Noise Removal in Speech signal using Fractional Fourier Transform

Prafulla Kumar
Electronics and Communication Department
Medicap Institute of Technology and Management
Indore, India
prafulla93ethane@gmail.com

Sarita Kansal
Electronics and Communication Department
Medicap Institute of Technology and Management
Indore, India
Sarita_1411@yahoo.com

***Abstract*—The Fractional Fourier Transform (FRFT) technique is derived from classical Fourier Transform. It is used in many applications. In this paper Fractional Fourier Transform is applied to denoise the noisy speech signal. The white Gaussian Noise is considered here which is affecting the speech signal. The performance of Fractional Fourier Transform is evaluated for different SNR and hard thresholding value. The results are presented by the measuring parameters MAE and PSNR. It is observed with fewer thresholds and SNR value, the MAE is minimum and PSNR is good.**

Keywords: **Fourier Transform, FRFT, PSNR, MAE etc.**

I. INTRODUCTION

In communication the reconstruction of speech signal is very crucial, if it is affected by noise. The amount of noise is more dominating than Speech signal. It is desired to denoise the speech signal, so that original message can be recovered properly. In this paper Fractional Fourier Transform is used to denoise the speech signal.

Fourier Transform is more suitable for stationary type of signal where as Fractional Fourier Transform deals with non stationary signals **[5]**. But for the non stationary signals FRFT is more suitable. In Fractional Fourier Transform the signal in time and frequency domain represents more clearly as overlapping is very small. The basic implementation of Fractional Fourier Transform and its applications are explained in **[1]**. The new algorithm introduced an algorithm known as Fractional Chirp Scaling Algorithm (FrCSA) for Fast Fourier transform for Small Aperture Radar imaging. This paper derived the mathematical expression for the azimuth-fractional Fourier transform of the new FrCSA in closed form. This allows high resolution imaging implementation. The closed-form expression in the derivation allows understanding the features required in azimuth direction for the Fractional Fourier Transform which includes fractional filtering, noise removal and residual phase compensation **[2]**. The performance is estimated for various values of fractional order, α and then the results are compared with the standard techniques by accessing different definitions of DFRFT. The paper concluded about the Spectral subtraction method as an efficient algorithm for speech enhancement **[3].** In **[4]** paper studied and concluded that most of the conventional de-noising methods are unable and inadaptable in the intense noise environment. Therefore the method for speech enhancement based on Fractional Fourier Transform (FRFT) filtering is suggested. The method is implemented by putting the noising speech in disperse degree of FRFT optimization. The acoustic voice signals are framed in the experiments and results disclose that this algorithm is capable to filter noise from available voice signals and can enhance the overall performance of speech recognition automation system significantly.

The paper is organized as follows; In Section II the brief introduction of Fractional Fourier Transform (FRFT) is given. Section III explains the detail of speech signal with its properties. Section IV describes the methodology used to remove the noise by using FRFT. Section V represents the simulation results and analysis. The last Section concludes the paper with conclusion.

II. FRACTIONAL FOURIER TRANSFORM (FRFT)

Fractional Fourier Transform is an extended form of Fourier Transform [1]. It is based upon the rotation angle denoted by 'α' whose period is between 0 to 2π. The rotation angle α is calculated as α = (aπ/2) where a € R. If time and frequency axis (t and ω) is rotated by an angle α in counter clockwise, then the rotated variables can be represented by u and v in the form of matrix as shown below:

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} t \\ w \end{bmatrix} \qquad (1)$$

Where, u and v should always orthogonal to each other.

If the signal in time domain is a rectangular pulse then it is realized as a sinc function in frequency domain. Then Fractional Fourier Transform converts the rectangular pulse to be in the domain between time and frequency shown in figure1.

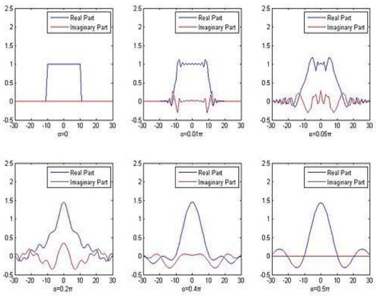

Figure 1: Waveforms of the FRFT

Mathematical Concept of FRFT

The FRFT of any function x with an angle α can be defined as the function $R^{\alpha} = X_{\alpha}$.

