

# Chapter 2

## Preliminaries

The trend of both curiosity and profit-driven human development has caused a surge in the amount of openly accessible raw data. More often than not, the data is generated much faster than it can be processed into something interpretable or useful. In the endeavor of keeping up with the inflow of information, there are two major factors that significantly hinder our progress. First, Moore’s law implicitly sets a physical limit to the number of transistors that can be placed on a chip, consequently limiting how powerful and how fast electronic systems can become (barring a paradigm shift in the fundamental design of semiconductors). The second is the Nyquist-Shannon sampling theorem (NST), which limits the range of frequencies a recording device can successfully capture. This states that given that you know a signal’s highest frequency component  $f_B$ , sampling it at a rate  $f_S$  that is at least twice this frequency is sufficient to capture all of the pertinent information regarding that signal: that is  $f_S \geq 2f_B$ , where  $f_B$  is known as the Nyquist rate [25], and also as the signal bandwidth. For signals that are not naturally bandlimited, such as images, the ability reproduce a signal is dependent on the device’s resolution

---

and still follows the same principle: there should be at least twice the number of pixels codimensional with the image's highest spatial frequency. For practical day-to-day use, the NST covers everything you need. However, issues arise when bandwidth and storage are at a premium. Typically, after sensing a signal, not all of the raw data is stored. Rather, this data is converted to a compressed format by systematically discarding values such that the loss of information is virtually imperceptible. Thus, the process of acquiring massive amounts of data followed by compression is extremely wasteful. CS aims to directly acquire the parts of the signal that would otherwise survive this compression stage in the classical sampling scheme.

Consider a signal  $\mathbf{x} \in \mathbb{R}^n$ ; this notation indicates that  $\mathbf{x}$  is real-valued (but can easily be extended to complex-valued signals), one-dimensional vector of size  $n$ . The process of acquisition or sensing this signal can be modeled as a linear system, where the physical signal properties we wish to capture are transformed into digital values by applying a linear transformation

$$\mathbf{y} = \mathbf{A}\mathbf{x} \tag{2.1}$$

or in the literature of signal processing [26], by correlating them with a waveform basis

$$y_k = \langle \mathbf{x} | \mathbf{a}_k \rangle, \quad k \in \mathbb{N} \leq n \tag{2.2}$$

In conventional sampling,  $\mathbf{a}_k$  are Dirac basis vectors which turn  $\mathbf{y}$  into a vector containing samples of  $\mathbf{x}$  in the temporal or spatial domain; if  $\mathbf{a}_k$  are Fourier basis vectors (i.e., sinusoids), then  $\mathbf{y}$  is a vector of Fourier coefficients. If the signal

has been sampled sufficiently in the sense that the number of measurements  $m$  is equal to the dimension  $n$  of the signal, then  $\mathbf{A}$  is a square matrix, and the original signal  $\mathbf{x}$  can be reconstructed from the information vector  $\mathbf{y}$  by inversion of (2.1). However, the process of recovering  $\mathbf{x} \in \mathbb{R}^n$  from  $\mathbf{y} \in \mathbb{R}^m$  becomes ill-posed when we consider the undersampled case ( $m \ll n$ ), as the sensing matrix  $\mathbf{A} \in \mathbb{R}^{m \times n}$ —whose row vectors are denoted as  $\mathbf{a}_m$ —causes the system to become underdetermined: there exist infinitely many candidate solutions  $\hat{\mathbf{x}}$  which satisfy (2.1). To add to this, we also consider the possibility that the measurements are not perfect, and are contaminated with noise. How then do we recover a signal from measurements which are incomplete and most likely inaccurate? The answer lies in enforcing constraints based on models of natural signals, as well as constraints based on optimization techniques.

## 2.1 Sparsity

Most natural signals, especially those with some underlying periodicity, can be represented sparsely when expressed in the appropriate basis. This process of “sparsifying” can be expressed as

$$f = \langle \mathbf{x} | \boldsymbol{\psi}(k) \rangle \quad (2.3)$$

Similar to (2.2), this involves correlating the signal with the appropriate basis function to yield a representation in the sparse domain. Image information, for example, are commonly expressed in the DCT domain by

$$f_k = \sum_{n=0}^{N-1} x_n \cos \left[ \frac{\pi}{N} \left( n + \frac{1}{2} \right) k \right], \quad 0 \leq k < N \quad (2.4)$$

and its corresponding inverse is

$$x_k = \frac{1}{2}f_0 + \sum_{n=1}^{N-1} f_n \cos \left[ \frac{\pi}{N} \left( k + \frac{1}{2} \right) n \right], \quad 0 \leq k < N \quad (2.5)$$

where the cosine term corresponds to  $\psi(k)$  in (2.3). We can express (2.4) more conveniently as  $\mathbf{f} = \mathbf{\Psi}\mathbf{x}$ , where  $\mathbf{\Psi} \in \mathbb{R}^{n \times n}$  is the sparsifying matrix. Figure 2.1 shows this sparsifying process in action: given a test image, taking its DCT and zooming into the first 20,000 coefficients show that most of the signal energy is concentrated in just a few of the coefficients close to the DC term and decay rapidly. All the other coefficients, when compared to the  $k$  highest coefficients, are practically zero; such a signal is referred to as  $k$ -sparse. The compressed image resulting from discarding all but the first 25,000 coefficients and performing the inverse transform shows that any difference from the original image is virtually imperceptible. A similar concept is used in JPEG compression, wherein an image is divided into  $8 \times 8$  blocks, and in each block, a certain number of DCT coefficients are discarded depending on the desired quality factor  $Q$  [27].

