See from a Pixel

Group2
Lorenzo Lu
ylu289

1 Introduction

Human civilization, used to be driven by the power of steam engine and electricity, now casts loads of attention onto communication and signal processing. As a milestone in the history of Signal Processing, Nyquist-Shannon Theorem labels a limit for engineers, even though it does seem to be possible to exceed. **NS-Theorem: If a function $x(t)$ contains no frequencies higher than B hertz, it is completely determined by giving its ordinates at a series of points spaced $\frac{1}{2B}$ seconds apart (Source: WikiPedia)**.

However, signals and images are compressible in most cases. Images, after discrete cosine transformation (DCT), will concentrate the intensity at the low frequency region. If only the significant and principal coefficients are kept, this image will be stored in JEPG format. The "sparse" property of images inspired people to seek a method with which we can directly measure these principal coefficients instead of measuring the whole image with compression coming later. Because fewer measurements are taken, this technology is thus called **compressed sensing (CS)**.

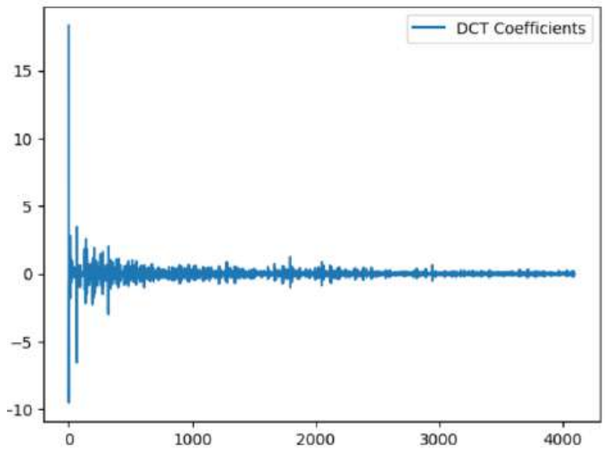

Figure 1: Typical image DCT coefficients distribution

Usually, a CS setup will require a tool to generate 0-1 masks, together with a single pixel sensor. DLP 7000 is a digital mirror device (DMD) produced by Texas Instruments which can produce 0-1 masks. This special device consists of 1024×768 mirror chips which can rotate to either $+12$ or -12 degrees. If a beam is cast onto the DMD, it will be split and reflected to two different directions. A single pixel sensor set at the focus of a lens can be adopted to collect the total photons in either direction. This experiment setup is not accessible due to the COVID-19, so we have to switch our work onto a 100% simulation.

2 Model

Imagine we can find m 0-1 masks where each two of them are not the same, and we use these masks measuring the same view, so every time we sum of photons from different parts of it. The measured intensity is notated by y, and ϕ, x for the mask and image respectively.

$$y = \sum_{i,j} \phi_{i,j} x_{i,j} \tag{1}$$

Reshaping the mask and image into a vector, we thus have

$$\boldsymbol{y}_{m\times 1} = \boldsymbol{\phi}_{m\times n} \boldsymbol{x}_{n\times 1} \tag{2}$$

The subscripts stand for the reshape of these matrices. m is the length of the mask sequences, which is the number of measurements we take. n is the number of pixels.

The image x can also be represented by matrices product.

$$\boldsymbol{x}_{n\times 1} = \boldsymbol{\psi}_{n\times n} \boldsymbol{a}_{n\times 1} \tag{3}$$

If $\boldsymbol{\psi}$ is DCT transform, $\boldsymbol{a}$ is the DCT coefficients. Similarly, if $\boldsymbol{\psi}$ is Haar wavelets transform, $\boldsymbol{a}$ is the wavelets coefficients. Consequently, this model is set to solve this following optimization problem

$$\underset{\boldsymbol{a}}{\text{Argmin}} \quad (\boldsymbol{y} - \boldsymbol{\phi}\boldsymbol{\psi}\boldsymbol{a})^2 + \lambda \|\boldsymbol{a}\|_1 \tag{4}$$

The λ is the optimizer. If $\boldsymbol{a}$ is sparse, the l_1 regularization will have an excellent performance when the matrix ϕ is underdetermined and $n \gg m$.

