AMATH 482 HOME WORK 5
COMPRESSED IMAGE RECOVERY

LANG QIN

Amazing Department, University of Washington, Seattle, WA
lq25@uw.edu

Abstract. The goal of this work is to recover an image from limited observations of its pixels using convex optimization. This work consists of three tasks: 1) image compression, 2) compressed image recovery, and 3) mysterious image reconstruction.

1. Introduction and Overview

This work primaries plays with two image inputs: René Magritte's "The Son of Man" of size and a mysterious image of size 50×50. In order to reduce run time complexity, this work uses rescaled "The Son of Man" of size 53×41. The goal of this work is to 1) investigate the compressibility of the discrete cosine transform (DCT) of "The Son of Man", 2) recover randomly corrupted "The Son of Man" via convex optimization, and 3) reconstruct the mysterious image from limited and indirect measurements using DCT and convex optimization.

2. Theoretical Background

To achieve the goal, this work primarily relies on the following two points: 1) image compression with discrete cosine optimization and 2) compressed image recovery by discrete cosine transformation with convex optimization.

2.1. Image Compression.

This section introduces discrete cosine transform and its inverse function that is used to investigate the compresibility of discrete cosine transform (DCT) of an image using various DCT coefficients.

2.1.1. *Discrete Cosine Transform (DCT).*

Given a discrete signal $\mathbf{f} \in \mathbb{R}^K$, the discrete cosine transformation (DCT) is defined as $\mathrm{DCT}(\mathbf{f}) \in \mathbb{R}^K$ [1], where

$$\mathrm{DCT}(\mathbf{f})_k = \sqrt{\frac{1}{K}}\left[f_0 \cos\left(\frac{\pi k}{2K}\right) + \sqrt{2}\sum_{j=1}^{K-1} f_j \cos\left(\frac{\pi k(2j+1)}{2K}\right)\right]. \tag{1}$$

This transformation is analogous to taking the real part of the FFT of $\mathbf{f}$. The inverse discrete cosine transform (iDCT) reconstructs the signal $\mathbf{f}$ from $\mathrm{DCT}(f)$, i.e. $\mathrm{iDCT}(\mathrm{DCT}(\mathbf{f})) = \mathbf{f}$. The iDCT of the DCT defined by Eq. 1 is defined as [1]

$$\mathrm{iDCT}(\mathbf{f})_k = \sqrt{\frac{1}{K}}\left[f_0 + \sqrt{2}\sum_{j=1}^{K-1} f_j \cos\left(\frac{\pi j(2k+1)}{2K}\right)\right]. \tag{2}$$

Date: March 17, 2022.

2.1.2. *Image Compression with DCT.* Let N_y, N_x denote the number of rows and columns (pixels) of the image with $N = N_y \times N_x$ denoting the number of pixels in the image. Let the discrete signal $F \in \mathbb{R}^{N_y \times N_x}$ denote the image and vectorize the 2D signal F as $\mathbf{f} \in \mathbb{R}^N$ by stacking the pixel values into a long vector. Then, one may obtain the DCT of the image by $\mathrm{DCT}(\mathbf{f})$ and reconstruct the transformed signal by $\mathrm{iDCT}(\mathrm{DCT}(\mathbf{f}))$ as defined in Section 2.1.1.

Given a compression threshold θ, one may compute the DCT coefficient at $(1-\theta)$ percentile c_θ and reset the DCT by

$$\widehat{\mathrm{DCT}}(\mathbf{f})_k = \begin{cases} 0 & (\mathrm{DCT}(\mathbf{f})_k < c_\theta) \\ \mathrm{DCT}(\mathbf{f})_k & \text{otherwis} \end{cases}. \tag{3}$$

Then, the compressed image signal with threshold of θ is obtained by

$$\hat{\mathbf{f}} = \mathrm{iDCT}(\widehat{\mathrm{DCT}}(\mathbf{f})). \tag{4}$$

