IJCTA, 10(07), 2017, pp. 235-243

Design of Fractional Fourier Transform based Filter for Speech Enhancement

Rashmirekha Ram* and Mihir Narayan Mohanty*

Abstract : Speech enhancement is a challenging issue due to many practical applications. Different methods have been used from several decades. Out of many methods filtering is proven more effective for the enhancement of speech. In this paper, Fractional Fast Fourier Transform (FrFT) based filtering is used instead of Fourier Transform (FrFT) to enhance the speech signal. FrFT is the generalization of FFT where the transform order is a fractional value in time-frequency plane. To analyze minutely and to extract the required information, the fractional order is chosen iteratively to get maximum Signal-to-Noise Ratio (SNR). Various noises with different SNR levels are considered for verification. Two different objectives measures are chosen for comparing the results. The results show that the Fractional FFT filtering provides better result than the FFT filtering. The work has been carried out in the platform of MATLAB environment. The comparative result shows the efficacy of the proposed method.

Keywords : Fractional Fourier Transform, Speech Enhancement, Filtering, Mean Square Error, Signal-to Noise Ratio, Perceptual Evaluation of Speech Quality.

1. INTRODUCTION

Speech signal is attenuated by different types of noise either during the recording or during the processing. The presence of noise reduces the quality of the speech signal. Speech enhancement aims to enhance the sound quality as well as reduce the noise level present in the speech signal. Speech enhancement methods are applied in the field of mobile communication, forensics, teleconferencing systems, speech recognition, hearing aids, etc. Different methods such as filtering based methods, estimation based methods, subspace based methods have been applied for speech enhancement. As the speech signal is both time and frequency variant, the Fourier transform is not suitable because of some limitations. Due to this, Fractional Fourier transform has applied from several decades as discussed in literature.

Fractional Fourier Transform (FrFT) is the generalization of the Fourier transform (FT). Other Fractional transforms are defined from it such as Fractional convolution and correlations, Fractional Hilbert transform linear canonical transform etc. Wigner distribution function (WDF), the ambiguity function (AF) etc. are some of the time-frequency distribution (TFDs). In [1] authors show the effects of the TFDs on the Fractional plane and also the relations between WDF/AF and the fractional/canonical operations are proposed. An algorithm for computing Fractional Fourier transform is proposed by Haldun M. Ozaktas *et.al.* Also the relationship between Discrete Fourier Transform (DFT) and the continuous FRFT is given. To avoid computational complexity of the FrFT, the integral transformations are decomposed into many sub operations [2]. The continuous FrFT is generalized from the FT whereas the discrete FrFT is generalized from the DFT. The DFT matrix is composed by taking the Eigen vectors and Hermite–Gaussian functions. A comparison is made between the continuous FrFT and discrete FrFT. The computational complexity is also reduced in case of discrete FrFT [3]. Duo-jia Ma *et.al.* have proposed an algorithm for selecting

* Department of Electronics and Communication Engineering, ITER, Siksha 'O' AnusandhanUniversity Bhubaneswar, Odisha, India. *E-mail : ram.rashmirekha14@gmail.com, mihir.n.mohanty@gmail.com*

the order of FrFT. The pitch and harmonics information of speech signal are used to obtain the optimum transform order. The adaptive Fractional Mel-frequency cepstral coefficients (MFCC) extraction algorithm based on fractional order is proposed. When the proposed algorithm is compared with the conventional MFCC, the proposed algorithm is found better for different fractional order [4]. The FrFT features are also used for speech recognition [5].

The FrFT is the speculation of the general Fourier transform. But this method depends on the fractional order and can be represented as the rotation in the time-frequency plane. How the time-frequency distributions are related to FrFT is mentioned in [6]. Filtering using window functions and other fractional transforms are presented in this. Fractional spectral subtraction is proposed for enhancing noisy speech [7]. The FrFT is applied to each frame of the noisy signal and then the estimated fractional noise spectrum is subtracted from the fractional noisy speech signal. The inverse FrFT is calculated using the phase of the noisy speech signal to get the enhanced speech signal [8]. In [9], authors are combined the MFCC and FrFT to find an acoustic feature for speech recognition. The FrFT order is adjusted adaptively according to the pitch change rate within a frame.

To enhance the speech signal, Convolution, filtering, and multiplexing of signals in fractional domains are discussed in [10-11]. FrFT is also related to the chirp and Wigner distribution. When the fractional order is set as zero then the FrFT is same as the conventional FT. Spatial filtering is used to separate the noise and the signal. In 2010, authors are used pitches, harmonies and formants of Gammatone Filterbank to determine the transform order. This method is used for speech enhancement and the computational complexity is reduced for searching the optimum transform order [12]. Different orders of FrFT and particle swarm optimization (PSO) algorithm are used to obtain the spectrograms. These spectrograms are converted into low dimensional vector space to reduce the complexity. The speech features are applied to the radial basis function (RBF) neural network for training and testing and finally to identify the speaker [14].

Different adaptive filters have been used for speech enhancement. An improved spectral subtraction filter and spectral attenuation filter is used in [16]. These two filters are combined for enhancing the noisy speech. LMS (Least mean squares), RLS (Recursive least squares) and NLMS (Normalized Least mean squares) adaptive algorithms are some commonly used adaptive algorithms for speech enhancement. Sayed A. Hadei *et.al.* have used FEDS (Fast Euclidean Direction Search) and FAP (Fast Affine Projection) algorithms for noise cancellation and speech enhancement [17]. These two algorithms are the new addition to the adaptive filtering. To minimize the MSE, the combinational adaptive filtering has used by Wu Caiyun in 2012. The WSLMS algorithm is designed to update the filter weight vector [18].

