I J C T A, 9(41) 2016, pp. 687-695 ISSN: 0974-5572

DESIGN OF FILTER USING FRACTIONAL-DCT FOR SPEECH ENHANCEMENT

Rashmirekha Ram and Mihir Narayan Mohanty

Abstract: Enhancement of speech signals corrupted by different types of noise is still a challenging task. The primary target of speech enhancement is to enhance the perceptual parts of the speech, for example, general quality, understandability and level of audience fatigue. The presence of background noise disgraces the nature of the original speech signal while recording or transmitting. Different algorithms have been proposed since several years to enhance the speech quality. But filtering methods are more effective than the others for speech enhancement. In this paper, Fractional Discrete Cosine Transform (FrDCT) filter is proposed to enhance the noisy speech. The fractional order for the filter is selected iteratively to obtain the maximum Signal-to-Noise-Ratio (SNR). Discrete Cosine transform (DCT) filter and FrDCT filter are compared to show the results. Different objective and subjective tests are performed for different types of noise levels. The results show that the Fractional DCT filter performs better than the DCT filter.

Key Words: *Discrete Cosine Transform, Fractional Discrete Cosine Transform, Filtering, Signal-to Noise Ratio, Perceptual Evaluation of Speech Quality, Log Spectral Distance, Speech Enhancement.*

I. INTRODUCTION

Speech enhancement is an essential method to reduce the background noise from the speech signal. Different speech enhancement methods are used for different applications such as mobile communication, teleconferencing systems, hearing aids, speech recognition etc. Speech enhancement algorithms increase the speech quality and reduce the effects of the background noise. From the several years, different algorithms are proposed by researchers for speech enhancement [1].

The noise signal changes very rapidly and the variation of this signal is random in nature, so the utilization of the adaptive algorithms give more converging results [2]. Spectral Subtraction (SS) is a well known single channel speech enhancement method used for reducing the additive noise. An estimated noise spectrum is subtracted from the noisy speech signal to get the clean speech spectrum. The annoying musical noise is the main drawback of this method. To overcome this, Junfeng Li *et.al.* proposed an adaptive spectral subtraction method. The input SNR depends on the variation of the spectral order β [3-4]. Different variations have also been made in spectral subtraction algorithm named as spectral over subtraction, nonlinear spectral subtraction, Multiband spectral subtraction, Iterative spectral subtraction and spectral subtraction based on the perceptual properties. Comparative results of these algorithms are presented in [5].

Both speech and noise signals are nonstationary in nature. So adaptive algorithms are essential in the field of speech enhancement. LMS (Least mean squares), NLMS (Normalized LMS) and RLS (Recursive least squares) algorithms are generally utilized. But in [6], authors have proposed fast affine projection (FAP) and FEDS (Fast Euclidean Search) algorithms for noise cancellation in speech enhancement. The fast

Department of Electronics and Communication Engineering, Siksha 'O' Anusandhan University, Bhubaneswar, Odisha, India
Emails- ram.rashmirekha14@gmail.com, mihir.n.mohanty@gmail.com

convergence rate is the main advantages of these algorithms. An adaptive filtering algorithm is used by Redha Bendoumia and Mohamed Djendi. The variable step sizes are used for noise reduction and it is a modification of the backward structure of the NLMS algorithm. The fast convergence rate is achieved from the proposed algorithm [7].

A second order adaptive Kalman filter is proposed in [8]. More matrix operations are calculated to reduce the convergence rate and SNR. The background noise is reduced more effectively than the conventional Kalman filtering. An adaptive beamformer using the second order extended H_∞ filter is proposed for speech enhancement. This method is derived from the second order Kalman filter [9]. Spectral subtraction method is compared with the adaptive RLS algorithm [10]. Different variations of speech enhancement algorithms are used by different authors. To cancel the impulsive noise, State-space RLS (SSRLS) algorithm is used. The proposed method out forms better enhanced result than the RLS algorithm. Also the short term energy (STE) is used to estimate the noise spectrum adaptively of the noisy speech. The SS method is used to enhance the noisy speech using the estimated noise spectrum of the STE [11-12].

