The Adaptive and Non-Adaptive Compressed Sensing Based on Average-Frame Signal to Noise Ratio (AFSNR)

Prof. R. Sateesh

Assistant Professor
Electronics and Communication Engineering
MLR Institute of Technology,
Hyderabad, Telangana-500043, India

Abstract: The objective of this work is to further improve the performance of the existing compressed sensing technique that uses non adaptive projection matrix. Average-Frame Signal-to-Noise Ratio (AFSNR) is calculated to evaluate the performance of the frame-based adaptive Compressed Sensing (CS) with the non-adaptive Compressed Sensing (CS) is an emerging signal acquisition technique that directly collects signals in a compressed form if they are sparse on some certain basis. It originates from the idea that it is not necessary to invest a lot of power into observing the entries of a sparse signal in all coordinates when most of them are zero anyway. Our proposed approach is adaptive projection matrix based on frame analysis which gives significantly improved speech reconstruction quality.

Keywords: compressed sensing, Adaptive Sampling, Projection Matrix, AFSNR

1. Introduction

Compressive sensing is an emerging and revolutionary technology that strongly relies on the sparsity of the signal. The key objective in compressed sensing is to reconstruct a signal accurately and effectively from a set of few non-adaptive linear measurements. Of course, linear algebra easily shows that, in general, it is not possible to reconstruct an arbitrary signal from an incomplete set of linear measurements. Thus, one must restrict the domain in which the signals belong.

There are several important traits that an optimal recovery algorithm must possess. The algorithm needs to be fast, so that it can effectively recover signals in practice. Of course, minimal storage requirements as well would be ideal. The algorithm should provide uniform guarantees, meaning that given a specific method of acquiring linear measurements, the algorithm recovers all sparse signals. Ideally, the algorithm would require as few linear measurements as possible. However, recovery using only this property would require searching through the exponentially large set of all possible lower dimensional subspaces, and so in practice is not numerically feasible. This means that if the signal or its measurements are perturbed slightly, then the recovery should still be approximately accurate.

2. Existing system

Compressed sensing is the technique used to overcome the constraints of conventional sampling theorem. The compressed sensing allows us to go beyond the Nyquist rate and sample the signal below Nyquist frequency. On the other hand, compressive sampling provides a new way to reconstruct the original signal from a minimal number of observations. CS is a sampling paradigm that allows us to go beyond the Shannon limit by exploiting the sparsity structure of the signal. It allows us to capture and represent the compressible signals at a rate significantly below the Nyquist rate. The signal is then reconstructed from these projections by using different optimization techniques. During compressive sampling only the important information about a signal is acquired, rather than acquiring the important information plus the information of a signal which will be eventually discarded at the receiver. But the existing compressed sensing uses non adaptive projection matrix and takes the same number of projections for all the frames. This leads to degradation in system’s efficiency.

3. Proposed system

Compressed Sensing (CS) is a rising focus in recent years for its simultaneous sampling and compression of sparse signals. Speech signals can be considered approximately sparse or compressible in some domains for natural characteristics. Thus, it has great prospect to apply compressed sensing to speech signals. Most work in compressed sensing focus on random projection matrix which is constructed by considering only the signals’ sparsity rather than other properties. In other word, the construction of projection matrix is non-adaptive. Observing that different kind speech frames have different intra-frame correlations, a frame-based adaptive compressed sensing framework, which applies adaptive projection matrix for speech signals, has been proposed. To do so, neighbouring frames are compared to estimate their intra-frame correlation, every frame is classified into different categories, and the number of projections for each frame is adjusted accordingly. The experimental results show that the adaptive projection matrix can significantly improve the speech reconstruction quality.

4. Implementation of Adaptive Compressive Sensing

In conventional compressed sensing process, the projection matrix which is used to generate the required compressed signal is generated randomly and considered to be fixed during the entire conversion process. That means the projection matrix is non-adaptive. Though this process
results in better performance when compared to conventional sampling process, even better results can be obtained by using adaptive projection matrix.

4.1 Adaptive Projection Matrix

Most work in CS research focus on random projection matrix which is constructed by considering only the signals sparsity rather than other properties. In other word, the construction of projection matrix is non-adaptive. Observing that different kind speech frames have different intra-frame correlations, a frame-based adaptive compressed sensing framework for speech signals has been proposed, which applies adaptive projection matrix. To do so, the neighbouring frames are compared to estimate their intra-frame correlation, every frame is classified into different categories, and the number of projections is adjusted accordingly.