And X_{α} (u) can be given by:

$$X_{\alpha}(u) = \int_{-\infty}^{\infty} x(t) K_{\alpha}(t,u) dt \quad (2)$$

Here

$$K_{\alpha}(t,u) = \begin{cases} \sqrt{\frac{1-j\cot\alpha}{2\pi}} e^{j\cot\alpha\frac{(t^2+u^2)}{2}} - jutcsc\alpha \end{cases} \quad (3)$$

When α is not multiple of π.

K_{α} (t, u) = δ (t-u),

When δ is the multiple of 2π

K_{α} (t, u) = δ (t+u),

When α + π is the multiple of π

A. Properties of FRFT

Zero rotation- $R^0 = I$, Identity operator

FT operator- $R^{\pi/2}$

Time reverse property- R^{π}

Inverse FT operator- $R^{3\pi/2}$

2π Rotation- $R^{2\pi} = I$

Additivity property- $R^{\alpha} R^{\beta} = R^{\alpha+\beta}$

III. SPEECH SIGNAL

The speech signal has three main characteristics and those are Zero Crossing Rate (ZCR), Energy and Auto correlation. The rate at which the speech signal crosses zero and provides information about the source of its origin is known as Zero crossing Rate. It should be minimum always so that source can easily and quickly be detected. It is revealed that unvoiced speech has a much higher ZCR than voiced speech. The energy of unvoiced signal is lower than the voiced signal. The Auto-correlation is calculated between two consecutive pitch cycles of the same signal. Its values are values between pitch cycles are lower (close to 0) in voiced signal than in unvoiced signal.

Among three characteristics mentioned above, energy and zero crossing rates are significant features that make the end point detection possible in speech signal analysis. Both these features are utilized in segmentation and classification of voiced and unvoiced signal.

IV. NOISE REMOVAL USING FRFT

The received signal can be distorted, noisy or degraded form of the transmitted signal. Therefore, to improve the strength of the signal, removal of noise is necessary. Thus to recover the desired signal at the receiver, filtering is done. The signals transmitted from transmitter end may vary at the receiving end. The message (voice) grabbed after transmission is corrupted with noise most of the time. Due to a noisy channel in the transmission medium, the noise gets introduced in the signal. Errors occur in the measurement process and quantization process of the data for digital storage. The signal obtained at the receiver demands for processing prior to use for further applications. Noise degrades the quality of the sound at the time of capturing or transmission of the voice.

The block diagram of the processing is shown in figure 2. The hard thresholding is applied on the spectrum received after Fractional Fourier Transform [13].This reduces the noise level in the signal. After thresholding inverse fractional Fourier transform is applied to the get the denoised speech signal.

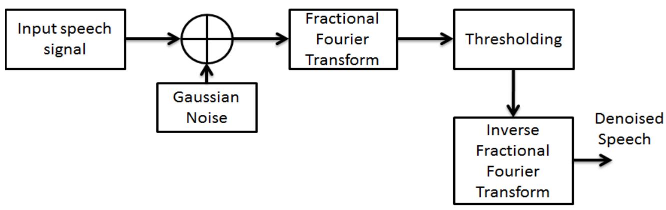

Figure 2: Block diagram of denoising process

Before applying signal processing tools to the speech signal, noise removal from voice signals is done at highest priority. To restore the original speech signal at the receiver end is the challenging task for the researchers.

V. SIMULATION AND RESULTS

The noise of the speech signal is removed by using the hard thresholding on Fractional Fourier transform domain signal for different signal to noise ratio. The parameters used for the simulation are given in the TABLE I.

TABLE I. SIMULATION PARAMETRS

	Parameters	*Values*
1	Sampling rate	44100 Hz

	Parameters	*Values*
2	Speech duration	18.46 Sec
3	Fractional coefficent	0.75
4	Threshold range	0.005 to 20 % of max signal
5	SNR range	0 to 30 dB

The speech signal is generated of duration 18.46 seconds as shown in the figure 3. The signal is passing through AWGN channel and the noisy signal is received at the receiver end.