For single pixel imaging system, Compressed Sensing, shows a great signal to noise (Gaussian noise) ratio (SNR) compared with Raster Scan (RS,scanning one pixel by one pixel), as CS sums intensity over half of the object during each measurement. Furthermore, CS saves time when compared with Basis Scan (BS) which requires a full rank ϕ. These advantages render CS a very charming imaging strategy in this era.

Meanwhile, an interesting question appears when people want to figure out if compressed sensing still robust for photon counting system, where the Poisson noise rather than Gaussian noise plays a role. In this case, the measured intensity is $\hat{\boldsymbol{y}} \sim \mathcal{P}(X = \boldsymbol{y})$. Poisson noise is similar to Gaussian noise at large mean values, but very different for smaller mean values because Poisson noise only generates integers. Also, Poisson noise is not addable, which is very different from Gaussian noise.

Another fact not favoring CS in photon counting is that the diagonalization of a matrix includes not only summation but also subtraction. Assuming we have $x_1 \sim \mathcal{P}(\mu_1)$ and $x_2 \sim \mathcal{P}(\mu_2)$, we have $x_1 + x_2 \sim \mathcal{P}(\mu_1 + \mu_2)$, but $x_1 - x_2$, which obeys Skellam distribution, has the mean equal to $\mu_1 - \mu_2$ but the variance equal to $\mu_1 + \mu_2$. SNR can be represented by $\frac{\mu}{\sigma}$, where μ and σ are the mean and standard deviation respectively. If a matrix subtraction happens during the diagonalization, the SNR will be worse than it supposed to be if we do a Raster Scan.

Therefore, the optimization is exactly written in this form

$$\underset{\boldsymbol{a}}{\text{Argmin}} \quad (\hat{\boldsymbol{y}} - \boldsymbol{\phi}\boldsymbol{\psi}\boldsymbol{a})^2 + \lambda \|\boldsymbol{a}\|_1 \tag{5}$$

3 Data

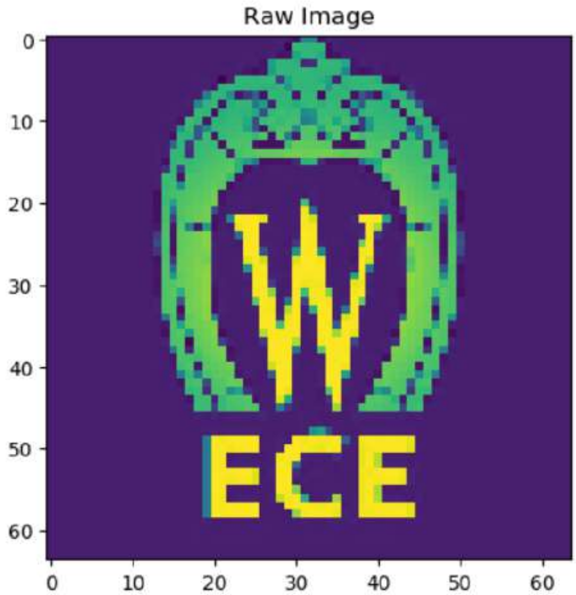

Figure 2: Data for simulation

4 Objectives

Here we want to design three projects to investigate is CS a good strategy for photon counting system.

1. (Primary) Assuming our photon counting system will be saturated when 10^6 photons arrives within one second. And our exposure time is, for instance, 100 seconds. And this constraints will only impact the $\hat{\boldsymbol{y}}$. The objective is still to find the best reconstruction of this image.

2. The coherence between ϕ and ψ can decide how many measurements we take for this image reconstruction. Researchers claims Noiselets and Wavelets has very low coherence. However, Noiselets matrix contains both real and imaginary parts, which may not be easy for us to implement our lasso regression problem. It would be interesting if we can find a binary matrix as ϕ to minimize the coherence between it and ψ. The coherence between ϕ and ψ is $\max\limits_{1\leq i,j\leq n} \langle \phi_i | \psi_j \rangle$.

$$\text{variables.} \quad \phi_{m\times n} \tag{6}$$
$$\text{s.t.} \quad \phi_{i,j} \in \{0,1\} \tag{7}$$
$$\text{rank}(\phi) = n \tag{8}$$
$$\min \max_{1\leq i,j\leq n} \langle \phi_i | \psi_j \rangle \tag{9}$$

   This question may not be easy to solve, as it might be very time consuming. If we cannot figure out a good approach to solve this, we may just try with three combinations of ϕ and ψ and find the best one.