2. METHODOLOGY

Design of Filter using Fractional Fourier Transform

FIR (Finite Impulse Response) filters are widely used for speech filtering. The frequency response of the speech signal is varied according to the require applications. The order of the filter is 25and Hamming window of length 512 is used to eliminate the noise. The impulse response and the frequency response H(ω) of the FIR filter is given as

$$h(m) = h_d(m)w(m) \tag{1}$$

And

$$\mathrm{H}(\omega) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathrm{W}(\theta)\,\mathrm{H}_d(\omega-\theta)\,d\theta \tag{2}$$

Where $w(m)$ is the window function, $h_d(m)$ and $\mathrm{H}_d(\omega)$ is the desired impulse response and desired frequency response respectively. The finite impulse response is convolved with the input signal. This convolution in the frequency domain is performed using FT.

The filter output $y(m)$ can be written as

$$y(m) = \int_{-\infty}^{\infty} h(m-\theta)\,x(\theta)\,d\theta \tag{3}$$

Where $x(m)$ is the input signal and $h(m)$ is the impulse response of the filter.

The above equation can be written in frequency domain as

$$y(m) = \frac{1}{\sqrt{2\pi}} \mathrm{IFT}(\mathrm{FT}(x(m)).\, \mathrm{H}(\omega)) \quad (4)$$

where $\mathrm{H}(\omega) = \mathrm{FT}\,(h(m)$

Then the Fractional filter is defined as

$$\mathrm{Y}(m) = \mathrm{F}^{-\alpha}\{\mathrm{F}^{-\alpha}\,\{x(m)\}.\, \mathrm{Ha}\,(\varepsilon)\} \quad (5)$$

where $\mathrm{H}_{\alpha}(\varepsilon) = \mathrm{F}^{\alpha}\{h(m)\}$

To filter out the noise the Wigner Distribution is rotated and in the fractional domain the noisy signal is filtered. By choosing the proper fractional order (α), the process is repeated. Fig.1 shows the speech enhancement method using filtering in fractional domain.

