have been successfully used in many scientific fields such as signal processing and image compression. DWT provides sufficient information both for analysis and synthesis and reduce the computation time sufficiently It analyzed the signal at different frequency bands with different resolutions and decompose the signal into a coarse approximation and detail information. The general form of DWT at  $L$ -level expressed in terms of  $L$  detail coefficients  $d_j(k)$ , and the  $L$ th level approximation  $a_L(k)$  coefficients can be written as [9]:

$$f(t) = \sum_{j=1}^L \sum_k d_j(k) \Psi_j(t) + \sum_k a_L(k) \phi_L(t) \quad (7)$$

Where, and  $\Psi_j(t)$  is the mother wavelet and  $\phi_L(t)$  is the scaling function. The approximation and detail at level  $j$  are expressed as:

$$a_{j+1}(k) = \sum_m h_0(m-2k) a_j(m) \quad (8)$$

$$d_{j+1}(k) = \sum_m h_1(M-2k) a_j(m) \quad (9)$$

Where  $h_0(k)$  and  $h_1(k)$  are known as wavelet filters

## III. NEW PROPOSED SPEECH ENHANCEMENT METHOD

In this research work, speech enhancement algorithm based on transformation is performed using three most commonly used steps: applying transformation (DOST or DWT), truncate coefficients (thresholding) and applying inverse transformation (IDOST or IDWT) to reconstruct the denoised signal (Figure 2).

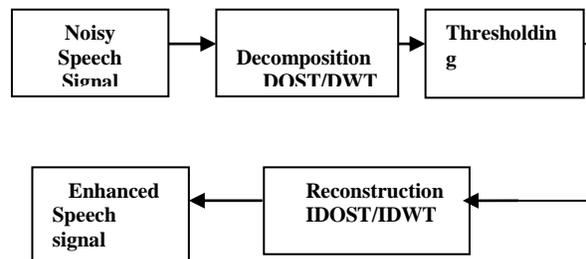


Fig. 1. Block diagram of the suggested method

Identify applicable sponsor/s here. If no sponsors, delete this text box (sponsors).

$$\text{Let: } y(n) = s(n) + w(n) \quad (10)$$

Finally, complete content and organizational editing before formatting. Please take note of the following items when proofreading spelling and grammar

Where  $s(n)$  is the clean speech  $w(n)$  the noise and  $y(n)$  the noisy speech signal. The proposed Discrete Orthonormal S-transform used in speech signal denoising can be represented as follows:

- In this step, the input speech signal is divided into stationary frames and then transformation method (DOST or DWT) is applied of each frame in order to extract coefficients.
- After performing the transformation method of the speech frame, denoising involves truncating the obtained coefficients below a given threshold values. For truncate the small valued coefficients, we calculate the appropriate threshold  $\tau$  to retain the original signal and restrict the noise and let all the coefficients in DOST domain compare with the threshold. In this case, the threshold value ( $\tau$ ) is manually adjusted and is chosen from coefficients ( $0 < \tau < C_{valMax}$ ), where  $C_{valMax}$  is the maximum value of the DOST coefficients or DWT Coefficients. . Soft Thresholding was carried out on the DOST or DWT coefficients before reconstructing the signal. In soft Thresholding, the elements whose absolute values are lower than the threshold are first set to zero. Then the nonzero coefficients are shrunk towards 0.

$$X_{soft} = \begin{cases} \text{sign}(x)(|x| - \tau), & \text{if } |x| > \tau \\ 0, & \text{if } |x| \leq \tau \end{cases}$$

Where  $X$  represents the DOST or DWT coefficients and  $\tau$  is the threshold value.

- After that, we can conduct inverse DOST or inverse DWT to get the denoised speech signal.

