Speech Enhancement using Compressive Sensing

K.Kiruthiga[1], J.Indra[2]

[1]*PG Scholar,Dept. Of EIE, Kongu Engineering College, Tamilnadu,India*
[2]*Assistant Professor (SLG),Dept. Of EIE, Kongu Engineering College, Tamilnadu, India*

---***---

Abstract - *Speech enhancement is a technique which is used to reduce the background noise present in the speech signal. The noises are additive noise, echo, reverberation and speaker interference. The aim of the proposed method is to reduce the background noise present in the speech signal by using compressive sensing. The goal of compressive sensing is to compress the speech signal at transmitter and decompress it at the receiver from far less samples than the nyquist rate. In this work, a speech signal is taken and then it is compressively sampled using a measurement matrix which in case is composed of randomly generated numbers. The output of the compressed sensing algorithm is the observation vector which is transmitted to the receiver. At the receiver section, signal is reconstructed from a significant small numbers of samples by using l1- minimization. MATLAB simulations are performed to compress the speech signal below the nyquist rate and to reconstruct it without losing any important information.*

Key Words: Speech enhancement, Compressive sensing, DCT, *l*1 –minimization, Measurement matrix.

1. INTRODUCTION

In recent years, various signal sampling schemes have been developed. However, such sampling methods are difficult to implement. So before sampling the signal it should have sufficient information about the reconstruction kernel. The emerging compressive sensing theory shows that an unevenly sampled discrete signal can be perfectly reconstructed by high probability of success by using different optimization techniques and by considering fewer random projections or measurements compared to the Nyquist standard. Amart Sulong et al proposed the compressive sensing method by combining randomized measurement matrix with the wiener filter to reduce the noisy speech signal and thereby producing high signal to noise ratio [1]. Joel A. Tropp et al demonstrated the theoretical and empirical work of Orthogonal Matching Pursuit (OMP) which is effective alternative to (BP) for signal recovery from random measurements [2]. Phu Ngoc Le et al proposed an improved soft – thresholding method for DCT speech enhancement [3]. Vahid Abolghasemi focused on proper estimation of measurement matrix for compressive sampling of the signal [4].

2. Compressive sensing

Compressive sensing involves recovering the speech signal from far less samples than the nyquist rate [8]. Fig.1 shows the basic block diagram of compressive sensing. Initially, the signal is sampled using nyquist rate, whereas with the help of compressive sensing the signal is sampled below the nyquist rate [5]. The signal is transformed into a domain in which it shows sparse representation. Then the signal is transmitted and stored in the channel by the receiver side [13].

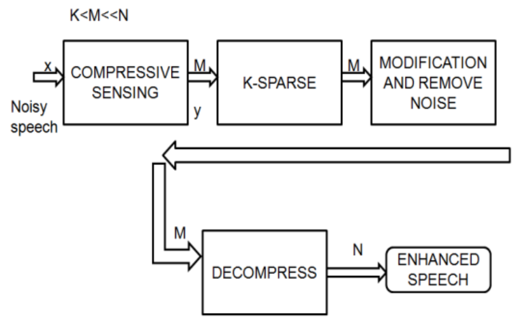


Fig.1 Basic block diagram of compressive sensing

Finally the signal is reconstructed from the samples by using one of the different optimization techniques available.

3. Noizeus Corpus

Thirty sentences from the IEEE sentence database (IEEE Subcommittee 1969) were recorded in a sound-proof booth using Tucker Davis Technologies (TDT) recording equipment [12]. The sentences were produced by three male and three female speakers. The sentences were originally sampled at 25 kHz and down sampled to 8 kHz and eight basic noise signals under different environmental conditions are taken from the AURORA database [9]. It has the recordings from different places like Babble, Car, Exhibition hall, Restaurant, Street, Airport, Train station and Train.

4. Proposed Method for Speech Enhancement Using Compressive Sensing algorithm

The proposed speech enhancement algorithm using compressive sensing is illustrated in Fig.2
Analysis filter bank uses gammatone filter due to its resemblance to the shape of human auditory filters. Then discrete cosine transform is chosen due to its simplicity. Subband modification is applied to produce subband coefficients for analysing the speech signal. On synthesis side, for solving convex optimization of compressive sensing, the gradient projection of sparse reconstruction algorithm is used [1]. Higher the processing power, higher is the quality of the signal synthesized.