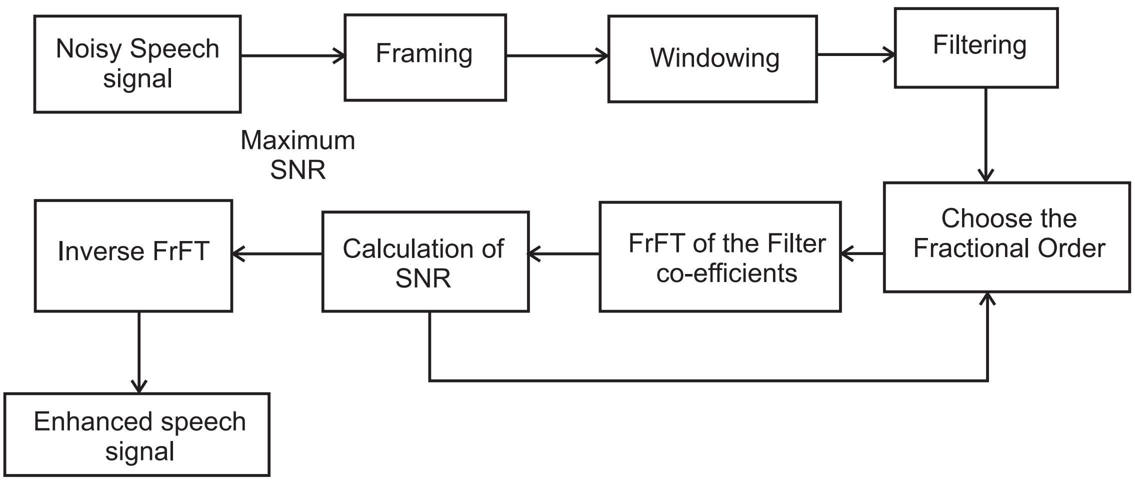

Figure 1: Filtering using Fractional Fourier Transform for Speech Enhancement

3. RESULTS & DISCUSSIONS

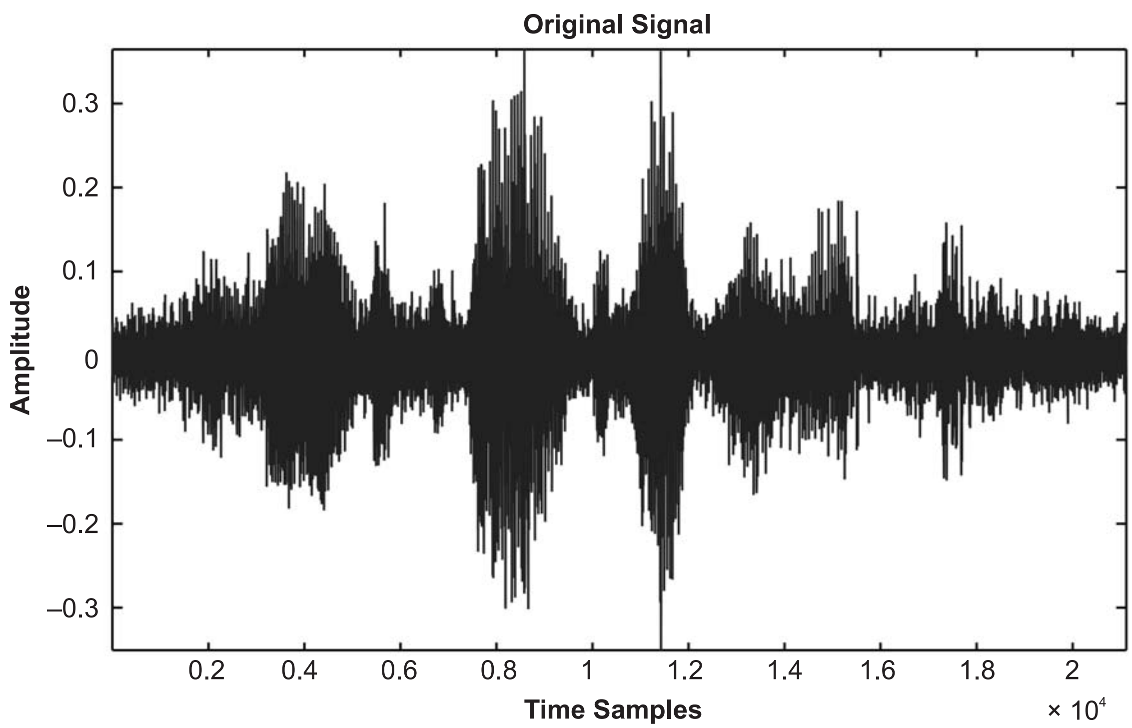

Figure 2: Speech signal for “Nice to meet you”

In order to perform the filtering in Fractional Fourier domain, different speech signals are recorded for testing. The clean speech spoken by a female speaker, "Nice to meet you" is sampled at 8 KHz is taken Fig.2. SNR of different levels are added with the speech signal. As the speech signal is nonstationary in nature, the recorded noisy speech signal is splitted into frames for processing. The speech signal after framing is shown in Fig.3. Then a Hamming window of length 512 is applied to each frame. FFT filtering is applied to the framed noisy speech signal to obtain the enhanced signal as in Fig.6.