#### IV. RESULTS AND EVALUATION

This section presents the experimental results of the proposed speech enhancement method at various SNR levels from -5 to 15 dB. The speech signal taken from the TIMIT Acoustic-Phonetic Continuous Speech Corpus [21], were used to evaluate the proposed algorithm. For this purpose, clean speech signal sampled at 16 kHz is used and recorded by female voice. To illustrate the performance of the proposed enhancement techniques, Several tests in different various noisy conditions, taken from Noisex-92 database: White Gaussian noise, F16 cockpit noise, Volvo car noise and Pink noise with different values of Signal to Noise Ratio (SNR) from -5dB to 15dB were used.

In order to evaluate the denoising performance of the DOST method and to compare it to DWT denoising and Spectral Subtraction; a number of objective tests used for speech enhancement technique evaluation, are presented in this study. Then, the proposed method is subjectively evaluated in terms of Informal listening tests in order to find the analogy between the objective metrics and subjective sound quality.

#### A. Objective evaluation

Objective measures [28] are based on mathematical comparison between the original and processed speech signals. The measure of the signal to noise ratio, SNR is one of the most extensively used. As the name suggests, it is computed as the ratio of the signal to noise powers in decibels:

- Signal-to-Noise Ratio

$$SNR = 10 \cdot \log_{10} \left[ \frac{\sum_n s^2(n)}{\sum_n [s(n) - \hat{s}(n)]^2} \right] \quad (11)$$

where  $s$  and  $\hat{s}$  are respectively the clean and the enhanced speech signals.

-Peak Signal to Noise Ratio (PSNR)

$$PSNR = 10 \log_{10} \frac{NS^2}{\|s - \hat{s}\|^2} \quad (12)$$

Where  $N$  is the length of reconstructed signal,  $S$  is the maximum absolute square value of signal  $s$  and  $\|s - \hat{s}\|^2$  is the energy of the difference between the original and reconstructed signal.

-Normalized Root Mean Square Error (NRMSE)

$$NRMSE = \sqrt{\frac{\sum_n (s(n) - \hat{s}(n))^2}{\sum_n (s(n) - \mu_s(n))^2}} \quad (13)$$

Here,  $s(n)$  is the speech signal,  $\hat{s}(n)$  is reconstructed speech signal and  $\mu_s(n)$  is the mean of speech signal

-Perceptual Evaluation of Speech Quality

PESQ (Perceptual Evaluation of Speech Quality) [27] is an objective quality measure that is approved as the ITU-T recommendation P.862. It is a tool of objective measurement conceived to predict the results of a subjective Mean Opinion Score (MOS) test. Particularly, PESQ was developed to model subjective tests commonly to assess the voice quality by human beings

#### B. Results

Several experiments using the TIMIT database were carried out to evaluate the performance of the proposed method and to compare it to DWT based speech enhancement methods [22], [23], [25] and [26] and Spectral Subtraction [1]. Indeed, in this work, for comparative purposes, the DWT algorithm given in [6], [7], [8] and [9], the used mother wavelet is "db10", five decomposition levels and Soft thresholding was implemented [24]. In this part of the paper, the obtained results from SNR, PSNR NRMSE and PESQ computation is reported These results are obtained by the application of the proposed speech enhancement technique, the Discrete Wavelet transform and Spectral Subtraction on a number of noisy speech signals which are obtained by corrupting the original signals by different types of noise (White, Pink, Volvo and F16) at different values of SNR (-5dB to 15dB). A comparative study between our proposed speech enhancement system using DOST and the DWT denoising proves that the proposed speech enhancement system using DOST outperforms the DWT and the experimental results are shown in tables and figures bellows;