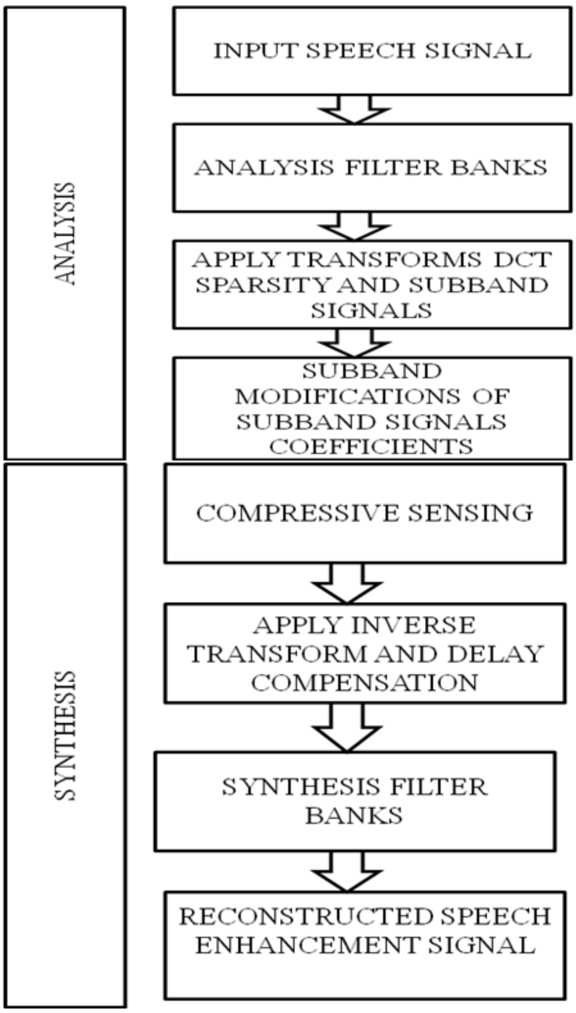


Fig.2 Flowchart for Proposed Speech Enhancement Algorithm

5. Measurement Matrix

A random matrix is a matrix with random entries. A random matrix (sometimes stochastic matrix) is a matrix-valued random variable, in which some matrix or all of whose elements are random variables.

$$y = \mathbf{\Phi x} = \mathbf{\Phi \Psi \alpha} \qquad \textbf{(1)}$$

Where, $\mathbf{\Phi}$ is a $M \times N$ measurement matrix with each row be a measurement vector φ_j^T

$\boldsymbol{\alpha}$ is the coefficient vector with K non zeroes element. The measurement matrix plays a major role in the process of recovering the original signal. In compressive sensing there are two types of measurement matrices namely, random measurement matrix and the predefined measurement matrix [14].

$$(P1)\ \min \| x \|_{l1} \text{subject to } \mathbf{\Phi x = y} \qquad \textbf{(2)}$$

which is also known as basis pursuit (P1).

$$\| x \|_{l1} = \sum_{i=1}^{n} | x_i | \qquad \textbf{(3)}$$

It is otherwise known as Taxicab norm Manhattan norm [2]. The distance obtained from this norm is called the Manhattan distance or $l1$ distance.

6. Optimization Techniques

Signal reconstruction plays an major role in compressive sensing theory where the signal is reconstructed or recovered from a less number of measurements [4]. By using optimization techniques it is possible to recover the signal without losing the information at the receiver.

6.1. $l1$ Minimization

$l1$ minimization is used to solve the under determined linear equations or sparsely corrupted solution to an over determined equations [11].

7. Conventional Thresholding

In the proposed method soft thresholding is followed due to its advantages [3].The soft thresholding is defined as

$$\text{Ysoft} = \begin{cases} sgn(x) * (|x| - |\lambda|), & |x| \geq |\lambda| \\ 0 & |x| < |\lambda| \end{cases} \qquad \textbf{(4)}$$

International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056
Volume: 04 Issue: 03 | Mar -2017 www.irjet.net p-ISSN: 2395-0072