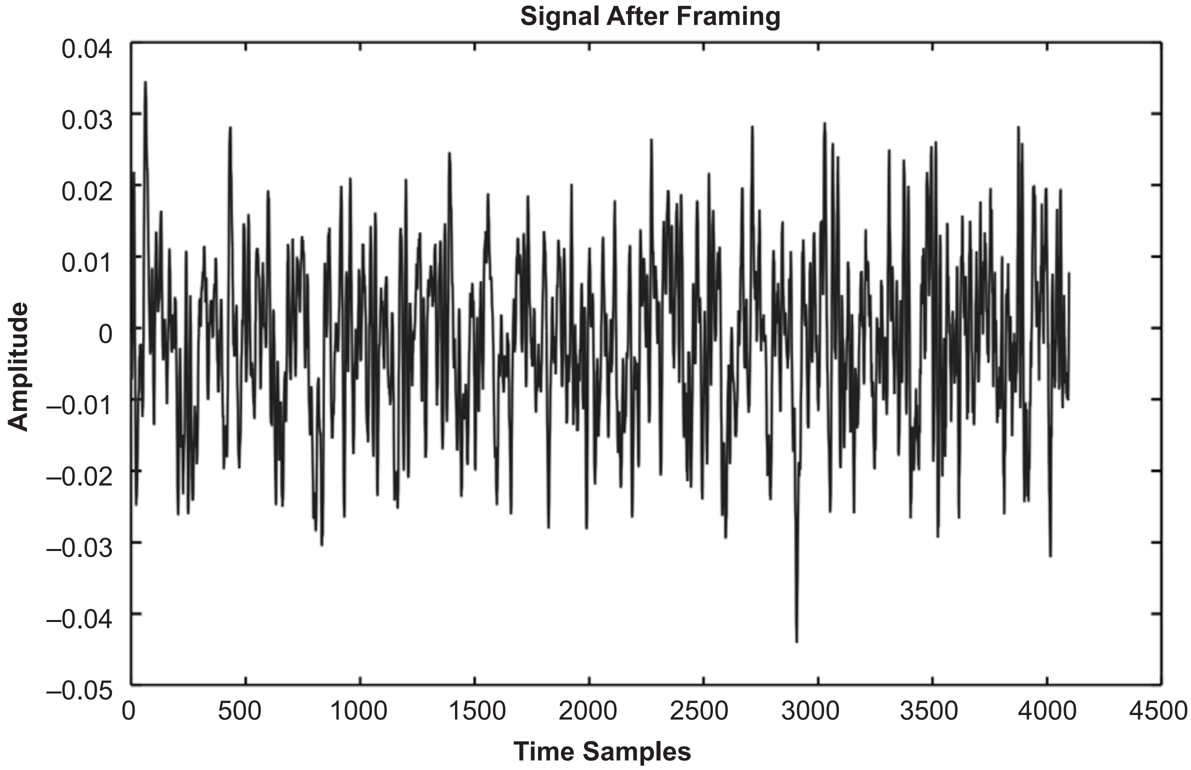

Figure 3: Speech signal after framing

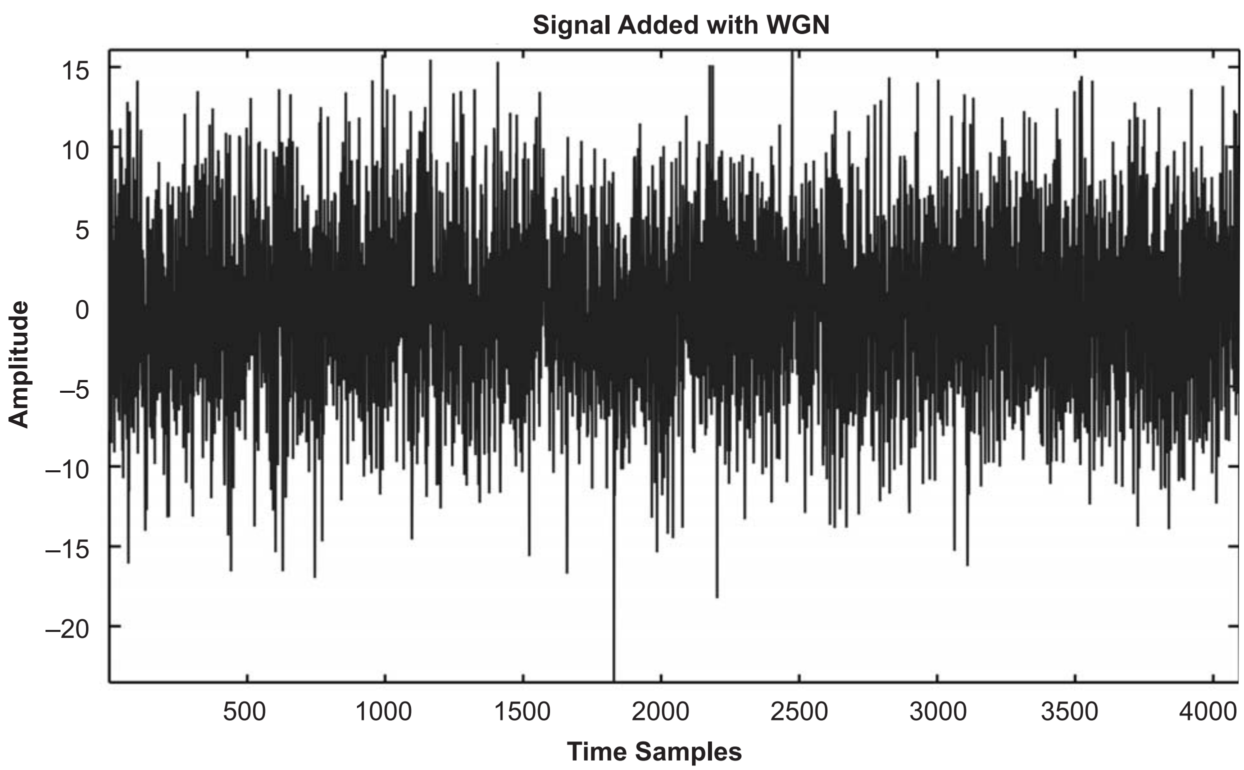

Figure 4: Speech signal added with WGN

Fig.4 shows the speech signal added with the WGN of SNR 5dB. To perform the filtering, the filter is taken in Fractional domain (Fig. 4). Different Fractional orders are tested to obtain highest SNR. The order is varied in between -2 to 2. The maximum SNR is obtained for $\alpha = 1.5$. The inverse FrFT is taken to obtain the enhanced output signal as shown in Fig.7.