TABLE I. SNR MEASURES OBTAINED FOR NOISY AND ENHANCED SPEECH SIGNAL

Noise type	Enhancement technique	Improved SNR(dB)				
		-5	0	5	10	15
White		-5	0	5	10	15
	DOST	3,70	6,50	11,63	14,1	17,26
	DWT	-1,21	2,85	7,49	12,43	17,25
	Spectral Subtraction	-0,18	4,08	8,92	13,62	18,23
F16		-5	0	5	10	15
	DOST	-1,65	2,8	6,94	10,98	15,63
	DWT	-1,66	2,4	6,75	10,34	15,08
	Spectral Subtraction	-0,07	4,51	9,06	13,49	16,83
Pink		-5	0	5	10	15
	DOST	-1,82	1,64	6,68	11,1	15,93
	DWT	-2,95	1,07	5,5	10,19	15,02
	Spectral Subtraction	0,33	4,54	8,92	12,59	16,15
Volvo		-5	0	5	10	15
	DOST	-2,9	1,23	5,52	10,25	15,9
	DWT	-2,95	1,07	5,5	10,19	15,02
	Spectral Subtraction	1,64	6,43	11,28	16,19	21,05

The proposed speech enhancement method is compared with a speech enhancement using DWT and Spectral Subtraction. SNR is calculated under different SNR inputs and the results are shown in Table I and Fig.2.

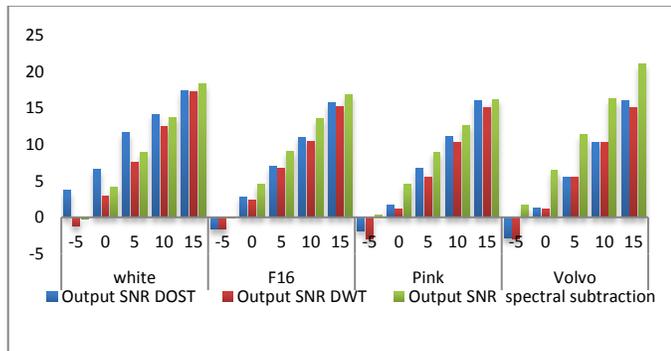


Fig. 2. Comparison of Output SNR for DOST, DWT and Spectral Subtraction

Table 1 and Figure2 below show that the three denoising techniques improve the signal to noise ratio (SNR). The results show also that the DOST based denoising technique is better than the DWT denoising. In fact, the DOST improves the output SNR for SNR input equal to -5dB and 0dB for various noise conditions. It's clear that the Spectral Subtraction denoising is better than the two denoising methods. In case of White Noise, the DOST denoising method is more efficient, however for Volvo noises; the Spectral Subtraction seems to be

more suitable and reliable. From these figures, it can be seen that the DOST method can greatly improve SNR with less distortion of the original speech signal.

TABLE II. PSNR MEASURES FOR ENHANCED SPEECH SIGNAL

Noise Type	Enhancement Technique	PSNR (dB)				
		-5	0	5	10	15
White		-5	0	5	10	15
	DOST	21,0455	23,84	27,97	31,44	34,5625
	DWT	16,1232	20,1927	24,8382	29,7722	33,5986
	Spectral Subtraction	17,1558	21,4244	26,9263	30,9617	34,5760
F16		-5	0	5	10	15
	DOST	16,4876	20,6554	24,0765	29,1221	34,9753
	DWT	15,2148	19,4083	23,1759	28,2993	33,4282
	Spectral subtraction	17,2635	21,8602	26,4096	30,8378	35,1769
Pink		-5	0	5	10	15
	DOST	15,4752	19,9254	24,4534	29,4447	32,2750
	DWT	15,5015	18,0198	23,0204	27,4171	32,2530
	Spectral subtraction	17,6733	21,8828	26,2621	30,9390	35,1534
Volvo		-5	0	5	10	15
	DOST	15,4371	19,4652	24,6547	28,6758	33,2426
	DWT	14,3875	18,4212	22,8447	32,2426	32,2426
	Spectral subtraction	18,9856	23,7729	28,8233	33,5331	38,3963