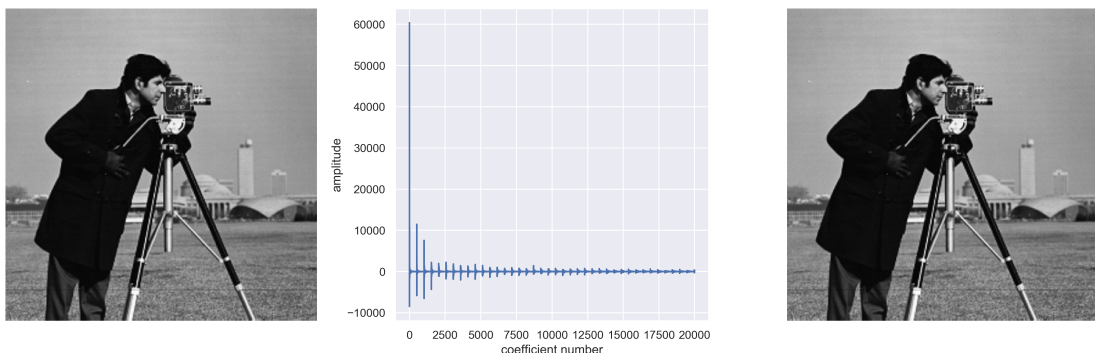


Figure 2.1: Original  $512 \times 512$ , 8-bit image (left), and its first 20,000 DCT coefficients (middle). Most of the signal energy is concentrated in just a few of the coefficients that are close to the origin. By discarding all but the first 25,000 coefficients and performing the inverse transform, the resulting image (right) is perceptually no different from the original.

## 2.2 Incoherence

Suppose we have two matrices  $\Phi$  and  $\Psi$  which are involved in the sensing of a signal. As before,  $\Psi$  is the sparsifying matrix which converts the signal into a sparse representation, and  $\Phi$  is the actual sensing matrix. The coherence between these two bases is expressed as

$$\mu(\Phi, \Psi) = \sqrt{n} \max_{0 \leq i, j < n} |\langle \varphi_i | \psi_j \rangle| \quad (2.6)$$

In other words, the coherence is the measure of the largest correlation between the column vectors of  $\Phi$  and  $\Psi$ . In compressive sensing, we are interested in low-coherence basis pairs (i.e., basis pairs for which  $\mu \rightarrow 1$ ). For example, in the classical sampling scheme,  $\Phi$  is the Dirac basis  $\varphi_k(t) = \delta(t - k)$ , and  $\Psi$  is the DCT basis (2.4). This basis pair in particular is also called a time-frequency pair and achieves maximum incoherence ( $\mu = 1$ ) regardless of the number of dimensions [16]. Additionally, any orthonormal basis  $\Phi$  containing independent identically distributed (i.i.d.) entries are also largely incoherent with a fixed basis  $\Psi$  [26]. The consequence of this is that CS performs most efficiently when sensing with incoherent and random systems.

## 2.3 Reconstruction strategies

Another measure of the sparsity of a signal is its  $\ell_0$  pseudo-norm, denoted  $\|\mathbf{x}\|_0$ , which simply counts the number of non-zero coefficients of  $\mathbf{x}$ . As such, the goal of the reconstruction stage in CS is to find the sparsest representation of the vector  $\mathbf{x}$  in terms of the sensing matrix  $\Phi$  by solving the combinatorial optimization problem

$$\min_{\mathbf{x}} \|\mathbf{x}\|_0 \quad \text{subject to} \quad \mathbf{y} = \Phi \mathbf{x} \quad (2.7)$$

which, as the name implies, requires one to enumerate all possible  $k$ -element combinations of the columns of  $\Phi$ , and determining the smallest combination which approximates the signal the closest. However, this process quickly becomes intractable even for a modestly-sized signal. This requirement is therefore relaxed by instead minimizing the  $\ell_1$  norm

$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{subject to} \quad \mathbf{y} = \Phi \mathbf{x} \quad (2.8)$$

where the  $\ell_1$  norm is defined as

$$\|\mathbf{x}\|_1 = \sum_{i=0}^{N-1} |x_i| \quad (2.9)$$

and is commonly called the taxicab or Manhattan distance. Aside from  $\ell_1$  minimization, other approximation methods include greedy approaches (e.g., basis pursuit), and convex programming.