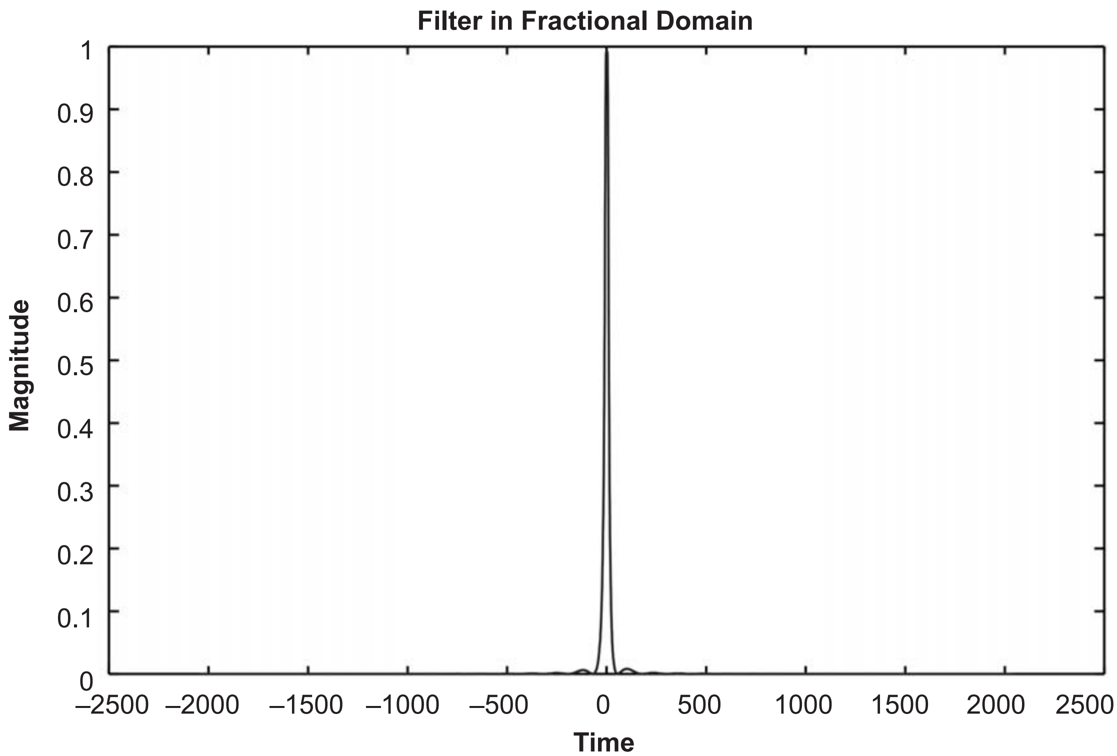

Figure 5: FFT filter in Fractional domain

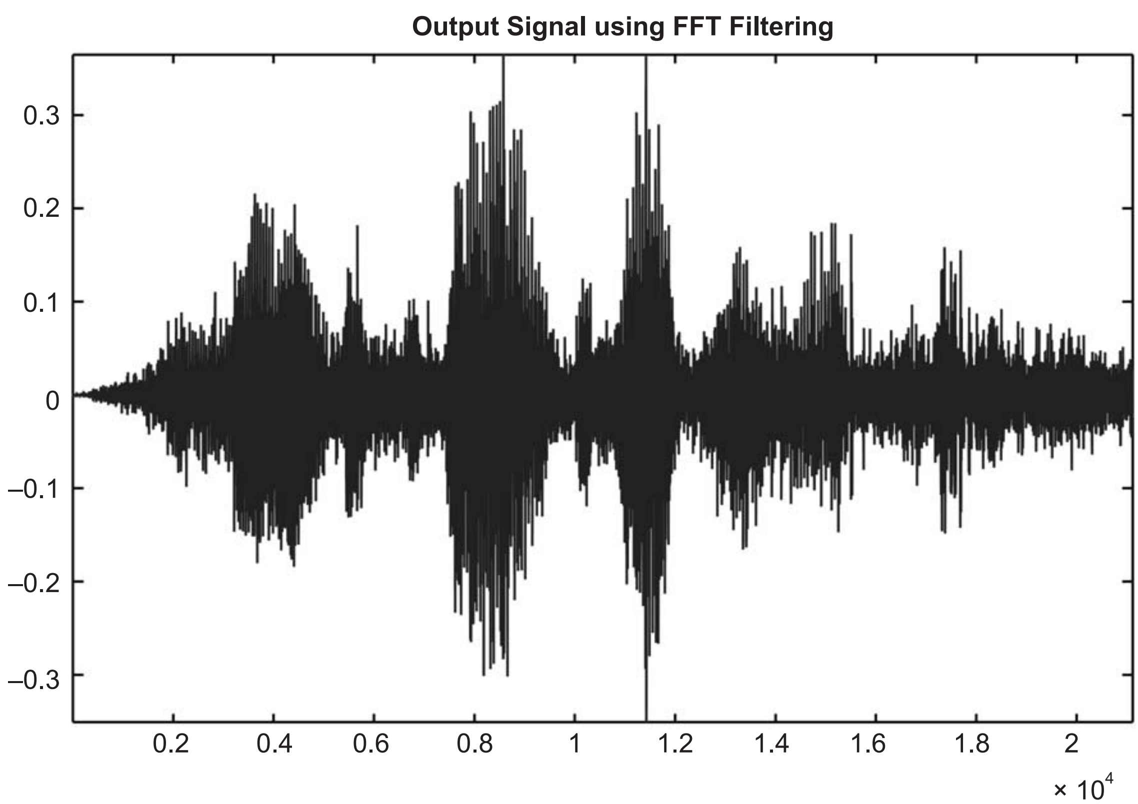

Figure 6: Enhanced signal after FFT filtering

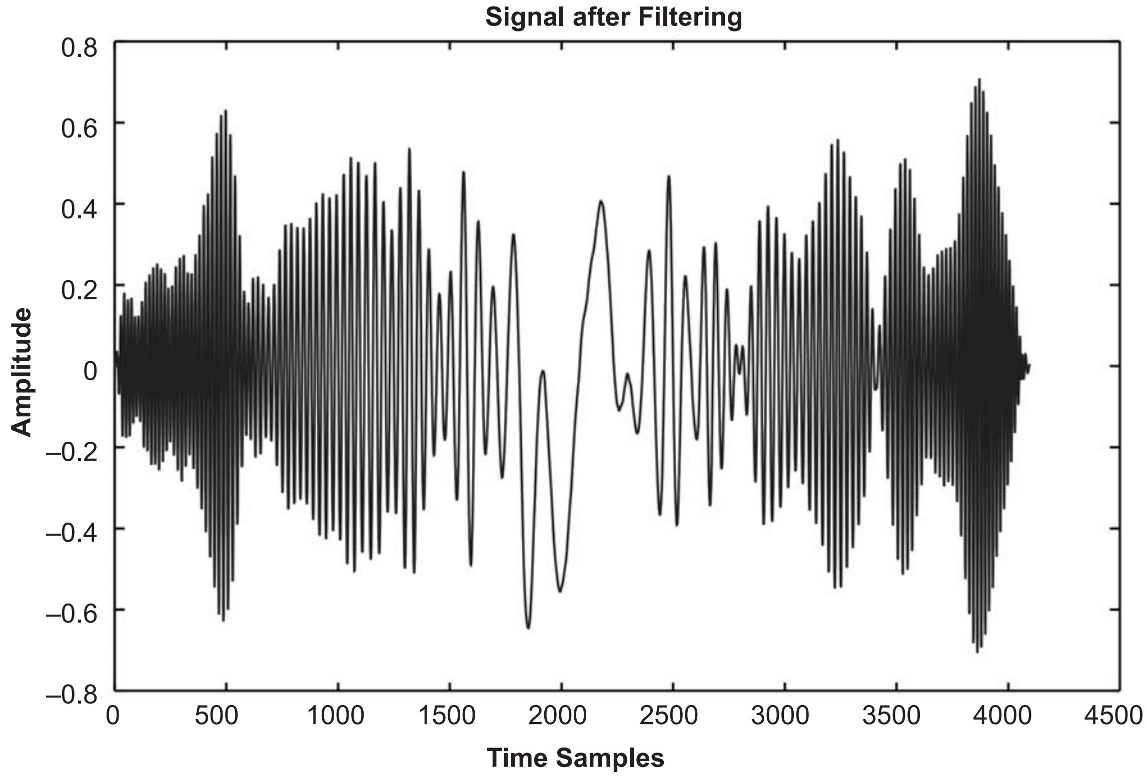

Figure 7: Enhanced signal after Fractional FFT filtering

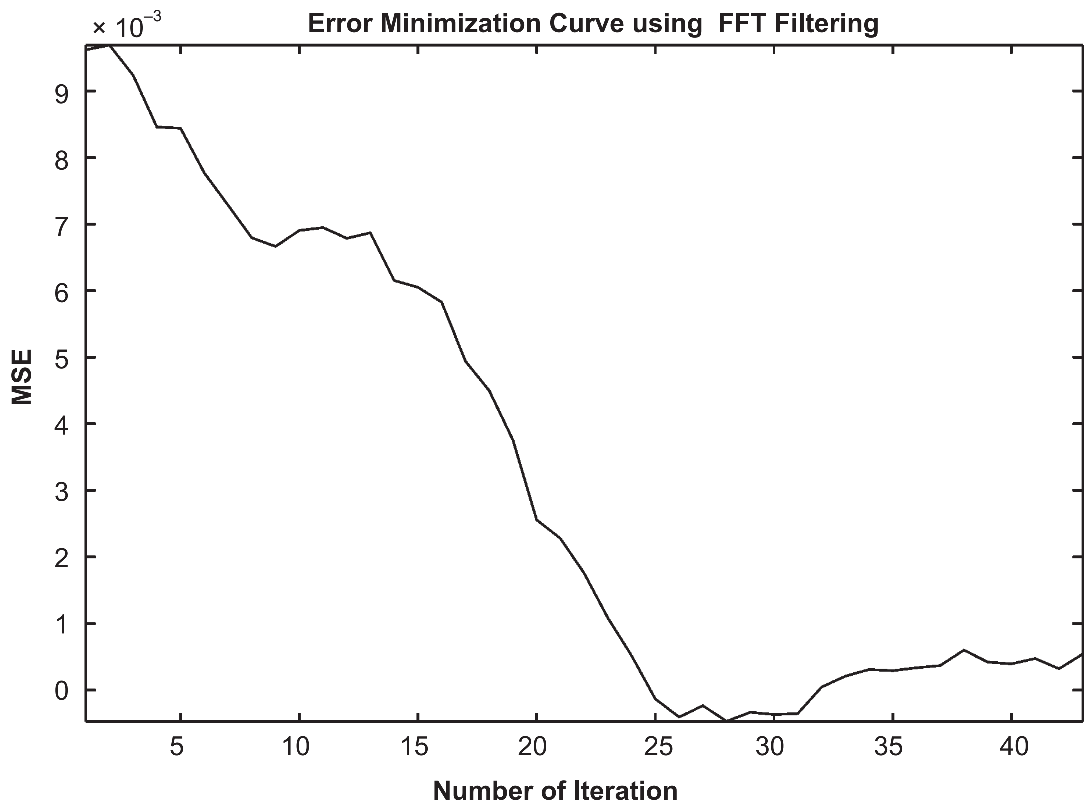

Figure 8: MSE curve for FFT filtering

The mean square error curve for FFT filtering and fractional FFT filtering has shown in Fig. 8 and Fig. 9 respectively. The MSE is observed between 40-50 iterations for FFT filtering and after 25 iterations MSE curve converges. But for Fractional FFT filtering above 300 iterations are performed and MSE curve converges with a continuous manner. After 260 iterations it is increased. The filtering in FrFT domain converges better than the filtering in FT domain.