A filter is designed to find the voiced and unvoiced regions of the signal. The empirical mode decomposition and the ACWA (adaptive centre weighted average) filters are combinely used to find the IMFs (Intrinsic mode functions). Both objective and subjective tests are performed to examine the speech quality [13]. To enhance the speech quality, subband filtering and spectral subtraction methods are used. Also the voice activity detector (VAD) is used to find the speech activity. The proposed filter algorithm is suppressed the background noise without changing the speech perceptuality [14].

In Discrete Cosine transform (DCT), the DCT coefficients present in the low frequency regions are the speech and the noise present in the high frequency region. So it is easy to separate the speech signal and the noise signal by choosing a threshold. The DCT coefficients are estimated to obtain the threshold parameter and by using this parameter, the speech presence probability is obtained. The DCT and DFT (Discrete Fourier Transform) are combined to form the discrete fractional Cosine transform (DFrCT). Using the fraction parameter, DFrCT is incorporated with Pitch Synchronous Analysis (PSA). The Pitch Synchronous Overlap and Add (PSOLA) strategy is utilized to improve the performance of PSA [15-16].

A fractional order digital FIR differentiator is implemented by using the Type-III Discrete Cosine Transform. To calculate the filter coefficients, the Fractional calculus of fractional order differentiator is used. Different windows are tested to change the filter response to obtain the better accuracy in FIR filter [17].

Filtering, convolution and multiplexing of signals in fractional Fourier transform domain is discussed in [18]. The relation of FrFT to the chirp transform and the wavelet transform is also presented. FrFT is nothing but the signal is rotated in the time frequency plane in counterclockwise. A. Rutwik *et.al.* proposed the spectral subtraction method in Fractional Fourier domain. The performance of the algorithm is measured for different fractional order. The proposed method finds better MSE and higher SNR as compared to the general spectral subtraction method [19]. The selection of the fractional order is varied according to the different applications and it is an iterative process. The Gammatone filterbank is used in FrFT domain for different values of α. The information of the pitches and the harmonics are used to find the correlogram of the filterbank which helps to find the transform order [20].

The FrFT is the modification of the Fourier transform. The signal is rotated in the time-frequency plane with different fractional order. There are many applications of FrFT in signal processing. In [23], authors presented how the time-frequency distribution is related to the FrFT and filtering using different window functions in fractional domain. FIR filters are also designed in fractional domain with different window functions. The adaptive LMS filter is also designed using FrFT. The fractional based adaptive filters provide better result than the general Fourier transform based filters [24].

This paper is organized as follows-Section I introduced the work along with some related literature. Method for speech enhancement is explained in section II. It follows the result in section III and section IV concludes the work.

II. FRACTIONAL DCT FILTERING FOR SPEECH ENHANCEMENT

Discrete Cosine transform (DCT) is the cosine version of the Discrete Fourier transform (DFT). Fractional Discrete Cosine transform (FrDCT) is the extension of the DFT and DCT. The interpolation method and the window method are generally used for DCT filtering. The DFT matrix is used to calculate the FrDCT matrix [21-22].

DCT for the discrete time sequence *x(0), x (1), x(2),…..x(L-1)* can be defined as

$$x(m) = \sum_{k=o}^{L-1} X(k)\cos\left(\frac{\pi mk}{L-1}\right) \qquad (1)$$

And the Inverse DCT is defined as

$$X(k) = \frac{2}{L-1} c_k \sum_{m=0}^{L-1} c_m x(m)\cos\left(\frac{\pi mk}{L-1}\right) \qquad (2)$$

Where *x(m)* and *X(k)* are formed DCT pairs and

$$c_m = \begin{cases} \frac{1}{2}, m = 0, L-1 \\ 1, otherwise \end{cases} \qquad (3)$$

The frequency response of the fractional DCT filter is defined as

$$G(\omega) = e^{-j\omega(A+\delta)} \qquad (4)$$

Where A is any positive integer and δ is the fractional order of the DCT. The transfer function of the filter is given by

$$H(z) = \sum_{a=0}^{L-1} h(a) z^{-a} \qquad (5)$$

The following steps are followed in fractional DCT filtering.

1. Different noisy speech signals are taken from the NOIZEUS databases.
2. Framing of the noisy signal is taken and Hamming window is applied to the framed speech.
3. For filtering the noisy signals fractional order is chosen arbitrarily between 0.2 to 3.
4. The windowed speech signal is passed through the FrDCT filter.
5. The Inverse FrDCT filter is taken to obtain the enhanced output signal.
6. Steps 3 to 5 is repeated until the maximum SNR is obtained for particular fractional order.