2.2. **Compressed Image Recovery.**

2.2.1. *Signal Recovery Problem.* The signal recovery problem takes the following form. Say $f^+ : [0,1] \to \mathbb{R}$ is a signal and the measurements $\underline{\mathbf{y}} \in \mathbb{R}^M$ with $M < N$ are of the form

$$y_j^+ = \underline{\xi}_j^T \underline{f}^+, \tag{5}$$

where $\underline{\xi} \sim N(0, I)$ and $\underline{\mathbf{f}}^+ = f^+(t_0), f^+(t_1), \ldots, f^+(t_{N-1}))$ for some points $t_j \in [0,1]$. The goal of this problem is to reconstruct $\mathbf{f}^+$ from given $\underline{\mathbf{y}}^+$.

2.2.2. *Solution to Signal Recovery Problem.* In order to solve the problem defined in Section 2.2.1, the first step is to assume a model for $\hat{f}$ by

$$f(t_n) = \sum_{m,n} \beta_{m,n} \psi_{m,n}(t_n), \tag{6}$$

where $\phi_{m,n}$-s are the Haar wavelet basis. Re-parametrizing the wavelet coefficients into a 1D array gives

$$f(t_n) = \sum_{j=0}^{J-1} \beta_j \psi_j(t_n). \tag{7}$$

Then, one may rewrite the signals $\underline{\mathbf{f}}^+ = f^+(t_0), f^+(t_1), \ldots, f^+(t_{N-1}))$ by $\underline{\mathbf{f}} = W\underline{\beta}$, where $W \in \mathbb{R}^{N \times J}$ and W takes the following form

$$W = \begin{bmatrix} \psi_0(t_0) & \psi_1(t_0) & \ldots & \psi_{J-1}(t_0) \\ \psi_0(t_1) & \phi_1(t_1) & \ldots & \psi_{J-1}(t_1) \\ \vdots & \vdots & & \vdots \\ \psi_0(t_{N-1}) & \psi_1(t_{N-1}) & \ldots & \psi_{J-1}(t_{N-1}) \end{bmatrix}. \tag{8}$$

Thus, for such an $\underline{\mathbf{f}}$, its measurements $\underline{\mathbf{y}}$ can be represented by

$$\underline{\mathbf{y}} = S\underline{\mathbf{f}},\ S = \begin{bmatrix} \underline{\xi}_0^T \\ \underline{\xi}_1^T \\ \vdots \\ \underline{\xi}_{M-1}^T \end{bmatrix} \in \mathbb{R}^{M \times N} \Longrightarrow \underline{\mathbf{y}} = A\underline{\beta}, \text{ where } A = SW. \tag{9}$$

As a result, this signal recovery problem is basically to solve the supervised learning/regression problem of

$$\text{find } \hat{\beta} \text{ s.t. } A\underline{\hat{\beta}} \approx \underline{\mathbf{y}}. \tag{10}$$

2.3. **Sparse Recovery.** One important property of the problem defined in Section 2.2.1 is the sparsity of $\mathrm{DCT}(\mathbf{f})$. Thus, to maintain the sparsity, there must be an addition constraint to the solution of problem defined by Eq. 10: $\hat{\beta}$ is sparse. Otherwise, the resulted reconstruction would no longer takes the form of some DCT and the iDCT of the solution will not give a proper image. Thus, lasso is introduced to solve the regression problem with sparsity.

2.3.1. *Lasso.* Regressions with Lasso are similar to regressions such as linear least squares in the way that they approaches the true value using 2-norms for error evaluation. The only difference takes place at the penalization part that Lasso minimizes the 1-norm of β, i.e. $\lambda||\beta||_1$, instead of 2-norm, i.e. $\lambda||\beta||_2^2$, which is often used in ridge regression. This is the key to ensure the sparsity of $\hat{\beta}$. Thus, the regression problem defined as Eq. 10 becomes the optimization problem as follows.

$$\begin{aligned} \min_{\beta \in \mathbb{R}^N} \quad & ||\beta||_1 \\ \text{s.t.} \quad & A\beta = \mathbf{y} \end{aligned} \tag{11}$$

Clearly, the result optimal vector β^* is the DCT vector of an image F^* that resembles the original image F.

