UNIVERSITY OF THE PHILIPPINES

COMPRESSIVE SENSING: APPLICATIONS FROM 1-D TO N-D

By

KENNETH V. DOMINGO

A Baccalaureate Thesis Submitted to the

National Institute of Physics

College of Science

University of the Philippines

Diliman, Quezon City

In Partial Fulfillment of the Requirements

For the Degree of

B.S. Applied Physics (Instrumentation)

April 2020

Acknowledgments

Abstract

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The recent trend of curiosity-driven human development has caused a surge in the amount of openly accessible data. More often than not, the inflow of information into digital systems happens much faster than the system can process the data. Moore's law implicitly sets a limit to how powerful and how quick our electronic systems can become (barring a significant breakthrough in the field of quantum computing), and the Nyquist-Shannon sampling theorem (NST) limits the range of frequencies a certain device can successfully recover. This study explores the use of compressed sensing (CS)—an emergent sampling theorem that allows recovery of signals from much fewer samples than required by the NST—as a viable sampling method for signals with an arbitrary number of dimensions, for applications such as compression, encryption, and enhancement. In this framework, the computational burden is shifted from the sampling device to the device performing reconstruction/decompression, and as such, there exist many ways to recover a signal from compressive measurements. The use of CS has been applied to simple audio signals containing pure tones [1, 2] and speech [3–5], images [6–8], and videos [9, 10]. Common implementations of CS utilize analytical measurement bases such as the discrete cosine transform (DCT) basis, but recent studies [11–13] have shown that learned bases perform much better on more complex signals, i.e., those one would typically encounter in practical situations. The learning algorithms associated with these bases range from the classical iterative methods, such as matching pursuit (MP) and principal components analysis (PCA), to the more contemporary machine learning methods, most notably recurrent neural networks (RNN) and associative memory neural networks (AMNN). The novelty of this study is to provide a generalization of compressive sensing methods on signals of arbitrary dimensions.

1.1 Related literature

In 2006, Candès, Romberg, Tao [14], and Donoho [15] kicked off the field of compressive sensing by answering the following question: "With the recent breakthroughs in lossy compression technologies, we now know that most of the data we acquire can be thrown away with minimal perceptual loss. Why bother to acquire all the data when we can just directly measure the part that will not be thrown away?" The methods in CS apply concepts from time-frequency uncertainty principles [16] and sparse representations [17]. CS can be viewed as a random but strategic undersampling method, where the sampling rate can be associated with a quasi-frequency which is significantly lower than the Nyquist rate, and the random samples usually follow Gaussian or uniform distribution. [18] demonstrated the use of deterministic chaos filters to acquire samples instead of random distributions. The chaotic behavior of the sampling function naturally led to exploring the use of CS as an encryption algorithm. This was used to achieve simultaneous compression and encryption in [6], and was extended in [7] to utilize higher-dimensional chaotic systems. Sampling using chaotic maps, such Gaussian-Logistic map, were applied to acoustic signals in [1]. In the methods above, sampling was performed in the real domain (i.e., time domain for audio, spatial domain for images), and the reconstruction was performed in the frequency domain. [2] proposed a method to perform both sampling and reconstruction in the time domain using differential evolution.

Due to the relatively large size of video information, as a consequence of its high dimensionality, it is impractical to apply image CS techniques on an entire frame-by-frame basis. Correlations between adjacent frames are utilized instead, and can be obtained using dictionary learning [11] or PCA [9]. For the same reason, the application of CS to videos naturally led researchers to look towards machine learning methods [19]. [20] created the DR2-Net architecture which trains on grayscale video patches to reduce dimensionality.

The application of CS for signal denoising was explored in [21], and for recorded speech enhancement in [3]. Aside from the frequency domain, signals have been shown to be sparse in the modulation domain as well, which is a more appropriate representation for speech signals [4], or other signals whose frequency contents may vary rapidly in time.

1.2 Novelty

Chapter 2

Preliminaries

2.1 Compressed sensing

Consider a discrete time signal $\mathbf{x} \in \mathbb{R}^N$. Consider also an orthonormal basis $\mathbf{\Psi} \in \mathbb{R}^{N \times N}$ whose column vectors are expressed as $\boldsymbol{\psi}_i$. The signal $\mathbf{x}$ can be expressed as

$$\mathbf{x} = \langle \mathbf{\Psi} | \boldsymbol{\alpha} \rangle = \sum_{i=1}^{N} \alpha_i^\top \boldsymbol{\psi}_i \tag{2.1}$$

