

Fixing of Sensing Matrix for Speech Signal Reconstruction Using OMP

K. V. Siddamal¹, Shubha P. Bhat²

¹(Associate Professor, Department of Electronics and Communication Engineering, Presidency University, Itgalpur, Rajankunte, Bangalore Karnataka, 560064, India,

²(Director, Karavali Institute of Engineering, Mangalore, Karnataka, 560064, India

Abstract: Compressive sensing is a technique for data acquisition that promises sampling a sparse signal from a far fewer measurements than its dimension. In this paper for speech signal "Hello" of length "n=16000" is reconstructed using Orthogonal Matching Pursuit. For better reconstruction an effort is made to fix the size of sensing matrix ϕ . Factors such as mean square error (MSR), signal to noise ratio (SNR) and peak signal to noise ratio (PSNR) are taken into considerations to measure the performance of the reconstructed signal.

Keywords: Compressive sensing, Sensing matrix, Sparse signal, OMP, Discrete Cosine Transform

I. Introduction

Compressive sensing is an emerging and revolutionary technology where the signal of interest is efficiently acquired from very few non-zero coefficients. CS provides both sampling and compression simultaneously [1], [2], [3]. Whatever is the field of application, most of the signal acquisition systems have been designed based on the Nyquist Shannon sampling theorem. The Nyquist sampling theorem states that the frequency used to sample a signal should be at least twice the maximum frequency contained within the signal. However, this is not an efficient way to compress the signal, as it increases the burden of sampling the entire signal while only a small percentage of the coefficients are needed to represent it. The importance of compressive sensing is that only the significant information about the signal/image is directly acquired rather than acquiring the significant information plus the information about the signal which will be eventually discarded at the receiver. Compressive Sensing basically relies on two things sparsity and incoherence [5]. Sparsity means the signal which comprises of small no of non-zero coefficients. On the other hand Incoherence means that, the signals which are represented sparsely should be spread out in a domain in which they are acquired. Compressive sensing can be used in medical imaging, image compression, radar system, mobile communication systems and in Analog information conversion [4]. The paper is organized as follows: section II provides a basic of compressive sensing. Section III discusses performance parameters. Section IV provides Experimental results. Section V discusses fixing of sensing matrix and conclusion is discussed in section VI.

II. Compressive Sensing Basics

The theory of CS was developed by Candes et al and Donoho in 2004 [4]. It involves taking random projections of the signal and recovering it from a small number of measurements using optimization techniques. This method is different from traditional Nyquist Shannon sampling method as it will sample the signal at a rate below the nyquist rate. In CS the signal is first transformed into a sparse domain and then the signal is reconstructed using different optimization techniques using small number of measurements. The basic principle of compressive sensing is as shown in the Figure 1.

The signal f with N samples is given as input at the transmitter side. Now this signal has to be converted into some domain in which it has sparse representation. For example DCT, FFT etc, after this conversion the signal is then transformed to a T -sparse signal using thresholding, where T represents the non-zero coefficients of f . Then this signal is multiplied with the sensing matrix Φ and the result will yield an observation vector y .

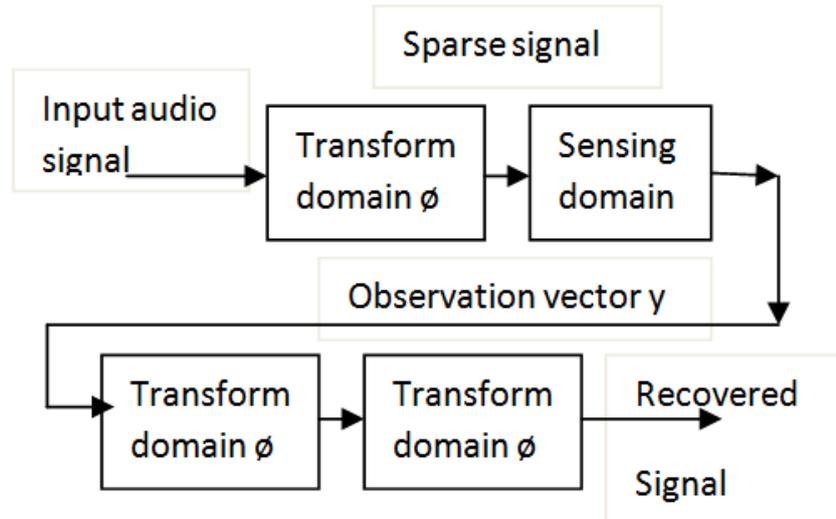


Fig1: Block Diagram of compressive sensing and reconstruction

At the receiver side different reconstruction algorithms are used to recover the original signal. At first the signal is multiplied with the sensing matrix Φ which gives N samples from m measurements. Then different optimization techniques are used to recover T -sparse signal. Finally the inverse sparsity is applied to obtain the original signal \hat{f} .