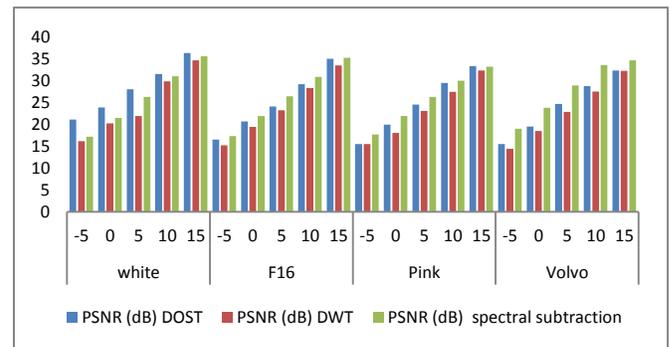


Fig. 3. Evaluating the proposed system based on the PSNR measure

It is observed as the level of Gaussian noise increased from -5dB to 15dB, the output Peak Signal to Noise Ratio (PSNR) values of our speech enhancement system based on DOST outperform the DWT and spectral subtraction. This demonstrates a significant improvement in signal quality and the powerful of the DOST.

However, in case of Volvo noise, the spectral subtraction has the best PSNR.

• Case White Noise

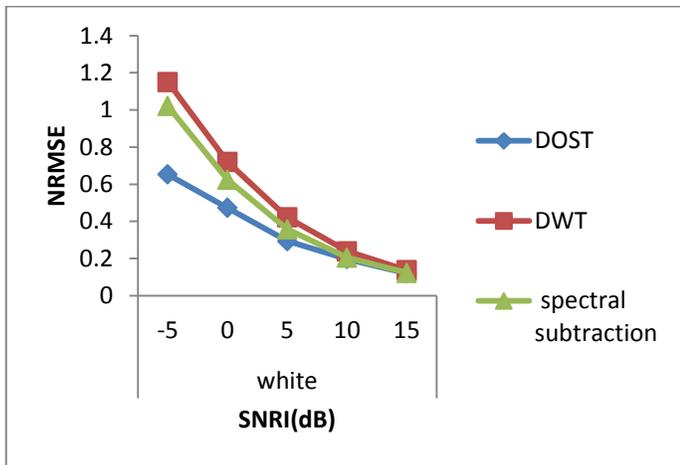


Fig. 4. Comparison of NRMSE for DOST, DWT and Spectral Subtraction for White noise

• Case F16 Noise

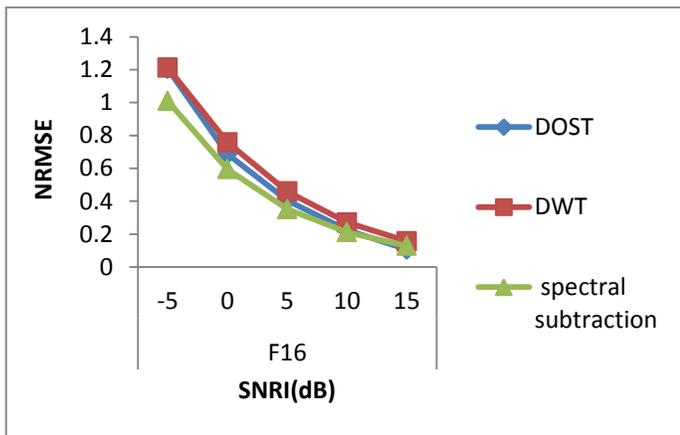


Fig. 5. Comparison of NRMSE for DOST, DWT and spectral subtraction

• Case Pink Noise

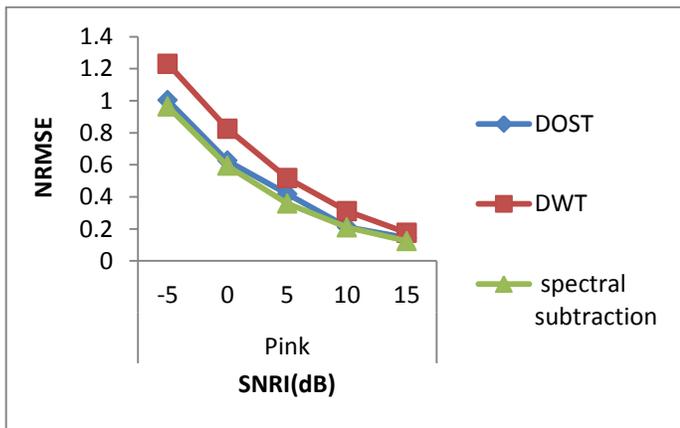


Fig. 6. Comparison of NRMSE for DOST, DWT and Spectral Subtraction for Pink noise

