

List of changes

Deleted: The trend of both curiosity and profit-driven [...]	3
Replaced: ,	3
Deleted: —as a viable method for compression, encryp [...]	3
Deleted: Recent studies, however, have shown that [...]	4
Added: both asked and answered the questions that [...]	4
Deleted: CS can be viewed as a strategic undersampl [...]	4
Added: Linh-Trung et al.	5
Deleted: Sampling using a Gaussian-Logistic map was [...]	5
Replaced: initial value parameters	5
Deleted: of numbers	5
Added: In this study, a logistic map was used to [...]	5
Added: including those in this study,	5
Deleted: Audio signals, compared to images, have hi [...]	5
Replaced: greater than	5
Replaced: completely	5
Replaced: significant duration	6
Deleted: or recordings with relatively static frequencies	6
Deleted: i.e., the signal’s spectrogram,	6
Deleted: , and percent overlap between adjacent subbands	6
Added: This method was adopted in this study, and [...]	6
Added: , for applications such as compression, encr [...]	6
Replaced: Contemporary	6
Replaced: signals	6
Deleted: as the target for CS	6

Replaced: and	6
Deleted: of most of the research in the field has been to	6
Added: on constructing effective sensing matrices,	6
Replaced: ing	6
Deleted: signal	6
Deleted: to compressively sample signals	6
Deleted: , namely, one-dimensional CS (1DCS) and [...]	7
Added: for image and audio signals	7
Deleted: It is shown that an N -dimensional signal [...]	7
Added: Furthermore, current research tend to evalu [...]	7
Added: Finally, this study is conducted in order to [...]	7
Added: Chapter 3 establishes basic workflows and [...]	8
Replaced: 4 & 5	8
Replaced: image-based CS and audio-based CS	8

Chapter 1

Introduction

1 The trend of both curiosity and profit-driven human development has caused a
2 surge in the amount of openly accessible raw data. More often than not, the data
3 is generated much faster than it can be processed into something interpretable
4 or useful. In the endeavor of keeping up with the inflow of information, there
5 are two major factors that significantly hinder our progress. First, Moore's law
6 implicitly sets a physical limit to the number of transistors that can be placed on
7 a chip, consequently limiting how powerful and how fast electronic systems can
8 become (barring a paradigm shift in the fundamental design of semiconductors).
9 The second is the Nyquist-Shannon sampling theorem (NST), which limits the
10 range of frequencies a recording device can successfully capture.

11 This study explores the use of compressive sensing (CS),— an emergent
12 sampling theorem that allows reconstruction of signals from much fewer samples
13 than required by the Nyquist-Shannon sampling theorem (NST) —as a viable
14 method for compression, encryption, and/or enhancement. In this framework, the
15 computational burden of encoding/decoding a signal is shifted from the sampling

16 device to the device performing reconstruction, decompression, or other modes
17 of post-processing. As such, there exist many ways to reconstruct a signal from
18 compressive measurements.

19 CS has found its applications in simple audio signals containing stable
20 frequencies (such as pure tones [1, 2]) and dynamic frequencies (such as speech
21 [3–5]), images [6–8], and grayscale videos [9, 10]. The formulation of a sensing
22 matrix in CS requires a basis conforming to some uniform uncertainty principle,
23 and most common starting points would be partial discrete cosine transform (DCT)
24 matrices or partial discrete wavelet transform (DWT) matrices. ~~Recent studies,~~
25 ~~however, have shown that learned bases perform much better on more complex~~
26 ~~signals, i.e., those that would be typically encountered in real life situations. The~~
27 ~~learning algorithms associated with the construction of these bases range from~~
28 ~~classical iterative methods, which have long been used in optimization problems,~~
29 ~~to the more contemporary machine learning methods.~~

30 1.1 Related literature

31 In 2004, Candès, Romberg, Tao [11], and Donoho [12] ~~both asked and answered~~
32 ~~the questions that birthed~~ the field which we now know as compressive sensing.
33 The methods in CS apply concepts from time-frequency uncertainty principles [13]
34 and sparse representations, which were studied rigorously by Donoho and Elad [14].
35 ~~CS can be viewed as a strategic undersampling method: the signal is sampled at~~
36 ~~random points in the real domain, and the ratio of the indices where it is sampled~~
37 ~~to the size of the signal can be associated with some quasi-frequency which may~~
38 ~~be lower than the Nyquist rate.~~