The experimental results show that the adaptive projection matrix can significantly improve the speech reconstruction quality. Intra-frame correlation of speech signals is explored to achieve efficient sampling. Because different kind speech signals may have different intra-frame correlations a frame-based adaptive CS framework that uses different sampling strategies in different kind speech frames, has been proposed.

4.2 Frame Analysis

Each speech sequence is divided into overlapping frames of size 1x n and all frames in a speech sequence are processed independently. The projection matrix is initialised by Gaussian random matrix $\Phi$ which has been proven to be incoherent with most sparse basis at high probability.

![Figure: Block diagram of Adaptive CS frame work for Speech](Image)

From Figure 1, for each frame in a speech sequence, a small number of projections is collected, and compared these projections with the projections collected for the previous frame. Based on the comparison results, the correlation between these two frames is estimated, and the correlation is classified into different categories. Then the sampling strategy is adjusted according to the correlation type and different numbers of samples are collected for the current frame.

4.3 Partial Sampling

For each frame in a speech sequence, a small number of projections are collected and compared with the projections collected for the previous frame. Based on the comparison results, the correlation between these two frames is estimated, and the correlation is classified into different categories. Then the sampling strategy is adjusted according to the correlation type and different numbers of samples are collected for the current frame.

4.4 Adaptive Sampling

Depending on their classified intra-frame correlation types, different number of projections is used for the speech frames. The frame is considered as surd frame if its intra-frame correlation type is surd vs. surd. A surd frame contains the least new information in the speech. Thus, the M0 measurements collected in the partial sampling stage are sufficient and additional sampling does not required. When its intra-frame correlation is sonant vs. sonant, the frame is considered as sonant and contains some new information, which requires more measurements to be collected.

For such frames, $M_1 (M_1 > M_0)$ measurements are collected. The $(M_0+1)^{th}$ to the $M_1^{th}$ rows of the Gaussian random matrix $\Phi$ are used and combined with $M_0$ to generate the final projection vector. The frames that experience large changes must contain the most new information. Therefore, a total of $M_2 (M_2 > M_1 > M_0)$ measurements are collected during the sampling process. The total projection matrix is the first $M_2$ rows of the Gaussian random matrix $\Phi$.

5. Results and discussion

To compare the performance of this proposed adaptive compressed sensing and the conventional non-adaptive CS, simulations are done. As a part of that, an arbitrary speech signal has been chosen Threshold values T1 and T2 are
chosen as 0.08 and 0.4 respectively which is tested through a great number of experiments. Average Frame Signal to Noise Ratio (AFSNR) is calculated and used to evaluate the reconstruction quality of speech signal. Average Frame Signal to Noise Ratio (AFSNR) is calculated using the formula

$$\text{AFSNR} = \frac{1}{K} \sum_{k=1}^{K} 10 \log_{10} (\frac{\| x_k \|}{\| x_k - x_k^1 \|})^2$$

Where K is the total frame number of a speech sequence x_k and x_k^1 represents the k\textsuperscript{th} frame speech and the k\textsuperscript{th} frame reconstructed speech. Under different compressed ratio(r = 0.2, r = 0.4 and r=0.6), which is defined as r = M/N, the different test results are obtained based on the proposed frame-based adaptive CS using OMP reconstruction algorithm.

**Input Signal**

The input signal which was already recorded is a combination of voiced and unvoiced sounds, and it lasts for seven seconds. Sampling frequency of the input signal is 8 kHz, and each sample is 8 bit encoded. So the bit rate of the signal is 64 kbps. Total no of samples present in the signal is equal to the product of sampling and total duration of the signal in seconds. Hence, total no of samples = 8000*7 =56,000. The input signal is undergone through Discrete Cosine transformation (DCT) technique which makes the signal sparse. The signal which is obtained by applying DCT to original input signal is still containing some low coefficients with low amplitude values. These coefficients have to be removed to make the signal further sparse. Hence the signal is undergone through a filtering process that eliminates the coefficients whose amplitude is below 0.04 and above -0.06.

**Frame Analysis**

The filtered signal is sparse enough to apply the compressed sensing. The entire signal is divided into 175 non overlapping frames in order to apply the adaptive compressed sensing. Each frame consists of 320 samples.