• Case Volvo Noise

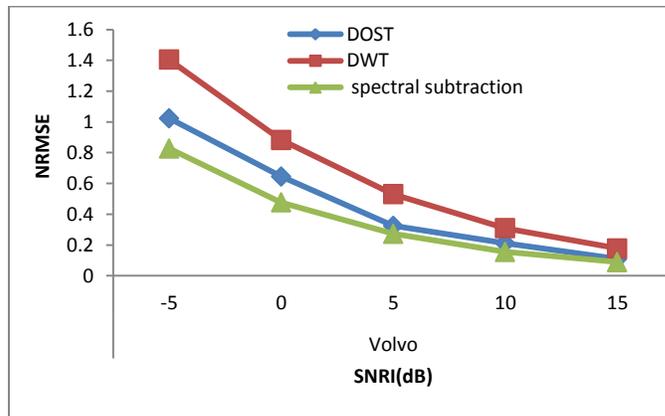


Fig. 7. Comparison of NRMSE for DOST, DWT and Spectral Subtraction for Volvo noise

Figure 4, 5, 6 and 7 below show that NRMSE is less for DOST compared to DWT methods for different types of noise for SNR levels ranging from -5dB, 0dB, 5dB, 10dB and 15dB. In the case of white noise the DOST outperform the DWT and Spectral Subtraction while for Volvo noise the Spectral subtraction has the best performance.

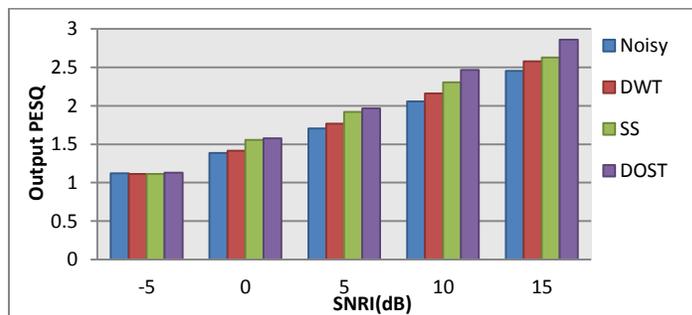


Fig. 8. Performance comparison of PESQ scores for different methods in the presence of White Noise

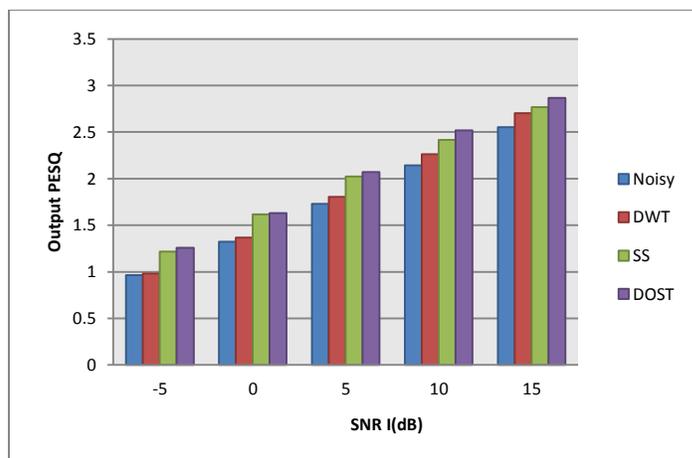


Fig. 9. Performance comparison of PESQ scores for different methods in the presence of Pink Noise