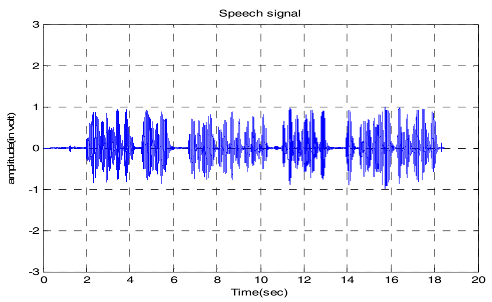

Figure 3

The Fractional Fourier transform is applied on the noisy signal and its spectrum is shown in the figure 4. The Fractional Fourier Transform is applied with coefficient value 0.75, SNR equal to 5 dB with threshold equal to 0.175.

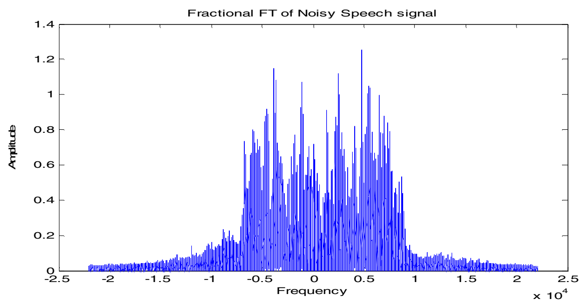

Figure 4

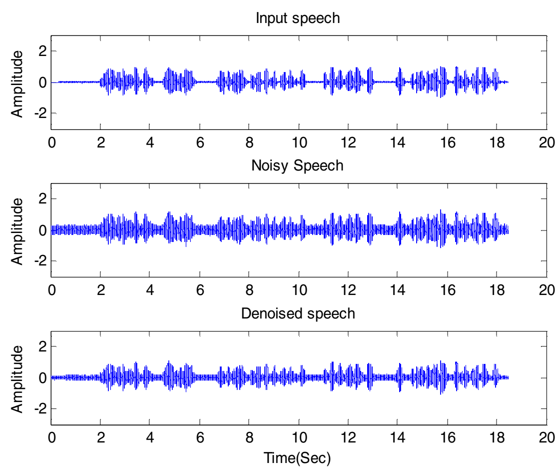

Figure 5

The figure 5 shows the performance of input speech signal, noisy speech and denoised speech signal. The measurement parameter MAE and PSNR is calculated with different SNR and thresholding value.

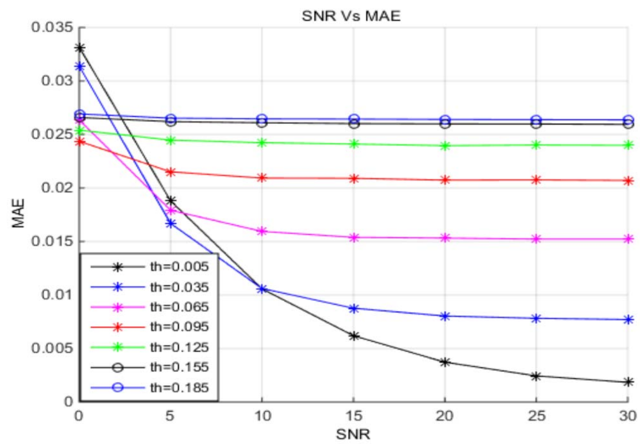

Figure 6: The Performance of variations in MAE with SNR

The Fractional Fourier Transform for denoising the signal is applied with different SNR values and its performance is shown in figure 6. The increase in SNR decreases MAE value at different thresholds.