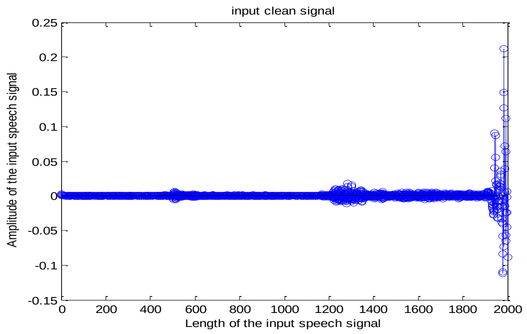
8. Experimental Results and Discussions

8.1. Input Speech and Noise Representation in Time Domain

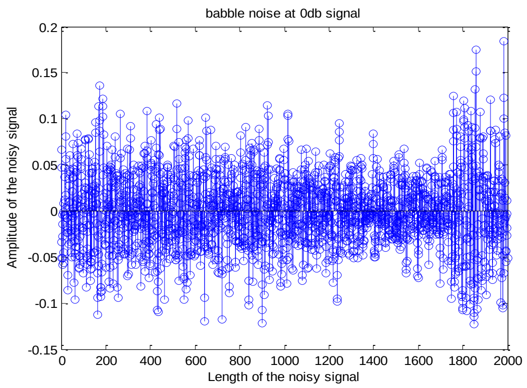
The sample sentence is **"Clams are small, round, soft and tasty"**

(i). Enhanced Output from Babble Noise (0db)

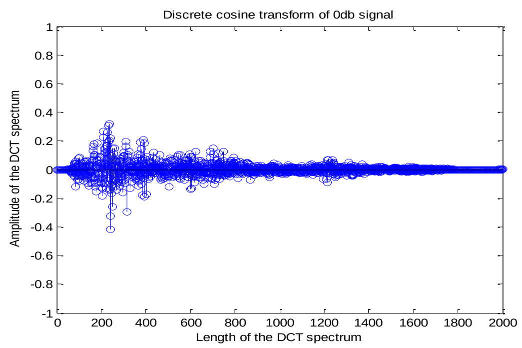
The enhanced output from babble noise is shown in the Fig.4



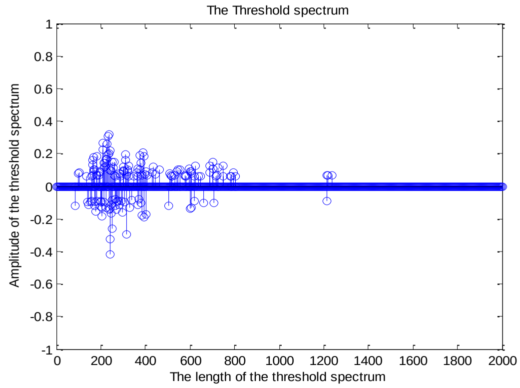
(a)



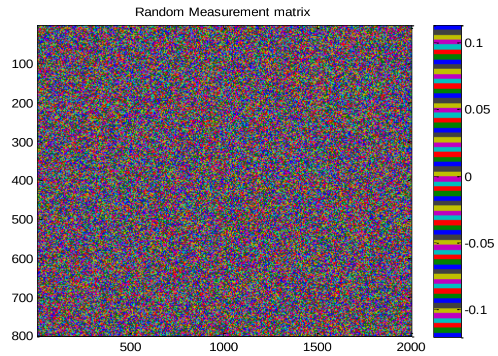
(b)



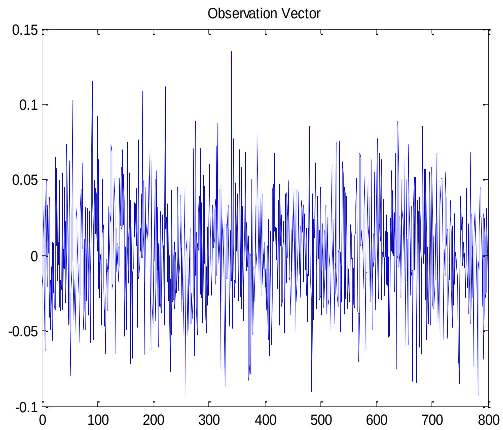
(c)



(d)



(e)



(f)

Fig.4 Enhanced Output From Babble Noise at (0db) (a) Clean Speech (b) Noisy Speech (c) Applying DCT (d) Thresholding (e) Random Measurement Matrix (f) Output for Compressive Sensing

Figure 4, shows input clean speech signal in (a) and then adding clean speech and babble noise at 0db which is composed of 2000 samples in (b), The recorded speech signal goes through DCT which transforms the sequence of real data points into real spectrum and is shown in (c). The threshold window is then applied to eliminate the small coefficients as shown in (d). Threshold spectrum is multiplied by measurement matrix which is composed of randomly generated numbers as shown in (e), and then compressive sensing algorithm is used in (f).