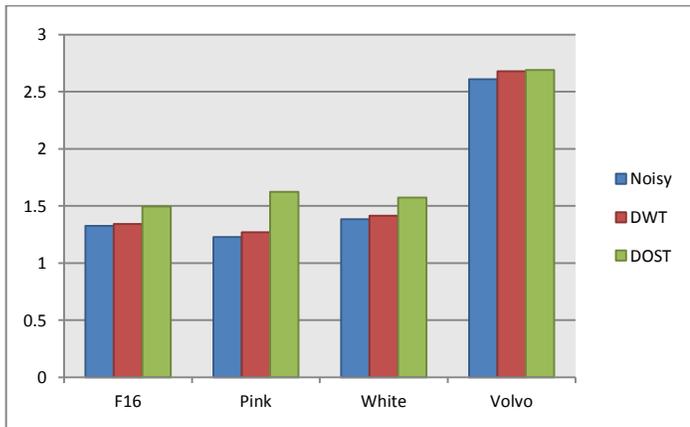


Fig. 10. Performance comparison of PESQ scores for different methods at 5 dB SNR in the presence of various noises

Fig.8 shows the PESQ scores for the noisy signal and the enhanced mentioned signals when the speech is degraded by white noise. As clearly shown in the figure below, PESQ scores of the proposed algorithm are better than PESQ scores given by the algorithm based on the Spectral Subtraction and DWT.

Fig. 9 presents the PESQ scores for the noisy signal and the enhanced mentioned signals when the speech is degraded by Pink noise. Clearly, the PESQ scores of the proposed algorithm outperform the PESQ scores of other methods.

Fig.10 shows the PESQ scores for the noisy signal and the enhanced signals using DWT and DOST over all noise conditions (F16, Pink, White, Volvo) at 5dB SNR. It can be seen from this figure that the proposed method is characterized by the highest PESQ scores showing that the enhanced speech by our method has a better perceived quality. Hence, the high PESQ scores perceived the quality of the enhanced speech.

### C. Subjective evaluation

In order to evaluate the enhancement quality of the noisy speech, we have used the Perceptual Evaluation of Speech Quality (PESQ) score which is a mean opinion score, showing high correlations with subjective listening tests. It ranges from 1 to 5. The higher PESQ score shows the higher perceptual quality and the lower speech distortions. Also we have conducted informal listening test where a group of 10 listeners (six women and four men) are permitted and disposed to perceptually evaluate 3 enhanced speech signals from the NOIZEUS database with three background noises (White, F16, and Pink ) at three SNR levels (0 ,5 and 10 dB). The listeners have used the MOS (Mean Opinion Score) method to evaluate the difference between the residual noise characteristics of the enhanced speech (1: Bad, 2: Poor, 3: Fair, 4: Good, 5: Excellent).Fig. 7 bellow shows the statistic results of subjective evaluation for 3 speech enhancement methods:

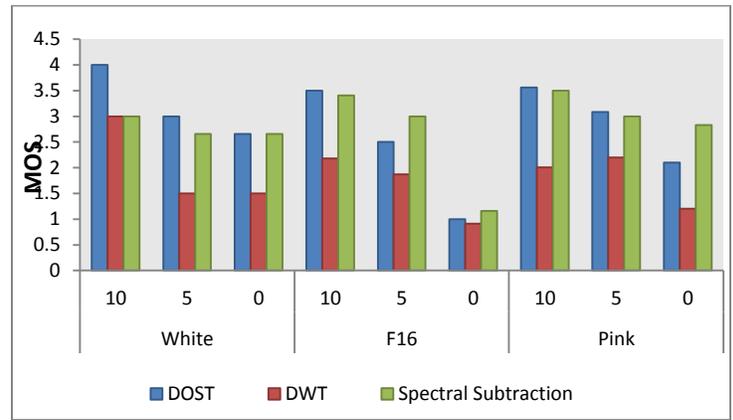


Fig. 11. Subjective evaluation of different speech enhancement methods

The test results show a favorite improvement in the auditory quality of our proposed speech enhancement. In fact, our approach provides a speech signal containing less musical noise while preserving the speech quality.

The listeners found that the subjective sound quality of our proposed method in denoising using DOST has the highest correlation with the objective evaluation in comparison with DWT and spectral subtraction in various noises conditions at different levels of SNR.