where α_i are the coefficients of $\mathbf{x}$ in the $\mathbf{\Psi}$ domain. For practical signal processing applications, a sampled signal is usually of lower dimension than its original representation. We define the compressed vector $\mathbf{y} \in \mathbb{R}^M$, $M < N$ which is obtained by

$$\mathbf{y} = \mathbf{A}\mathbf{x} \tag{2.2}$$

where $\mathbf{A} \in \mathbb{R}^{M \times N}$ $(M \leq N)$ is the sensing matrix. According to the NST, a periodic signal—which may be composed of a linear superposition of sinusoids with different parameters—has a characteristic bandwidth f_B, defined by its highest frequency component. The signal can be successfully reconstructed by a sampling device if the signal is sampled at a fixed rate with frequency f_s, which is at least twice f_B; that is $f_s \geq 2f_B$, where $2f_B$ is known as the Nyquist rate [22]. Under the NST, the original signal $\mathbf{x}$ in (2.2) can be recovered by inversion, least squares approximation, or more commonly, sinc interpolation. More often than not, the inversion of (2.2) is ill-posed because it is an underdetermined system. However, recovery is still possible provided that $\mathbf{x}$ is sparse or can be represented sparsely in some domain, and $\mathbf{A}$ satisfies one of the following properties [23]:

- **Low mutual coherence**: The mutual coherence of a matrix $\mathbf{A}$ is defined as

$$\mu(\mathbf{A}) = \max_{i \neq j} |\langle \mathbf{a}_i | \mathbf{a}_j \rangle| \tag{2.3}$$

- **Restricted isometry property** (RIP): Matrix $\mathbf{A}$ satisfies the RIP with sparsity parameter k and restricted isometry constant δ if for all k-sparse vectors $\mathbf{x}$,

$$(1-\delta)\|\mathbf{x}\|_2^2 \leq \|\mathbf{A}\mathbf{x}\|_2^2 \leq (1-\delta)\|\mathbf{x}\|_2^2 \tag{2.4}$$

  The sparsity parameter k entails that a vector has, at most, k nonzero coefficients, in which case the vector is said to be k-sparse [14].

Once these have been satisfied, $\mathbf{x}$ can be recovered exactly from the compressive measurements $\mathbf{y}$ by solving the combinatorial minimization problem

$$\min \|\mathbf{x}\|_0 \quad \text{subject to} \quad \mathbf{A}\mathbf{x} = \mathbf{y} \tag{2.5}$$

where $\|\mathbf{x}\|_0$ denotes the ℓ_0 pseudo-norm of $\mathbf{x}$, which extracts the number of nonzero coefficients. However, this problem is computationally intractable even for a small signal. A more tractable solution is to reduce this to a convex problem

$$\min \|\mathbf{x}\|_1 \quad \text{subject to} \quad \mathbf{A}\mathbf{x} = \mathbf{y} \tag{2.6}$$

where $\|\mathbf{x}\|_1 \equiv \sum_i |x_i|$ denotes the ℓ_1 norm of $\mathbf{x}$. For a sufficiently sparse signal, the solutions to (2.5) and (2.6) are identical [24]. A plethora of algorithms exist dedicated to solving (2.6). The scope of this study utilizes the following algorithms used commonly in the literature:

- **Least absolute shrinkage and selection operator** (LASSO): recasts the ℓ_1 minimization problem as an ℓ_1-regularized least squares problem:

$$\min_{\mathbf{x}} \frac{1}{2N}\|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \alpha\|\mathbf{x}\|_1 \tag{2.7}$$

  where α is the regularization hyperparameter. Optimization is performed via coordinate descent [25].

2.2 Associative memory networks

Chapter 3

Two-dimensional compressive sensing (2DCS)

Chapter 4

One-dimensional compressive sensing (1DCS)

9

Chapter 5

Generalization to N-dimensions

Chapter 6

Conclusions

Bibliography

[1] M. R. Mathew and B. Premanand, Sub-Nyquist Sampling of Acoustic Signals Based on Chaotic Compressed Sensing, *Procedia Technol.* **24**, 941 (2016), ISSN 22120173.

[2] I. Andráš, P. Dolinský, L. Michaeli, and J. Šaliga, A time domain reconstruction method of randomly sampled frequency sparse signal, *Meas. J. Int. Meas. Confed.* **127**, 68 (2018), ISSN 02632241.

[3] S. Y. Low, D. S. Pham, and S. Venkatesh, Compressive speech enhancement, *Speech Commun.* **55**, 757 (2013), ISSN 01676393.

[4] S. Y. Low, Compressive speech enhancement in the modulation domain, *Speech Commun.* **102**, 87 (2018), ISSN 01676393.