3. Algorithm Implementation and Development

The implementation of this project can be divided into three parts: 1) Environment Setup, 2) Image Compression, 3) Compressed Image Recovery, and 4) Mysterious Image Reconstruction. This programming work is done by Google Colab. The `Python` libraries imported for this work include: `google.colab.drive`, `numpy`, `matplotlib.pyplot`, `scipy`, `cvxpy`, and `skimage`.

3.1. **Environment Setup.** Import image from `SonOfMan.png`, convert to grayscale, and resize the image by 53×41 as F. Import information of the mysterious image from `UnknownImage.npz` and extract the measurement matrix $B_i \in \mathbb{R}^{2000\times 2500}$ and vector of measurement $\mathbf{y}_i \in \mathbb{R}^{2000}$.

3.2. **Image Compression.** This step is to use DCT to compress the image by using the n-th largest DCT coefficients and reconstruct the compressed image via iDCT. The algorithm of this task is as follows.

i. As described in Section 2.1.2, the 2D pixel matrix F need to be converted into a 1D pixel vector $\mathbf{f} \in \mathbb{R}^{53\times 41}$;
ii. Initialize the forward and backward DCT matrices $D, D^{-1} \in \mathbb{R}^N$ defined in Section 2.1.1 using the given helper functions;
iii. Derive the DCT of the flattened image signals $\text{DCT}(\mathbf{f}) = DF$ and plot the absolute values of $\text{DCT}(\mathbf{f})$;
iv. For each threshold $\theta = 5, 10, 20, 40$,
   (a) Find $(1-\theta)$ percentile largest coefficient of $\text{DCT}(\mathbf{f})$ in absolute form and obtain $\widehat{\text{DCT}}(\mathbf{f})$ as defined by Eq. 3;
   (b) Reconstruct the compressed pixel information by taking the inverse of $\widehat{\text{DCT}}(\mathbf{f})$, i.e. $\mathbf{f}_{\text{compressed}} = D^{-1}\widehat{\text{DCT}}(\mathbf{f})$, according to Section 2.1.1;
   (c) Reshape $\mathbf{f}_{\text{compressed}}$ by 53×41 to obtain the compressed image $F_{\text{compressed}}$ and plot it.

3.3. **Compressed Image Recovery.** This step is to to recover the image F from limited random observations of its pixels using the prior knowledge that the $\text{DCT}(\mathbf{f})$ is nearly sparse. In order to complete this task, this work must 1) construct an image with limited random observation that forms a signal recovery problem discussed in Section 2.2.1, 2) generate a optimization problem as defined in Eq. 11 using theories introduced in Section 2.2.1 and 2.3, and 3) reconstruct the image by solving the optimization problem with Lasso shown in Section 2.3.1.

The algorithm of this step is as follows. For each ratio $r = 0.2, 0.4, 0.6$, perform the follow three times:

i. Construct a measurement matrix $B \in \mathbb{R}^{M\times N}$ by randomly selecting M rows of the identity matrix $I \in \mathbb{R}^{N\times N}$, where M is the nearest integer of $r \times N$;
ii. Generate a vector of random measurements $\mathbf{y} \in \mathbb{R}^M$ of the signal $\mathbf{f}$ by applying B to $\mathbf{f}$;
iii. Define the matrix $A = BD^{-1} \in \mathbb{R}^{M\times N}$ and use CVX to obtain β^* by solving the optimization problem formed by $A, \mathbf{y}$ defined as Eq. 11, using the following parameters: "`verbose=True, solver='CVXOPT', max_iter=1000, reltol=1e-2, featol=1e-2`";
iv. Reconstruct the image signals $\mathbf{f}^* = D^{-1}\beta^*$, resize $\mathbf{f}^*$ to obtain F^* by 53×41 and plot it.