Mathematically speaking let $f \in R^N$ be an N -dimensional signal that is represented in proper orthogonal basis $\psi = \{\psi_1, \psi_2, \dots, \psi_N\}$. (i.e. ψ_i 's are orthogonal basis vectors spanning R^N). The signal is said to be sparse if,

$$f_T(t) = \sum_{i=1}^N x_{T_i} \psi_i(t) \quad t = 1, 2, \dots, N \quad (1)$$

Where x_T is the vector of coefficients (x_i) with all but the T largest coefficients set to zero. In other words, x_T is a sparse vector with only T non-zero elements where ($T \ll N$). Based on CS theory, sample the signal f in the sampling domain $\psi = \{\varphi_1, \varphi_2, \dots, \varphi_N\}$ and reconstruct the signal at the receiver with the full knowledge of random bases. The measurement signal is defined as

$$y(k) = \sum_{i=1}^N f_i \varphi_i(k) \quad k = 1, 2, \dots, m \quad (2)$$

Where φ is a measurement matrix of dimension $m \times N$. The φ is made up of orthogonal basis vector. The vector y can be reconstructed perfectly if $m > T \log N$ measurements (i.e. incoherence condition between φ and ψ should be satisfied).

For the reconstruction of sparse signals convex optimisation techniques are used. One such algorithm to solve CS problem “basic pursuit or l_1 – minimiation technique” which can be formulated as follows

$$\hat{x} = \operatorname{argmin} \|\hat{x}\|_1 \quad (3)$$

Subject to $y = \varphi \psi \hat{x} = A \hat{x}$

$\hat{f} = \psi \hat{x}$ where A is a reconstruction matrix and \hat{f} is a reconstructed signal.

Another efficient algorithm to solve CS problem is “Orthogonal Matching Pursuit” (OMP) which can be formulated as follows

$$\hat{x} = \operatorname{argmin} \|y - \varphi \psi \hat{x}\|_2 \quad (4)$$

Subject to $\|\hat{x}\|_0$

The idea behind this greedy algorithm is to pick the column in greedy fashion. Let the sensing matrix be $\psi = \{\varphi_1, \varphi_2, \dots, \varphi_N\}$ where φ_i is a $m \times 1$ column. At each of the T iterations, select the column φ_i that is most strongly correlated with y , then subtract its contribution to y and iterate on the subtraction residual.

III. Performance Evaluation

To evaluate the performance of the proposed speech compression scheme, several objective tests were made. Factors such as mean square error (MSR), signal to noise ratio (SNR) and peak signal to noise ratio (PSNR) are taken into considerations to measure the performance of the reconstructed signal.

1.1 Mean Square Error

The mean square error is defined by

$$MSR = \frac{\sum_n (a[n] - b[n])^2}{n} \quad (5)$$

Where $a[n]$ is original speech signal, $b[n]$ is reconstructed signal and n is the length of the signal.

1.2 Signal to Noise Ratio (SNR)

The signal to noise ratio is defined by

$$SNR = 10 \log_{10} \left[\frac{\sum_n a[n]^2}{\sum_n (a[n] - b[n])^2} \right] \quad (6)$$

Where $a[n]$ is original speech signal, $b[n]$ is reconstructed signal and n is the length of the signal.

1.3 Perceptual Evaluation Speech Quality (PESQ)

The Perceptual Evaluation of Speech Quality is an objective method of measuring speech quality. It is calculated using subjective Mean Opinion Scores (MOS). The range of PESQ lies within 0 to 5. Here the lower values are interpreted as poor.

IV. Experimental Results

The experiment is simulated on MATLAB 2010a. The audio signal is divided into number of frames. Each frame contains certain number of samples in it. The algorithm is applied for each of the frames using random measurement matrix. The signals of each frame are reconstructed using orthogonal matching pursuit (OMP) algorithm. Each reconstructed frame is combined to obtain complete reconstructed signal.

Fixing of Sensing Matrix

To study compressive sensing and to fix the sensing matrix an experiment was conducted on a audio signal. A wave recorder was used to record a short piece of speech "Hello" of length $N = 16000$ as shown in figure 2. Further, a time domain speech signal is transformed into frequency domain by using Discrete Cosine Transform (DCT) technique. Signals of zero and lower amplitude do not contribute much to the signal. Hence thresholding is done with threshold window of -0.2 to $+0.2$.