The block diagram of the FrDCT filter for speech enhancement is shown in Figure1.

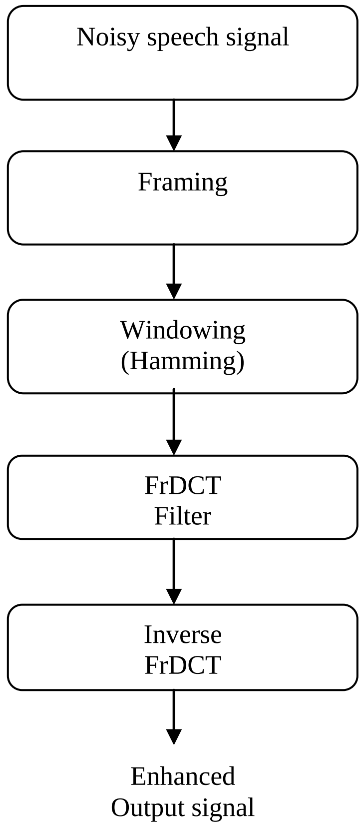

Figure 1. Fractional DCT filter for Speech Enhancement

III. Results and Discussion

To obtain the enhanced signals different NOIZEUS databases are collected. Filtering by using the fractional DCT algorithm, first the speech signals are framed. The frame length of 256 is taken to process the whole signal. Hamming window of length 256 is applied to the framed speech. Figure 2 shows the speech signal after windowing. After filtering using DCT and FrDCT, the filtered speech signals for only one frame are shown in Figure 3 and Figure 4.

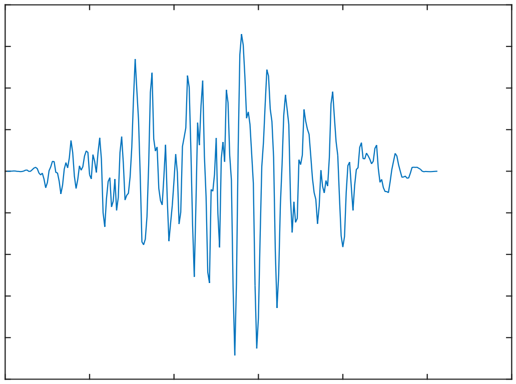

Figure 2. Speech signal after windowing

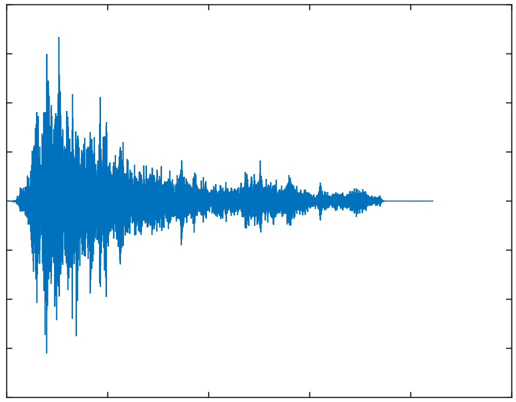

Figure 3. Filtered Speech signal after applying DCT

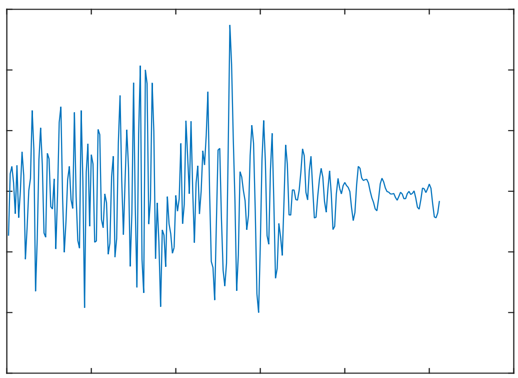

Figure 4. Filtered Speech signal after applying FrDCT

Different noisy speeech signals are tested for speech enhancement using DCT filtering and Fractional DCT filtering. Some results are only shown here. Figure 5 shows the DCT of the speech signal without filtering.The filtered signals using DCT and FrDCT are shown in Figure 6 and Figure 7 respectively. The fractional order is selected iteratively to get the highest SNR.