   (a) ϕ is a random 0-1 matrix and ψ is DCT/Haar basis.

   (b) ϕ is 0-1 Walsh matrix and ψ is DCT/Haar basis.

   (c) ϕ is 0-1 Walsh matrix but has priority for the low frequency basis and ψ is DCT/Haar basis. I am still skeptical is this strategy can be called as Compressed Sensing, which should choose the ϕ basis randomly. But as long as it can give good reconstruction, we can try to explain why this works.

3. As an expectation, the longer measurement time renders a better reconstruction under Poisson noise, but it seems to be impossible that we reconstruct an image with infinitely long time in order to achieve the minimum errors, ϵ. On the other hand, we would like to use less time for measurements.

$$\text{variables.} \quad t \geq 0 \tag{10}$$

$$\text{s.t.} \quad \int_{\tau}^{\tau+1} n(t)dt \leq 10^6 \quad \forall \tau \in \mathbb{R} \tag{11}$$

$$\min \quad \epsilon(t) + \lambda t \tag{12}$$

The constraint stands for the saturation boundary of the sensor. Again, the λ is the optimizer. This optimization problem is hard to formulate, and seems to be time and memory consuming. If so, we will switch to find a numerical answer.

4. Compare strategies such as RS BS, and CS to see which one has the best performance. If CS is not the best, try to find an approach to improve it.

5 Early outcomes and analysis

An unexpected problem with memory appeared, so that original objective function should be modified in the following form

$$\min \quad \|\boldsymbol{a}\|_1 \tag{13}$$

$$\text{s.t.} \quad (\boldsymbol{\phi}\boldsymbol{\psi}\boldsymbol{a} - \hat{\boldsymbol{y}})^2 \leq \epsilon \tag{14}$$

The first reconstruction without noise is like

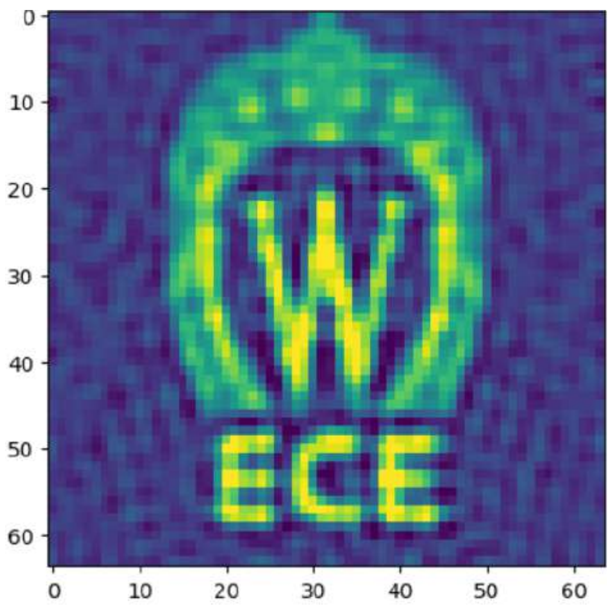

Figure 3: Walsh and DCT basis, 64×64, measurements/pixels=0.25

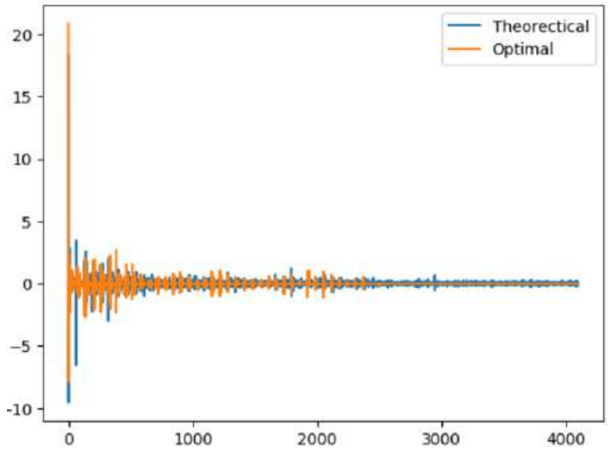

Figure 4: The reconstructed DCT coefficients

From this result we can see the sparse coefficients are reconstructed with most peaks at correct positions. The amplitudes now are not satisfying, and we may compare gradient descent method with JuMP optimizer later.

Later we will add noise during the virtual measuring process, and we will show the objectives posted above in the final report.