Table 1: Amount of signals compressed by using signal parameters

NOISE SIGNAL AT DIFFERENT (DB)		LENGTH OF THE SIGNAL (L)	THRESHOLD WINDOW UL	LL	COMPESSED SAMPLE(K)	ERROR(%)	COMPRESSION(%) (K/L)
Babble	0db	2000	0.04	-0.06	800	0.1089	40%
	5db	2000	0.04	-0.06		0.1361	
	10db	2000	0.04	-0.06		0.1931	
	15db	2000	0.04	-0.06		0.1821	
Airport	0db	2000	0.04	-0.06	800	0.1037	40%
	5db	2000	0.04	-0.06		0.1900	
	10db	2000	0.04	-0.06		0.1777	
	15db	2000	0.04	-0.06		0.1975	
Car	0db	2000	0.04	-0.06	800	0.0795	40%
	5db	2000	0.04	-0.06		0.1738	
	10db	2000	0.04	-0.06		0.2193	
	15db	2000	0.04	-0.06		0.1963	

Table 2: Amount of signals compressed by using signal parameters

NOISE SIGNAL AT DIFFERENT (DB)		LENGTH OF THE SIGNAL (L)	THRESHOLD WINDOW UL	LL	COMPRESSED SAMPLE(K)	ERROR (%)	COMPRESSION(%) (K/L)
Babble	0db	2000	0.04	-0.06	800	0.1047	40%
	5db	2000	0.04	-0.06		0.0802	
	10db	2000	0.04	-0.06		0.0810	
	15db	2000	0.04	-0.06		0.1179	
Airport	0db	2000	0.04	-0.06	800	0.0835	40%
	5db	2000	0.04	-0.06		0.0708	
	10db	2000	0.04	-0.06		0.0608	
	15db	2000	0.04	-0.06		0.0693	
Car	0db	2000	0.04	-0.06	800	0.1273	40%
	5db	2000	0.04	-0.06		0.0717	
	10db	2000	0.04	-0.06		0.0568	
	15db	2000	0.04	-0.06		0.0746	

From Table 1, it is observed that for babble noise with noise level as 5db, length of the signal as 2000, threshold window from 0.04 to -0.06 and compressed samples as 800, the error is 13.61% and compression is upto 40%. From Table 2, it is observed that for babble noise with noise level as 5db, length of the signal as 2000, threshold window from 0.04 to -0.06, and compressed samples as 800, the error is 8.02% and compression is upto 40%.

Table 3: Amount of Signals Compressed by using Signal Parameters

NOISE SIGNAL AT DIFFERENT (DB)		LENGTH OF THE SIGNAL (L)	THRESHOLD WINDOW UL	LL	COMPRESSED SAMPLE(K)	ERROR (%)	COMPRESSION(%) (K/L)
Babble	0db	2000	0.04	-0.06	800	0.0866	40%
	5db	2000	0.04	-0.06		0.0863	
	10db	2000	0.04	-0.06		0.1009	
	15db	2000	0.04	-0.06		0.0887	
Airport	0db	2000	0.04	-0.06	800	0.0894	40%
	5db	2000	0.04	-0.06		0.0767	
	10db	2000	0.04	-0.06		0.0939	
	15db	2000	0.04	-0.06		0.0834	
Car	0db	2000	0.04	-0.06	800	0.1410	40%
	5db	2000	0.04	-0.06		0.0995	
	10db	2000	0.04	-0.06		0.0879	
	15db	2000	0.04	-0.06		0.0967	

From Table 3, it is observed that for babble noise with noise level as 5db, length of the signal as 2000, threshold window from 0.04 to -0.06, and compressed samples as 800, the error is 8.63% and compression is upto 40%.