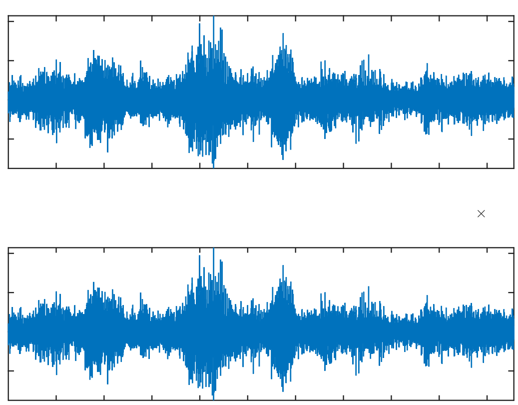

Figure 5. DCT of the noisy speech signal

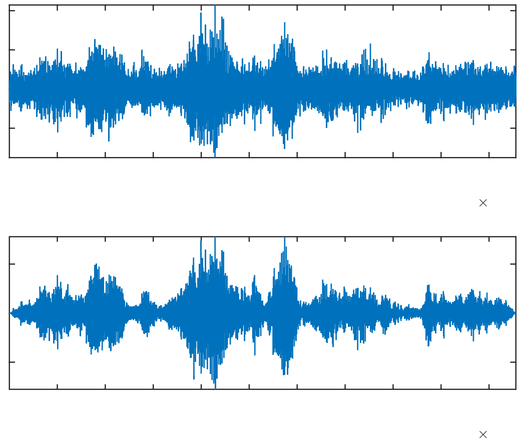

Figure 6. Enhanced speech signal after filtering using DCT

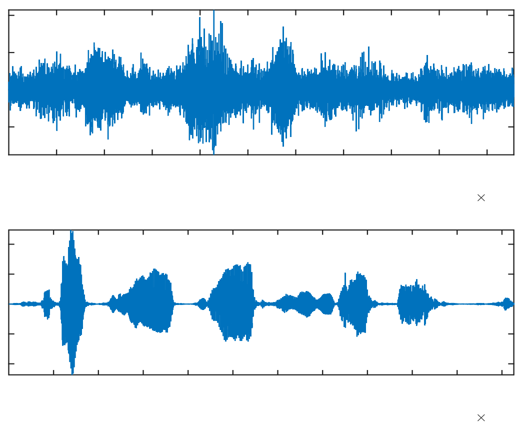

Figure 7. Enhanced speech signal after filtering using FrDCT

The error minimization curve for DCT filtering is shown in Figure 8 . The error converges after 14-15 iterations. But after that the behaviour is nonlinear. In case of FrDCT filtering the error curve converges after 25-30 iterations which is shown in Figure 9. The better error minimization curve is obtained from the FrDCT filtering.

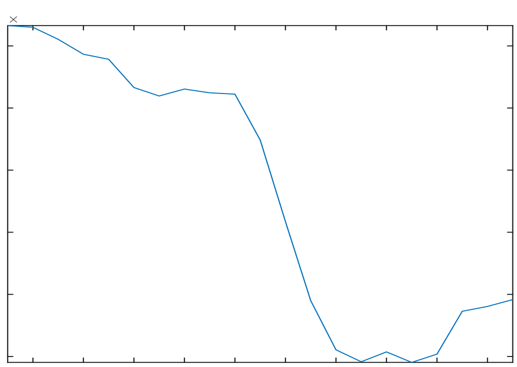

Figure 8. MSE curve using DCT filtering

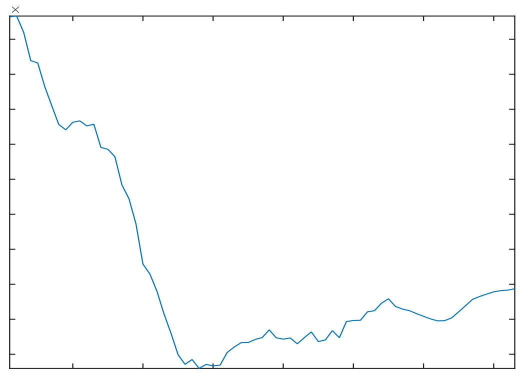

Figure 9. MSE curve using FrDCT filtering

Different objective measures are performed to compare the results. Table-1 shows the Segmental SNR (SegSNR) of different types of noise level. The SegSNR calculates the average SNR of framed signal instead of calculating the whole signal. The maximum SNR 9.9941 dB is obtained for train noise in the proposed method.