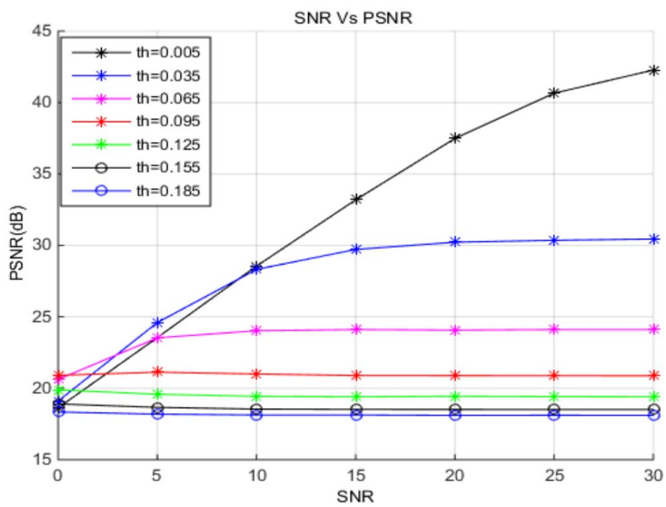

Figure 7: The Performance of PSNR with SNR.

The Fractional Fourier Transform for denoising the signal is applied with different SNR values to obtain PSNR and its performance is shown in figure 7. The increase in SNR increases PSNR value at different thresholding. The maximum PSNR is obtained at maximum value of SNR. The improvement of the above parameters is significant at low

SNR. The SNR less than 7 dB is the area where this technique shown a significant improvement.

VI. CONCLUSION

The work is based on FRFT technique to remove the noise from speech signal. The Gaussian noise is considered here which affected the speech signal. The results are presented with different values of SNR and hard thresholding is also used with different SNR values to denoise the speech signal. The PSNR is improved by 3.6 dB with applying the Fractional Fourier transform and thresholding. The thresholding level is decreasing monotonically with increase in SNR value. This method offers many advantages over the other denoising method due to less perturb to the original content of the speech. Hence other technique may be hybrid with this technique.

REFERENCES

[1] V. Ashok Narayanan, K.M.M. Prabhu, “The Fractional Fourier Transform: Theory, Implementation and Error Analysis”, www.elsevier.com/locate/micpro, June 2003.

[2] A. S. Amein, John J. Soraghan, “The Fractional Fourier Transform and its Application to High Resolution SAR Imaging”, IEEE, 2007.

[3] A. Rutwik and B.T. Krishna, “Speech Enhancement using Fractional Order Spectral Subtraction Method”, IJCTA, 2015.

[4] Jingfang Wang, “Speech Enhancement based on Fractional Fourier transform”, WSEAS Transactions on Signal Processing, Vol.10, 2014.

[5] M. F. Edren, M. A. Kutay and H. M. Ozaktas, “Repeated Filtering in Consecutive Fractional Fourier Domains and its Applications to Signal Restoration, IEEE Trans. Signal Processing, 1999.

[6] V. Namias, “The Fractional Order Fourier transform and its Application to Quantum Mechanics,” J. Inst. Maths. Applics., vol. 25, pp. 241-265, 1980.

[7] E.U. Condon, “Immersion of the Fourier Transform in a Continuous Group of Functional Transformations”, Proc. National Academy Sciences, 1937.

[8] Rachit Patel, “Review Paper on Fractional Fourier Transform”, International Journal of Advance Research in Science and Engineering, vol.no.5, 5 May 2016.

[9] Pao-Yen Lin, “The Fractional Fourier Transform and Its Applications”, IJCTA, 2010.

[10] Rainer G. Dorsch, Adolf W. Lohmann, Yigal Bitran, David Mendlovic, and Haldun M. Ozaktas, “Chirp Filtering in the Fractional Fourier Domain”, Vol. 33, No. 32, 10 November 1994.

[11] Chi Chang-yanab, Zhang Ji-xiana , Liu Zheng-juna, “Study on Methods of Noise Reduction in a Stripped Image”, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Vol. XXXVII. Part B6b, 2008

[12] Wang, Yao; "Source Coding Basics and Speech Coding", Polytechnic University, Brooklyn, NY11201, ©Yao Wang, (2006).

[13] Milind U. Nemade , Prof. Satish K. Shah, “Performance Comparison of Single Channel Speech Enhancement Techniques for Personal Communication”, International Journal of Innovative Research in Computer and Communication Engineering Vol. 1, Issue 1, March 2013.

[14] Haldun M. Ozaktas, Orhan Arikan, “Digital Computaion of the Fractional Fourier Transform”, IEEE Transactions on Signal Processing, Vol.44,No.9,September 1996.