## V. CONCLUSION

In this paper, a new method for speech enhancement using the Discrete Orthogonal Stockwel Transform has been presented. The evaluation of the proposed technique is performed by comparing it to the speech enhancement technique based on DWT and the technique based on Spectral Subtraction. Both objective and subjective methods used for evaluation of DOST performance in speech denoising. Hence, this evaluation is based on the use of a number of objective criterions which are the SNR, PSNR, NRMSE and PESQ. Also, in this evaluation, a speech signal with a female speaker from the TIMIT database is used and corrupted it by different types of noises which are Pink, White, F16 and Volvo noises at various input SNR levels ranging from -5dB to 15dB. Simulation results show that the proposed method provides better results in terms of higher output SNR, higher output PSNR, higher PESQ score, and lower NRMSE values than the DWT and Spectral subtraction based denoising methods and results in a better enhanced speech. Also, Informal listening tests justify the efficiency of the proposed method that results in a better enhanced speech than that obtained by the other methods.

In the future research works, we will cooperate the proposed speech enhancement method with speech recognition systems in order to increase their recognition rate under noisy environments.

REFERENCES

- [1] G. Ramesh Babu, "Modified Kalman Filter-based Approach in Comparison with Traditional Speech Enhancement Algorithms from Adverse Noisy Environments," *International Journal on Computer Science and Engineering (IJCSSE)* Vol. 3 No. 2 Feb 2011.
- [2] I. Almajai and B. Milner, "Visually derived wiener filters for speech enhancement," *IEEE Trans. Audio, Speech, and Language Processing*, vol. 19, no. 6, pp. 1642-1651, Aug. 2011.
- [3] S. Ben Jebara, "A perceptual approach to reduce musical noise phenomenon with wiener denoising technique," *IEEE Int. Conf. Acoustics, Speech and Signal Processing, (ICASSP)*, vol. 3, may 2006, p. III.
- [4] S. Tabibian, A. Akbari, and B. NaserSharif, "A new wavelet thresholding method for speech enhancement based on symmetric kullback-leibler divergence," in *Proc. 14th Int. CSI Computer Conference, CSICC*, Oct. 2009, pp. 495-500.
- [5] Y. Hu and P. C. Loizou, "Speech enhancement based on wavelet thresholding the multitaper spectrum," *IEEE Transactions on Speech and Audio Processing*, 12(1):59-67, 2004.
- [6] Slavy G. Mihov, Ratcho M. Ivanov, Angel N. Popov, "Denoising Speech Signals by Wavelet Transform," *Annual Journal Of Electronics; 2009; ISSN 1313-1842*.
- [7] Mahesh S. Chavan, Manjusha N. Chavan, M.S. Gaikwad, "Studies on Implementation of Wavelet for Denoising Speech Signal," *International Journal of Computer Applications; Vol. 3: No.2: 2010; 1-7*
- [8] Matko Saric, Luki Bilicic, Hrvoje Dujmic, "White Noise Reduction of Audio Signal Using Wavelets Transform With Modified Universal Threshold," R. Boskovic B. B Hr 21000 Split, Croatia. 2005
- [9] A. Malayeri, "Noise speech wavelet analyzing an special time ranges," in *Int. Conf. Advanced Communication Technology (ICACT)*, vol. 1, feb. 2010, pp. 525-528.
- [10] Y. Wang and J. Orchard, "On the Use of the Stockwell Transform for Image Compression," *SPIE Electron. Imaging Algorithms System VII*, p. 7245, 2009.
- [11] R. G. Stockwell, "A basis for efficient representation of the S-transform," *Digital Signal Processing*, vol. 17, no. 1, pp. 371-393, January 2007.
