Characterization of reconstruction error in compressively sampled speech

Kenneth V. Domingo[*] and Maricor N. Soriano

National Institute of Physics, University of the Philippines, Diliman, Quezon City, Philippines
[*]Corresponding author: kdomingo@nip.upd.edu.ph

Abstract

Modern signal acquisition technologies are made possible by the Nyquist-Shannon sampling theorem (NST). However, this paradigm is extremely wasteful as the signal is compressed before storing it by systematically discarding the imperceptible information. Compressive sensing (CS) aims to directly sense the information that would otherwise survive this compression stage. Current literature focus exclusively on either audio or image signals, as well as quantifying their reconstruction quality with ambiguous metrics. In this paper, we compressively sample signals of arbitrary dimensions such as those consisting of combinations of audio and images, as well as quantify the reconstruction quality using metrics that are perceptually intuitive.

Keywords: compressive sensing, signal processing, compressive speech enhancement

1 Introduction

Conventional sensing devices are based on the Nyquist-Shannon sampling theorem (NST), which states that given a signal's bandwidth, one can capture all the pertinent information about that signal if it is sampled at a rate

$$f_S \geq 2f_B \tag{1}$$

where f_B is the signal's highest frequency component, and $2f_B$ is known as the Nyquist rate [1]. After this sampling process, the information is compressed by exploiting the signal's natural compressibility in some domain. For regular consumer-grade and even commercial-grade applications, this tried-and-true method of sampling and systematically discarding the imperceptible information works flawlessly. However, in certain situations when transmission bandwidth and/or storage comes at a premium, this process is highly wasteful [2]. Candés et al. [3] and Donoho et al. [4] independently pioneered compressive sensing (CS), which could directly sample the portions of a signal that would otherwise survive the compression stage in conventional sampling. In this new sampling paradigm, we consider the linear model of signal acquisition

$$\mathbf{y} = \mathbf{A}\mathbf{x} \equiv \langle \mathbf{x} | \mathbf{a}_k \rangle \tag{2}$$

In other words, the signal vector $\mathbf{x} \in \mathbb{R}^n$ is correlated with the basis waveforms $\mathbf{a}_k$ to yield a compressed information vector $\mathbf{y} \in \mathbb{R}^m$, where $m \ll n$. This causes the task of recovering $\mathbf{x}$ from $\mathbf{y}$ impossible since there exist an infinite number of candidate solutions $\hat{\mathbf{x}}$ which satisfy (2). This can be circumvented by enforcing constraints based on models of natural signals, as well as those based on optimization techniques. In particular, reconstruction can be achieved by enforcing sparsity and incoherence constraints [2]. From this, the process of reconstruction is then recast into a general minimization problem

$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{subject to} \quad \mathbf{A}\mathbf{x} = \mathbf{y} \tag{3}$$

where $\|\mathbf{x}\|_1$ denotes the ℓ_1 norm of $\mathbf{x}$.

2 Methodology
3 Results and Discussion
4 Conclusions

Here is the Summary or Conclusions section.

Proceedings of the Samahang Pisika ng Pilipinas
38th Samahang Pisika ng Pilipinas Physics Conference
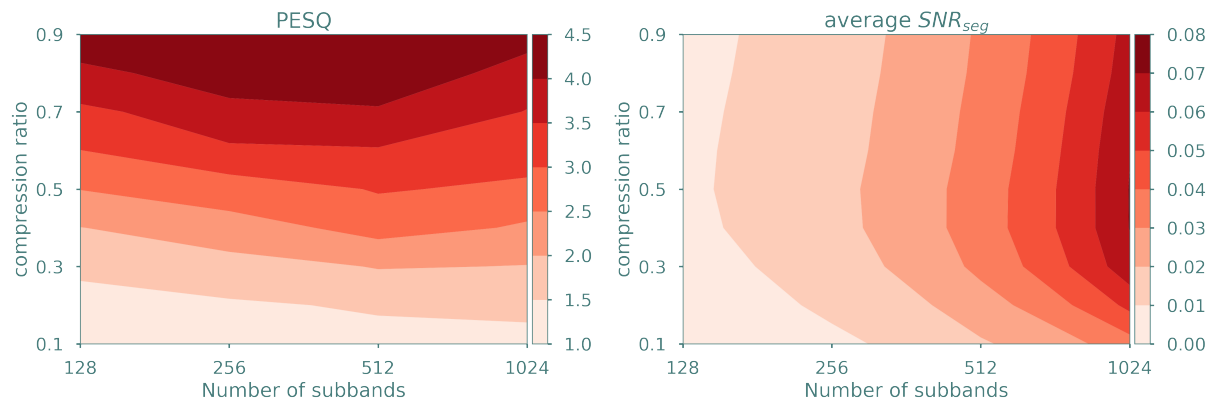Legazpi City, Albay, 3–6 June 2020
1

Figure 1: Two most commonly used metrics in evaluating quality of reconstructed speech recordings. PESQ is sensitive to the signal compression (left), while the average segmental SNR is sensitive to the number of subbands (right).

References

[1] C. E. Shannon, Communication in the presence of noise, *P. I.R.E.* **37**, 10 (1949).

[2] E. J. Candes and M. B. Wakin, An introduction to compressive sampling: A sensing/sampling paradigm that goes against the common knowledge in data acquisition, *IEEE Signal Proc. Mag.* **25**, 21 (2008).

[3] E. J. Candès, J. K. Romberg, and T. Tao, Stable signal recovery from incomplete and inaccurate measurements, *Commun. Pure Appl. Math.* **59**, 1207 (2006), arXiv:0503066v2.

[4] D. Donoho, M. Elad, and V. Temlyakov, Stable recovery of sparse overcomplete representations in the presence of noise, *IEEE Trans. Inf. Theory* **52**, 6 (2006).