9. Conclusion and Future Scope

During the design process, this module went through different tests and analysis in order to find the most adequate optimization technique to reconstruct the speech signal with few random measurements without losing the information. For simulation purposes, code was created in order to compress and transmit the speech signal below the Nyquist rate by taking only a few measurements of the signal. As a result, it shows that by keeping the length of the signal (L) and threshold window (Th) constant we can achieve the desired compression of the signal by making the signal sparse (K) to a certain amount which in turn increases the data rates. After multiple simulations, it was found that the system worked as expected and the speech signal was reconstructed efficiently with a minimum error.

The speech signal was reconstructed without losing important information that leads to increase in data rate. Some of the future works are as follows. Different transformations need to be tested in order to find the most efficient one for this application. A measurement matrix that will be optimum for speech signals is to be designed. The proposed method has to be tested with other existing methods to prove its efficiency.

10. References

[1] Amart Sulong, Teddy Surya Gunawan, Othman O. Khalifa, Mira Kartiwi,' Speech Enhancement based on Wiener filter and Compressive Sensing', Indonesian Journal of Electrical Engineering and Computer Science, Vol. 2, issue.2, pp. 367-379, 2016.

[2] Joel A. Tropp and Anna C. Gilbert, 'Signal Recovery from Random Measurements via Orthogonal Matching Pursuit', IEEE Transactions on Information Theory, pp. 4655- 4666, Dec 2007.

[3] Phu Ngoc Le, Eliathamby Ambikairajah and Eric Choi, 'An Improved Soft Threshold Method for DCT Speech Enhancement', IEEE, pp. 268-271, July 2008.

[4] Vahid Abolghasemi, Saideh Ferdowsi, Bahador Makkiabadi, and Saeid Saneis, 'On Optimization of the Measurement matrix for Compressive Sensing', 18th European Signal Processing Conference (EUSIPCO-2010), semantic scholar, pp. 427-431, Aalborg, Denmark ,Aug 2010.

[5] Amart Sulong, Teddy S.Gunawan, Othman O.Khalifa, and Jalel Chebil, 'Speech Enhancement based on Compressive Sensing', 5th International Conference on Mechatronics (ICOM'13), pp. 1-10. 2013.

[6] Anna maria jose and Mathurakani M, 'Compressive Sensing Based OFDM Channel Estimation', International Journal of Modern Sciences and Engineering Technology, Vol. 3, issue. 2, pp.13-20, 2016.

[7] Donoho D.L, 'Compressed Sensing', IEEE Transactions on Information Theory, vol. 52, pp.1289-1306, April 2006.

[8] Emmanuel J. Candes and Michael B. Wakin, 'An Introduction to Compressive Sampling', IEEE Signal Processing Magazine, Vol. 25, issue 2, pp. 21-30, March 2008.

[9] Hirsch H and Pearce D (2000), 'The Aurora Experimental Framework for the Performance Evaluation of Speech Recognition Systems under Noisy Conditions', ISCA ITRW ASR2000.

[10] Jagdeep Kaur, Kamaljeet Kaur, Monika Bharti, Pankaj Sharma and Jatinder Kaur 'Reconstruction Using Compressive Sensing: A Review', International Journal of Advanced Research in Computer and Communication Engineering, Vol.2, issue.9, pp.3648-3650,2013.

[11] Massimo Fornasier and Holger Rauhut, 'Compressive Sensing', Springer link handbook of mathematical methods and imaging, pp. 187-228, 2010.

[12] IEEE Subcommittee 1969 IEEE Recommended Practice for Speech Quality Measurements. IEEE Trans. Audio and Electroacoustics, AU-17(3), 225-246.

[13] Pooja C. Nahar, Dr. Mahesh and Kolte.T, 'An Introduction to Compressive Sensing and its Applications', International Journal of Scientific and Research Publications, Volume 4, Issue 6, pp . 1-5, June 2014.

[14] Richard G. Baraniuk, 'Compressive Sensing', IEEE Signal Processing Magazine, lecture notes, pp.118-120, 124, 2007.

[15] Siddhi Desai and Prof. Naitik Nakrani 'Compressive Sensing in Speech Processing: A Survey Based on Sparsity and Sensing Matrix', International Journal of Emerging Technology and Advanced Engineering, Volume 3, Issue 12, pp . 18-23,2013.

[16] Sabir Ahmed, 'Compressive sensing for speech signals in mobile systems', The University of Texas at Tyler, pp.1-51, 2011.