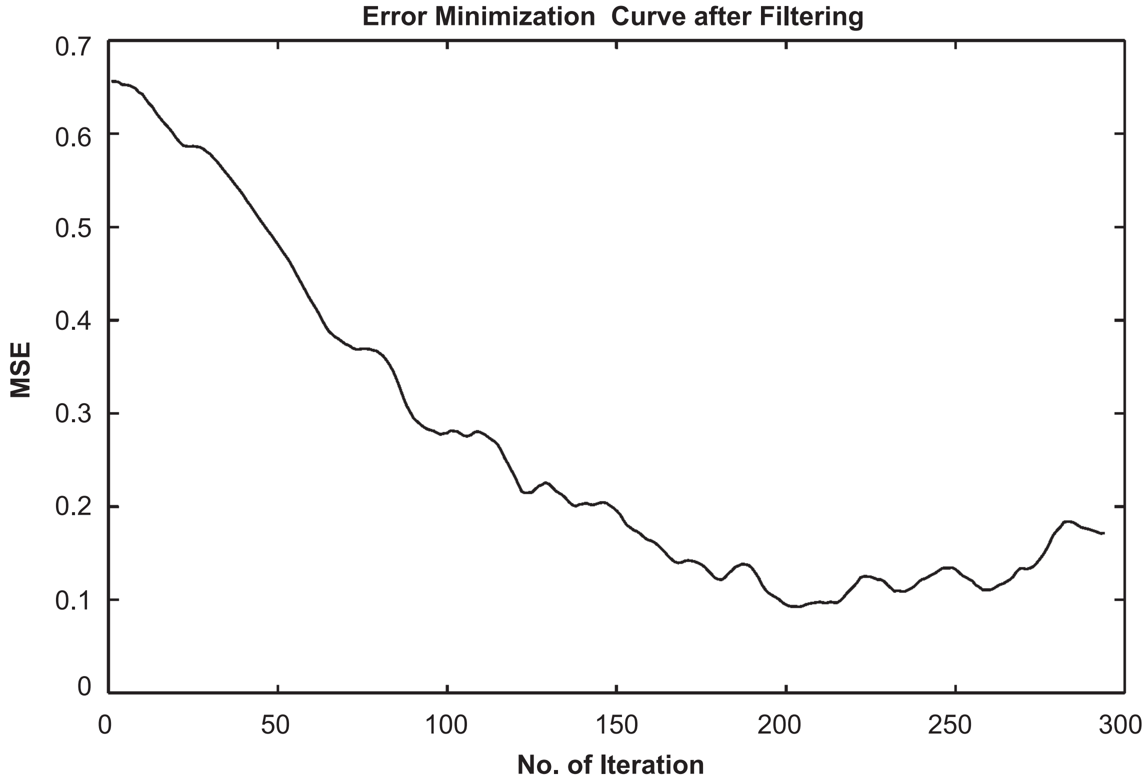

Figure 9: MSE curve for Fractional FFT filtering

Table 1 shows the Segmental SNR (SegSNR) for FFT filtering as well as for Fractional FFT filtering. SegSNR is one of the objective measures for speech quality. SNR of 0 dB, 5dB and 10dB of babble noise, white gaussian noise and car noise are added to the speech signal. For all SNR levels, FFT filtering in Fractional domain provide better results than FFT filtering.

TABLE 1
Segmental SNR Improvement for Different types of Noise Level

Noise	*Input Segmental SNR (dB)*	*Output Segmental SNR (dB)*	
		FFT Filtering	*FFT filtering in Fractional domain*
Babble	0 5 10	2.384 6.240 12.417	2.502 7.185 12.921
WGN	0 5 10	2.611 6.637 11.382	3.172 7.356 12.283
Car	0 5 10	1.894 7.462 12.073	3.203 8.246 12.989

The Perceptual Evaluation of Speech Quality (PESQ) is a standard methodology for testing the voice quality. Generally the PESQ score ranges from -1 to 4.5. Table 2 shows the PESQ score for both FFT filtering and fractional FFT filtering of different SNR levels. The maximum PESQ score is 3.64 which is obtained for car noise of 0 dB SNR.

TABLE 2
PESQ Scores for Different Types of Noise Level

Noise	*Input Segmental SNR (dB)*	*PESQ Scores*	
		FFT Filtering	*FFT filtering in Fractional domain*
Babble	0	2.22	2.35
	5	2.41	2.49
	10	2.45	2.52
WGN	0	2.61	3.13
	5	2.25	2.51
	10	3.12	3.61
Car	0	3.04	3.64
	5	2.91	3.27
	10	2.28	2.32

4. CONCLUSION

Filtering using FrFT is proposed for speech enhancement. Noisy speech signal has been enhanced successfully using the proposed method. First the noisy speech signal is processed through the FFT filtering. Then the filtering is performed in FrFT domain. The fractional order is chosen iteratively and the order 1.5 provides maximum SNR. The MSE curve, Segmental SNR and PESQ score are taken to compare fft filtering with Fractional FFT filtering. Results proved that the proposed method is more effective for enhancing the noisy signal. Filters with different fractional transform can be considered for future work.

5. APPENDIX

Fractional Fourier Transform:

Fractional Fourier transform is the generalization and transformation of the Fourier transform. FrFT can be taken as the Fourier transform to the a^{th} power. Where the integer a is the order of the Fourier transform. It rotates the time domain representation of the signal by π/2 in the counterclockwise direction. If $a = 1$, then the signal rotation is by π/2 and if $a = 2$ then the rotation is by π. But in FrFT, a can take any fractional value [13].