Table-1 Segmental SNR Improvement for Different types of Noise Level

Noise	Input Segmental SNR (dB)	Output Segmental SNR (dB)	
		Filtering using DCT	Filtering using FrDCT
Airport	0	6.7845	7.1294
	5	4.2306	5.6943
	10	3.2167	4.3423
Car	0	7.3142	8.6510
	5	6.3498	7.3455
	10	5.2104	6.5618
Train	0	9.4532	9.9941
	5	7.1736	8.6519
	10	6.2827	7.2716

Table-2 PESQ Scores for Different types of Noise Level

Noise	Input Segmental SNR (dB)	PESQ Scores	
		Filtering using DCT	Filtering using FrDCT
Airport	0	2.35	2.90
	5	3.67	3.84
	10	2.90	3.21
Car	0	2.23	2.67
	5	2.98	3.21

	10	**3.10**	**3.32**
Train	**0**	**2.24**	**2.34**
	5	**2.65**	**2.94**
	10	**2.92**	**3.10**

Table-2 shows the PESQ scores of two different filtering methods. The comparison is made for different types of noise levels. PESQ is one of the subjective quality test measurement for speech enhancement. The maximum value is 3.84 which is obtained for airport noise.

Table-3 Log Spectral Distance for Different types of Noise Level

Noise	**Input Segmental SNR (dB)**	**Log Spectral Distance (dB)**	
		Filtering using DCT	**Filtering using FrDCT**
Airport	**0**	**3.6745**	**3.1342**
	5	**3.2938**	**3.0126**
	10	**2.3293**	**1.9920**
Car	**0**	**4.7465**	**4.1274**
	5	**3.2430**	**2.9934**
	10	**4.3565**	**4.1230**
Train	**0**	**3.9783**	**3.2342**
	5	**3.5434**	**3.1279**
	10	**3.5463**	**3.2136**

Table-3 shows the Log spectral distance measure for different types of noise levels. It is nothing but the log spectral distortion and the distance is measured in dB. The LSD in airport noise is 3.2938 dB for DCT filtering. But in FrDCT filtering the value of LSD is 3.0126 dB which is minimum as compared to others.

IV. CONCLUSION

The Fractional DCT filter is proposed for speech enhancement in this paper. DCT filter and FrDCT filter are compared for simulation results. Some noisy speech signals are taken from the databases for different objective and subjective measures. The fractional order is chosen iteratively between 0.2 to 3. The maximum Segmental SNR is obtained for the order 0.5. The FrDCT filter is more effective than the DCT filter as results suggest. For better speech signal enhancement, different fractional transforms may be tested. Also different filtering algorithms can be designed in fractional domain and kept for a future work.

REFERENCES

[1] P. Loizou. Speech Enhancement: Theory and Practice. *CRC Press*, 2007.
[2] S.Haykin, “Adaptive Filter Theory”, *Prentice Hall, Upper Saddle River*, 3rd Edition, 1996.
[3] S. F. Boll, “Suppression of Acoustic Noise in Speech using Spectral Subtraction", *IEEE Transaction ASSP*, pp 113-120, April 1979.
[4] Junfeng Li, Shuichi Sakamoto, Satoshi Hongo, Masato Akagi, Yoiti Suzuki, “Adaptive β-order generalized spectral subtraction for speech enhancement”, Elsevier Signal Processing(J), No. 88, pp 2764– 2776, 2008.
[5] Navneet Upadhyay, Abhijit Karmakar, “Speech Enhancement using Spectral Subtraction-type Algorithms: A Comparison and Simulation Study”, Elsevier, Procedia Computer Science, 54, pp. 574 – 584, 2015.
[6] Sayed. A. Hadei and M. Lotfizad, “A Family of Adaptive Filter Algorithms in Noise Cancellation for Speech Enhancement”, *International Journal of Computer and Electrical Engineering*, Vol. 2, No. 2, pp.1793-8163, April 2010.

[7] Redha Bendoumia and Mohamed Djendi, “Speech Enhancement using Backward Adaptive Filtering Algorithm: Variable Step-Sizes Approaches”, *IEEE Conference*, 2013.