39 Linh-Trung et al. [15] demonstrated the use of deterministic chaos
40 filters to acquire samples instead of random distributions. ~~Sampling using~~
41 ~~a Gaussian-Logistic map was applied to acoustic signals in [1].~~ Normally,
42 a deterministic chaotic function needs one or more initial value parameters
43 ~~initialization values as a “seed”,~~ and the sequence ~~of numbers~~ produced by different
44 combinations of initial values rapidly diverge from each other. This phenomenon
45 led to investigation of the use of CS as an encryption algorithm. Simultaneous
46 compression and encryption was achieved by [7], and it was found that the produced
47 sequences were sensitive to initial value perturbations on the order of 10^{-15} . Their
48 image compression-encryption model via CS was shown to have a key space on
49 the order of 10^{83} , making it extremely resistant to brute force and other types of
50 attacks. ~~In this study, a logistic map was used to encode and construct the sensing~~
51 ~~matrix. The encryption system exhibited similar key sensitivity and robustness~~
52 ~~characteristics mentioned in the former.~~ In the methods above, ~~including those in~~
53 ~~this study,~~ sampling was performed in the signal domain (i.e., temporal domain
54 for audio, spatial domain for images), and the reconstruction was performed in the
55 frequency domain.

56 ~~Audio signals, compared to images, have higher information density.~~
57 Whereas images are not naturally bandlimited and rather, are dependent on
58 the spatial resolution and bit depth of the imaging device, audio size scales
59 proportionally with time and takes on a wider range of values. The accepted
60 frequency range of human hearing is from 20 Hz to 20 kHz, so by the NST, a
61 sampling frequency ~~greater than at least~~ 40 kHz is needed to ensure that an audio
62 sample is recorded ~~completely correctly~~. Any meaningful audio recording, especially
63 those containing speech, will certainly have a ~~significant duration~~~~duration of a few~~

seconds up to a few hours, so one cannot straightforwardly apply methodologies used for images or recordings with relatively static frequencies. The first challenge this would pose for electronic systems is insufficient memory to process the entire signal all at once. Low circumvented this problem [3, 4] by transforming the signal to the modulation domain, i.e., the signal’s spectrogram, essentially raising a one-dimensional signal to N -dimensions, where N is dependent on the desired spectrogram resolution and number of subbands, and percent overlap between adjacent subbands. This method was adopted in this study, and additionally, each subband was multiplied with a window function to suppress potential boundary artifacts when reconstructing. In earlier experiments, reconstruction would often completely fail when windowing was not used. In the cases, however, that were successful, the reconstruction exhibited severe boundary artifacts in the form of distortion, aliasing, or noise.

1.2 Novelty

This study aims to provide a generalization for applying CS techniques to signals of arbitrary dimensions, for applications such as compression, encryption, and enhancement. Contemporary Previous CS research work exclusively on either audio or image signals sequences as the target for CS, and—due to the computational demands—the focus of most of the research in the field has been to on constructing effective sensing matrices, optimizing the computational complexity for real-time applications, and improving signal reconstruction quality. In the establishment of CS methods, two different general frameworks to compressively sample signals arise, namely, one-dimensional CS (1DCS) and two-dimensional

87 CS (2DCS) for image and audio signals. It is shown that an N -dimensional
88 signal can be decomposed into factors of one-dimensional and two-dimensional
89 signals, and can be processed using methods appropriate for each type of signal.
90 Furthermore, it is shown that N is bound not only by the type of signals being
91 worked with, but also the computational power of the decoding/decompressing
92 device. In particular, large values of N are useful in encryption, where a signal is
93 first raised to a high dimension in a certain basis, the sensing matrix is derived
94 from another high-dimensional basis, and the result is cast back to either one or
95 two dimensions to yield the encrypted message.

96 Furthermore, current research tend to evaluate signal reconstruction quality
97 using statistical metrics, such as mean-squared error (MSE) and its variants.
98 Arguably, the final interpreter of all signals are humans, and it is important to be
99 able to tell how well any compressive/reconstructive algorithm performs just by
100 looking at the metrics without directly observing the signal contents. In light of
101 this, the study also aims to evaluate the reconstruction quality of CS algorithms
102 using perceptually accurate metrics. This class of objective metrics are usually
103 built upon now-obsolete subjective scoring systems, and allows human observers
104 to make an informed estimate of the signal quality without directly accessing the
105 signal itself.

106 Finally, this study is conducted in order to lay out a unified, standardized
107 workflow for similar applications of CS on signals with arbitrary content. This
108 includes signals containing a combination of audio and images, such as color videos
109 and hyperspectral images.

110 **1.3 Thesis overview**

111 The next chapter establishes the relevant mathematical concepts and notation
112 to be used throughout this study, algorithms used in signal reconstruction, and
113 appropriate metrics per type of signal. Chapter 3 establishes basic workflows
114 and studies the effect of random sampling on CS reconstruction. Chapters 4 & 5
115 ~~3-5~~ respectively focus on image-based CS and audio-based CS~~two-dimensional~~
116 ~~CS, one-dimensional CS, and N -dimensional CS~~. Each of these chapters are
117 self-contained methodologies, results, and discussions to emphasize that the
118 methods can work independently of each other. Conclusions of the study and
119 recommendations for future studies are presented in Chapter 6.