Let f and F are the two functions and form a Fourier transform pair:

$$F(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-i\omega x} dx \tag{1}$$

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F(\omega) e^{i\omega x} d\omega \tag{2}$$

If ξ is the FT operator, then

$$\begin{aligned} \xi^1[f(x)] &= F(\omega) \Rightarrow \text{Rotation over an angle } \pi/2 \text{ (FT)} \\ \xi^2[f(x)] &= F(-x) \Rightarrow \text{Rotation over an angle } \pi \\ \xi^3[f(x)] &= F(-\omega) \Rightarrow \text{Rotation over an angle } 3\pi/2 \\ \xi^4[f(x)] &= F(x) \Rightarrow \text{Rotation over an angle } 2\pi \end{aligned} \tag{3}$$

The Fractional FT defined as

$$F_\lambda(\omega) = \xi_\lambda f(x) = \int_{-\infty}^{\infty} f(x) K_\lambda(x,\omega) dx \tag{4}$$

Where $K_\lambda(x, \omega$ = Kernel function and defined as [16]

$$K_\lambda(x,\omega) = \begin{cases} \sqrt{\frac{1-j\cot\lambda}{2\pi}}\exp\left(j\frac{x^2+\omega^2}{2}\cot\lambda - j(\omega.x)\csc\lambda\right) & ,\lambda \text{ is not multiple of } \pi \\ \delta(x-\omega), & \lambda \text{ is multiple of } \pi \\ \delta(x+\omega), & \lambda \text{ is a multiple of } 2\pi \end{cases} \quad (5)$$

$\delta(x)$ is the Dirac delta function. The FrFT operator is linear as well as additive. The integral transformations are characterised by complex exponential kernels. The α^{th}- order FrFT is obtained by rotating the Wigner Distribution by an angle $\varphi = \alpha\pi/2$i n the clockwise direction [15].

6. REFERENCES

1. Soo-Chang Pei, Jian-Jiun Ding., "Relations between fractional operations and time-frequency distributions, and their applications", IEEE Trans. Signal Processing, pp 1638-1655, 2001.
2. H.M.Ozaktas, O. Arikan, M. A. Kutay and G. Bozdag, " Digital computation of the fractional Fourier transforms", IEEE Trans. Signal Processing, 44, pp 2141–2150,1996.
3. C.Candan, M.A.Kutay, and H.M.Ozaktas. "The discrete fractional Fourier transform". IEEE Trans. Signal Processing, vol.48, No.5, pp 1329-1337, May. 2000.
4. Duo-jia Ma, Xiang Xie, Jing-ming Kuan, "A Novel Algorithm of Seeking FRFT Order for Speech Processing", IEEE Conference, ICASSP, 2011.
5. Ruhi Sarikaya, Yuqing Gao, George Saon, "Fractional Fourier Transform features for Speech Recognition", IEEE Conference, ICASSP, 2004.
6. Rajiv Saxena, Kulbir Singh. "Fractional Fourier transform: A novel tool for signal processing" Indian Journal of Inst. Sci., Jan.–Feb. 85, pp 11–26, 2005.
7. Wang Zhenli and ZhangXiongwei, "On the application of Fractional Fourier Transform for Enhancing Noisy Speech", IEEE International Symposium on Microwave, Antenna, Propagation and EMC Technologies for Wireless Communications Proceedings, 2005.
8. A. Rutwik and B.T. Krishna, "Speech Enhancement using Fractional Order Spectral Subtraction Method", International Journal of Computer Technology and Applications, 8(3), pp. 1063-1070, 2015.
9. Hui Yin, Climent Nadeu, Volker Hohmann, Xiang Xie, Jingming Kuang, "Order Adaptation of the Fractional Fourier Transform using the Intraframe Pitch Change Rate for Speech Recognition", IEEE Conference, 2008.
10. 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, pp 547-559, February 1994.
11. M. Alper Kutay, Haldun M. Ozaktas, Levent Onural, and Orhan Arzkan, "Optimal Filtering in Fractional Fourier Domains", IEEE Conference, 1995.
12. Duo-jia Ma, Xiang Xie, Jing-Ming Kuang, "A Novel Algorithm of Seeking FrFT Order for Speech Enhancement", IEEE Conference, 2010.
13. Soo-Chang Pei1, Chun-Lin Liu2, and Yun-Chiu Lai, "The Generalized Fractional Fourier Transform", IEEE Conference, ICASSP, 2012.
14. Penghua Li, Yuanyuan Li, Dechao Luo, Hongping Luo, "Speaker Identification Using FrFT-based Spectrogram and RBF Neural Network", Proceedings of the 34th Chinese Control Conference, pp 28-30, July, 2015.
15. Ervin Sejdic, Igor Djurovic, LJubisa Stankovic, "Fractional Fourier transform as a Signal Processing Tool: An Overview of Recent Developments", Elsevier, Signal Processing 91, pp 1351–1369, 2011.
16. Sana Alaya, Novlene Zoghlami and Zied Lachiri, "Adaptive Filter for Perceptual Speech Enhancement", International Multi-Conference on Systems, Signals & Devices, 2015.
17. Sayed. A. Hadei, 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, April 2010.
18. Wu Caiyun, "Study of Speech Enhancement System Using Combinational Adaptive Filtering", International Conference on Computer Distributed Control and Intelligent Enviromental Monitoring, 2012.