- [12] Y. Wang and J. Orchard, "The Discrete Orthonormal Stockwell Transform for Image Restoration," in *16th IEEE International Conference on Image Processing*, 2009, pp. 2761-2764.
- [13] Y. Yan and H. Zhu, "The Generalization of Discrete Stockwell Transforms," *EURASIP*, pp. 1209-1213, 2011.
- [14] H. Huang, F. Sun, P. Babyn, Z. Zhou, and L. Wang, "Medical-Image Denoising and Compressing Using Discrete Orthonormal S transform," in *2nd International Conference on Electrical, computer Engineering and Electronics (ICECEE 2015)*, 2015, vol. 291, no. Icecee, pp. 291-296.
- [15] Y. Wang and J. Orchard, "Fast Discrete Orthonormal Stockwell Transform," *SIAM. J. Science Computation*, 31(5):4000-4012, 2009
- [16] Drabycz, R. G. Stockwell, and J. R. Mitchell, "Image Texture Characterization Using The Discrete Orthonormal S-Transform.," *J. Digit. Imaging*, vol. 22, no. 6, pp. 696-708, Dec. 2009.
- [17] R. G. Stockwell, L. Mansinha, and R. P. Lowe, "Localization of complex spectrum: the S transform," *IEEE Transactions on Signal Processing*, vol. 144, no. 4, pp. 998-1001, April 1996.
- [18] L. Mansinha, R. G. Stockwell, and R. P. Lowe, "Pattern analysis with two-dimensional spectral localization: Application of two-dimensional S transforms," *Physica A*, vol. 239, pp. 286-295, 1997.
- [19] L. Mansinha, R. G. Stockwell, R.P. Lowe, M. Eramian, and R. A. Schincariol, "Local S-spectrum analysis of 1-D and 2-D data," *Phys. Earth Planet. Inter.*, vol. 103, pp. 329C336, November 1997.
- [20] M. Eramian, R. A. Schincariol, R. G. Stockwell, R. P. Lowe, and L. Mansinha, "Review of applications of 1D and 2D S-transforms," *Proc. SPIE*, vol. 3078, pp. 558-568, April 1997.
- [21] J. Garofolo, L. Lamel, W. Fisher, J. Fiscus, D. Pallett, and N. Dahlgren, et al., "TIMIT Acoustic-Phonetic Continuous Speech Corpus: Linguistic Data Consortium, 1993.
- [22] S. Ayat, M.T. Manjuri-Shalmani and R. Dianat, "An Improved Wavelet Based Speech Enhancement by using Speech Signal Features," *Computer and Electrical Engineering*, 23 May 2006, 411-425.
- [23] P. Sunitha, T. Sekhar, "Speech Signal Enhancement Using Wavelet Threshold Methods," *International Journal of Ethics in Engineering & Management Education Website: www.ijee.in (ISSN: 2348-4748, Volume 1, Issue 3, March 2014) 59*
- [24] T. Farah Sanam, C. Shahnaz, "An Approach for Noisy Speech Enhancement through soft thresholding function by Employing the Teager Energy Operation on Wavelet Packet," *International Journal of Scientific & Technology Research* Volume 1, Issue 2, March 2012.
- [25] R. Dhivya, Judith Justin, "A novel speech enhancement technique," *IJRET: International Journal of Research in Engineering and Technology* eISSN: 2319-1163 | pISSN: 2321-7308 Volume: 03 Special Issue: 07 | May-2014, Available @ <http://www.ijret.org> 98
- [26] Kanya Dubey, "A Review on Speech Denoising Using Wavelet Techniques," *International Journal of Engineering Trends and Technology (IJETT) - Volume 4 Issue 9- Sep 2013* ISSN: 2231-5381 [Http://www.ijettjournal.org](http://www.ijettjournal.org) Page 3882
- [27] ITU-T P.862, "Perceptual evaluation of speech quality (PESQ), and objective method for end-to-end speech quality assessment of narrow band telephone networks and speech codecs," *ITU-T Recommendation p.862*, 2000.
- [28] Y. Hu and P. Loizou, "Evaluation of objective quality measures for speech enhancement," *IEEE Trans. Audio, Speech, and Language Processing*, vol. 16, no. 1, pp. 229-238, Jan. 2008.