[8] Chanchal Pandey, Sandip Nemade, “Enhancement of the Speech Quality by the Implementation of Second Order Fast Adaptive Kalman Filter Algorithm”, *IEEE India Conference (INDICON)*, 2014.

[9] Jwu-Sheng Hu, Ming-Tang Lee, and Chia-Hsin Yang, “Robust Adaptive Beamformer for Speech Enhancement Using the Second-Order Extended H_∞ Filter”, *IEEE Transactions on Audio, Speech, and Language Processing*, Vol. 21, No. 1, 2013.

[10] Rashmirekha Ram, Mihir Narayan Mohnaty, “Speech Enhancement using RLS Algorithm based Adaptive Filter”, *IEEE Conf. , ICECS*, 2016 (communicated).

[11] Rashmirekha Ram, Sarthak Panda, Hemanta Kumar Palo, Mihir Narayan Mohanty, “Impulsive Noise Cancellation for Speech Enhancement using State Space Adaptive Algorithm”, *IEEE Conf., SCOPES*, 2016,(communicated).

[12] Shambhu Shankar Bharti, Manish Gupta, Suneeta Agarwal, “A New Spectral Subtraction Method for Speech Enhancement using Adaptive Noise Estimation”, *3rd Int. Conf. on Recent Advances in Information Technology*, 2016.

[13] Kais Khaldi, Abdel-Quahab boudraa, Monia Turki, “ Voiced/Unvoiced Speech Classification –based Adaptive Filtering of Decomposed Empirical Modes for Speech Enhancement”, *IET Signal Processing*, Vol.:10 , Issue 1, pp.69-80, 2016.

[14] Vignesh Ganesan and Sangeetha Manoharan, “Surround Noise Cancellation and Speech Enhancement using Sub Band Filtering and Spectral Subtraction”, *Indian Journal of Science and Technology*, Vol 8(33), December 2015.

[15] Li Xue Yao, Xie Hua, Zhang Ru-bo, **“**Speech Enhancement based on Discrete Cosine Transform**”,** *Journal of Harbin Engineeing University*, 2007.

[16] V. R. Balaji , S. Subramanian, “A Discrete Fractional Cosine Transform Based Speech Enhancement System Through Adaptive Kalman Filter Combined with Perceptual Weighting Filter with Pitch Synchronous Analysis”, *International Review on Computers and Software*, Vol.8, No.9, 2013.

[17] Shaily Khare, Manjeet Kumar, Tarun Kumar Rawat, “Window based design of fractional order FIR digital differentiator using type- III discrete cosine transform” *IEEE International Conference on Signal Propagation and Computer Technology (ICSPCT)*, 2014 .

[18] Haldun M. Ozaktas and Billur Barshan, “Convolution, Filtering, and Multiplexing in Fractional Fourier Domains and their Relation to Chirp and Wavelet Transforms”, *J. Opt. Soc. Am*. A/Vol. 11, No. 2, February 1994.

[19] A. Rutwik and B.T. Krishna, “Speech Enhancement using Fractional Order Spectral Subtraction Method”, *International Journal of Computer Technology and Applications*, Vol. 8, No.3, 2015, pp. 1063-1070.

[20] Duo-jia Ma, Xiang Xie, Jing-Ming Kuang, “A Novel Algorithm of Seeking FrFT Order for Speech Enhancement”, *IEEE Conference*, 2010.

[21] Ing Yann Soon, Soo Ngee Koh, Chai Kiat Yeo, “Noisy speech enhancement using discrete cosine transform”, *Speech Communication 24, Elsevier(J)*, pp 249-257,1998.

[22] B.Chitprasert, K.R.Rao, “Discrete Cosine Transform Filtering”, *IEEE Inter. Conf. on Acoustics, Speech, and Signal Processing*, 1990.

[23] Rajiv Saxena, Kulbir Singh, “Fractional Fourier transform: A Novel Tool for Signal Processing”, *J.of Indian Inst. Sci.*, No. 85, pp 11–26, Jan.–Feb. 2005

[24] K.Yashas Kumar, P.V.Muralidhar, D.V.L.N.Sastry, “Performance Analysis of FrFT based Adaptive Filters with LMS Algorithm”, *International Journal of Computer Applications*, Vol.99, No.6, August 2014.