**Projection Matrix**

Projection matrix is a randomly generated matrix in which the no of columns is equal to N, and the no of rows is equal to M. Where N is equal to the no of samples present in the single frame, in this case the frame size is 320 and hence N=320. M is equal to r * N, where r is the compression ratio.

**For compression ratio r=0.2** N=320, r = M / N = 0.2. Hence M= r * N = 0.2 * 320 = 64

**For compression ratio r=0.4** N=320, r = M / N = 0.4. Hence M= r * N = 0.4 * 320 = 128

**For compression ratio r=0.6** N=320, r = M / N = 0.6. Hence M= r * N = 0.6 * 320 = 192

**Compressed Sensed Signal**

The compressed sensed signal is obtained by multiplying the input frame with the projection matrix corresponding to the required compression ratio. Here the compression ratio r is equal to 0.2 which indicates that the number of projections M = 64 for the frame of samples N = 320. The below figures shows the time domain waveform of the original speech signal and adaptive CS reconstructed speech with compressed ratio of 0.2. The average frame signal to noise ratio for compression ratio r = 0.2 with adaptive projection matrix is 12.030 which is 4.970 with non adaptive projection matrix. Thus the AFSNR can be increased by adaptive projection matrix.

**Case (i) for compression ratio r=0.2**

![Original speech signal & Reconstructed signal with compression ratio r = 0.2](image)

**Case (ii) for compression ratio r=0.4**

Here the compression ratio r is equal to 0.4 which indicates that the number of projections M=128 for the frame of samples N = 320. The below figure 3 shows the time domain waveform of the original speech signal and adaptive CS reconstructed speech with compressed ratio of 0.4. The average frame signal to noise ratio for compression ratio r=0.4 with adaptive projection matrix is 15.1714 which is 12.103 with non adaptive projection matrix. Thus the AFSNR can be increased by adaptive projection matrix. From the above figure 2, it is observed that, when the compressed ratio r = 0.4, the quality of the reconstructed signal has been increased than the reconstructed signal obtained from compressed ratio r = 0.2.

![Original speech signal & Reconstructed signal with compression ratio r = 0.4](image)
Case (iii) for compression ratio r=0.6

Here the compression ratio ‘r’ is equal to 0.6 which indicates that the number of projections M=192 for the frame of samples N = 320. The below figure 4 shows the time domain waveform of the original speech signal and adaptive CS reconstructed speech with compressed ratio of 0.6. The average frame signal to noise ratio for compression ratio r = 0.6 with adaptive projection matrix is 25.1721 which is 21.630 with non adaptive projection matrix. Thus the AFSNR can be increased by adaptive projection matrix. From the above figure 3, it is observed that, when the compressed ratio r ≤ 0.6, the quality of the reconstructed signal has been increased than the reconstructed signal obtained from compressed ratio r = 0.2 and r = 0.4.

Figure 4: Original signal & Reconstructed signal with compression ratio r = 0.6

Comparison of AFSNR for adaptive CS and non adaptive CS

The results obtained by using adaptive projection matrix to implement the compressed sensing are compared with that of conventional non adaptive compressed sensing, for the three compressed ratios r = 0.2, r=0.4, r= 0.6.

Table 1: AFSNR of the reconstructed speech using non-adaptive CS and the Adaptive CS

<table>
<thead>
<tr>
<th>Compression ratio</th>
<th>0.2</th>
<th>0.4</th>
<th>0.6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non Adaptive CS</td>
<td>5.00</td>
<td>13.00</td>
<td>22.00</td>
</tr>
<tr>
<td>Adaptive CS</td>
<td>2.9185</td>
<td>41.7593</td>
<td>43.6308</td>
</tr>
</tbody>
</table>

Thus it is clearly observed that the proposed adaptive compressed sensing is more efficient than the conventional non adaptive compressed sensing.

6. Conclusion

The adaptive projection matrix has been applied to the conventional compressed sensing and improved the average frame signal to noise ratio. It is also proved that the quality of the reconstructed signal increases as the compressed ratio increases. Thus the conventional non adaptive compressed sensing can be replaced by the adaptive compressed sensing to improve the efficiency of the system. 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 the speech signal below the Nyquist rate by taking only a few measurements of the signal.

7. References