Various experiments are conducted to fix this threshold window. This makes the transformed signal sparse which is a prerequisite for compressive sensing. Table 1 shows that of 16000 samples after thresholding only 2989 samples are used for reconstruction. In compressive sensing, instead of uniformly sampling the signal f the inner products of the signal measured against the set of measurement vectors $\phi_1, \phi_2, \phi_3, \dots, \phi_m$. When the no of measurements $M < N$, the signal f is efficiently compressed. Measurement vectors are taken as rows of the measurement matrix $\phi \in R^{M \times N}$. This vector is multiplied with the sparse representation of the signal to yield the observation vector $y = \phi f$ containing CS measurements. Now we aim to recover the signal f from fewest possible measurements y . Fixing the measurement matrix is a challenging task in CS.

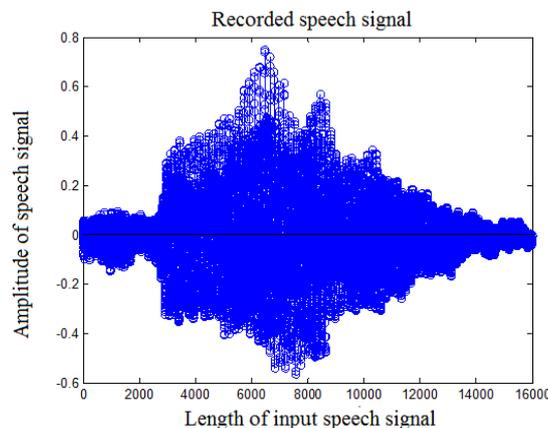


Fig2: Recorded speech signal

Various experiments are conducted on input speech signal “Hello”. For a fixed threshold of $+0.02$ to -0.02 the size of measurements are varied and performance parameters are measured. As shown in table 1 for 350 measurements i.e $M \times N = 350 \times 16000$ the mean square error is 0.0458, signal to noise ratio is 37.95 and the Perceptual Evaluation Speech Quality is 1.7/5. As the no of measurements are increased to 600 there is an improvement in these parameters. 21.66% of improvement in observed in the recovered signal. The recovered speech is as shown in fig. 3. For the measurements above 600 the reconstructed signal observed has the same SNR of 60.47. Hence the measurements matrix is fixed to $M \times N = 600 \times 16000$.

Table1: Performance analysis of signal *Hello* of length 16000

Length of signal (N)	Threshold Window	No of non zero Coefficients (T)	No of Measurements (M)	MSE	SNR	PESQ
16000	+0.02 to -0.02	2928	350	0.0458	37.95	1.7
16000	+0.02 to -0.02	2928	400	0.0391	39.08	1.9
16000	+0.02 to -0.02	2928	550	0.0291	40.34	2.7
16000	+0.02 to -0.02	2928	600	0.133	60.47	3.5

V. Conclusion

An audio signal “Hello” of length $N = 16000$ was considered for experimenting. This paper discusses fixing of sensing matrix. For faithful reconstruction of audio fixing the number of measurements M is vital. The experiment shows that for faithful reconstruction minimum number measurements required are 600. The performance of the recovered signal is measured using factors such as mean square error, signal to noise ratio and speech quality.

References

- [1] David L. Donoho. 2004. Compressive sensing. Department of statistics, Stanford University.
- [2] Emmanuel J. Candes and Michael B. Wakin, — “An Introduction to Compressive sampling.” *IEEE signal Processing Magazine*, Vol.25, Issue.2, pp.21-30, 2008.
- [3] Heung-No lee. 2011. Introduction to compressive sensing. Lecture notes. 26-29.
- [4] Ahmed Sabir. 2011. Compressive sensing for speech signals in mobile system. M.S. Thesis. The University of Texas.
- [5] Compressive Sensing in speech processing: “A survey based on sparsity and sensing matrix”. *IJETAE* vol. 3 issue 12. December-2013.
- [6] R.G.Moreno-Alvarado, Mauricio Martinez-Garcia “DCT-Compressive Sampling of frequency sparse audio signal”, *WCE* 2011.
- [7] Liban A. kassim and T. S. Gunawan. 2012 “Evaluation of sparsifying algorithm for speech signals”, International Conference on Computer and Communication Engineering (ICCCCE), Kuala Lumpur, pp.308.313,
- [8] J. Stanislaus and T. Mohsenin “High performance compressive sensing reconstruction hardware with QRD process”, in *Proc. ISCAS*, May 2012, pp. 2932.
- [9] Shwetha A. Gangannawar, Saroja V. S, “Compressed Sensing Reconstruction of an Audio Signal using OMP”, *International Journal of Advanced Computer Research (IJACR)*, Volume-5, Issue-18, March-2015.