[5] V. Abrol, P. Sharma, and A. K. Sao, Voiced/nonvoiced detection in compressively sensed speech signals, *Speech Commun.* **72**, 194 (2015), ISSN 01676393.

[6] Y. Mo, A. Zhang, F. Zheng, and N. Zhou, An image compression-encryption algorithm based on 2-D compressive sensing, *J. Comput. Inf. Syst.* **9**, 10057 (2013), ISSN 15539105.

[7] N. Zhou, S. Pan, S. Cheng, and Z. Zhou, Image compression-encryption scheme based on hyper-chaotic system and 2D compressive sensing, *Opt. Laser Technol.* **82**, 121 (2016), ISSN 00303992.

[8] R. A. Romero, G. A. Tapang, and C. A. Saloma, in *Proceedings of the Samahang Pisika ng Pilipinas Physics Conference* (University of the Philippines Visayas, Iloilo City, Philippines, 2016), vol. 34, SPP–2016–PA–14.

[9] S. Liu, M. Gu, Q. Zhang, and B. Li, Principal component analysis algorithm in video compressed sensing, *Optik (Stuttg).* **125**, 1149 (2014), ISSN 00304026.

[10] J. Chen, K. X. Su, W. X. Wang, and C. D. Lan, Residual distributed compressive video sensing based on double side information, *Zidonghua Xuebao/Acta Autom. Sin.* **40**, 2316 (2014), ISSN 02544156.

[11] H. Liu, B. Song, H. Qin, and Z. Qiu, Dictionary learning based reconstruction for distributed compressed video sensing, *J. Vis. Commun. Image Represent.* **24**, 1232 (2013), ISSN 10473203.

[12] P. Sharma, V. Abrol, Nivedita, and A. K. Sao, Reducing footprint of unit selection based text-to-speech system using compressed sensing and sparse representation, *Comput. Speech Lang.* **52**, 191 (2018), ISSN 10958363.

[13] N. Eslahi, A. Aghagolzadeh, and S. M. H. Andargoli, Image/video compressive sensing recovery using joint adaptive sparsity measure, *Neurocomputing* **200**, 88 (2016), ISSN 18728286.

[14] E. J. Candès, J. K. Romberg, and T. Tao, Stable signal recovery from incomplete and inaccurate measurements, *Commun. Pure Appl. Math.* **59**, 1207 (2006), ISSN 00103640, arXiv:0503066v2.

[15] D. L. Donoho, Compressed sensing, *IEEE Trans. Inf. Theory* **52**, 1289 (2006), ISSN 00189448.

[16] D. Donoho and X. Huo, Uncertainty principles and ideal atomic decomposition, *IEEE Trans. Inf. Theory* **47**, 2845 (2001), ISSN 00189448.

[17] D. L. Donoho and M. Elad, Optimally sparse representation in general (nonorthogonal) dictionaries via L1 minimization, *Proc. Natl. Acad. Sci. U. S. A.* **100**, 2197 (2003), ISSN 00278424.

[18] N. Linh-Trung, D. Van Phong, Z. M. Hussain, H. T. Huynh, V. L. Morgan, and J. C. Gore, Compressed sensing using chaos filters, *Proc. 2008 Australas. Telecommun. Networks Appl. Conf. ATNAC 2008* 219–223 (2008).

[19] H. Yao, F. Dai, S. Zhang, Y. Zhang, Q. Tian, and C. Xu, DR2-Net: Deep Residual Reconstruction Network for image compressive sensing,

Neurocomputing **359**, 483 (2019), ISSN 18728286.

[20] M. Iliadis, L. Spinoulas, and A. K. Katsaggelos, Deep fully-connected networks for video compressive sensing, *Digit. Signal Process. A Rev. J.* **72**, 9 (2018), ISSN 10512004.

[21] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, Image denoising by sparse 3-D transform-domain collaborative filtering, *IEEE Trans. Image Process.* **16**, 2080 (2007), ISSN 10577149, arXiv:arXiv:1011.1669v3.

[22] C. E. Shannon, Communication in the presence of noise, *Proceedings of the Institute of Radio Engineers* **37**, 10 (1949).

[23] A. Mazumdar and A. S. Rawat, Associative memory via a sparse recovery model, *Adv. Neural Inf. Process. Syst.* 2701–2709 (2015), ISSN 10495258.

[24] E. J. Candès, J. Romberg, and T. Tao, Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information, *IEEE Trans. Inf. Theory* **52**, 489 (2006), ISSN 00189448, arXiv:0409186.

[25] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, et al., Scikit-learn: Machine learning in Python, *Journal of Machine Learning Research* **12**, 2825 (2011).