3.4. **Mysterious Image Reconstruction.** This step is to reconstruct the mysterious image from the input measurement matrix B_i and vector of measurement $\mathbf{y}_i$, given the original image is of size 50×50. The algorithm is as follows.

i. Initialize the iDCT matrix of the mysterious image D_m^{-1} using the given helper function;
ii. Construct $A_m = B_m D_m^{-1}$ and form the optimization problem of $A_m, \mathbf{y}_m$ defined as Eq. 11;

iii. Solve the optimization problem of $A_m, \mathbf{y}_m$ for β_m^* with CVX using parameter: "`verbose=True`";
iv. Reconstruct the image signals $\mathbf{f}_m^* = D_m^{-1}\beta_m^*$, resize $\mathbf{f}_m^*$ to obtain F_m^* by 50×50 and plot it.

4. Computational Results

4.1. Image Compression. A plot of DCT over the grayscale resized image *Son of Man* is generated in Section 3.2 and shown as Figure 1. One may observe that, in Figure 1, there are only very limited among of coefficients that is significantly larger than 0 and most of the coefficients are 0 or nearly 0. In particular, DCT of the image has coefficients greater than 0 in the first 500 ones, where the rest about 1900 coefficients are close to 0.This confirms the sparsity of the DCT of the image signals, implying that image compression using only those non-zero coefficients is practical.

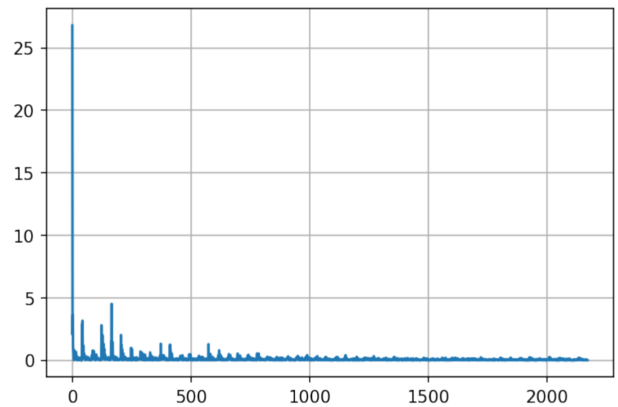

Figure 1. DCT coefficients of *Son of Man*

In Section 3.2, a plot of series compressed image of *Son of Man* at threshold equals $5, 10, 20, 40$ percent of DCT coefficients is generated and shown as Figure 2. This plot clearly demonstrates an improvement in image quality, which aligns with the assumed result at various threshold. Since the threshold decides the among of information to keep, then the larger the threshold is, the more information the compressed image will contain, resulting in better quality.

4.2. Compressed Image Recovery. In Section 3.3, the original signals are randomly corrupted at some ratios and recovered for three times. The plot for the total 9 recovered images is shown as Figure 3. Firstly, the reason for repeating the experiments is to make sure that the random measurement matrices are not too luck/unlucky. According to Figure 3, the recovered images' qualities are very similar to each other at the same information ratio. Hence, the experiments of this work is neither too lucky nor unlucky, implying that the resulting recovered images are convincing. In addition, one observe that as the ratio increase from 0.2 to 0.6, the quality of recovered images improve. This also matches our assumption that, the higher the ratio, the more information will be preserved in the corrupted image, i.e. the amount of r original information will randomly preserved in corrupted images.

4.3. Mysterious Image Reconstruction. The reconstructed image from the given measurement matrix and vector of measurement is shown as Figure 4. Though this image is merely of 50×50, one may simply tell that it is an image of the "Nyan Cat". Thus, it is safe to conclude that the mysterious image is successfully reconstructed.

5. Summary and Conclusions

Through both forward and backward discrete cosine transform and sparse recovery with Lasso, this work successfully 1) investigated the compressibility with DCT, 2) recovered randomly compressed images, and 3) reconstructed the image of "Nyan Cat" using only measurement data.

The limitation of this work is that this work played with only small-sized gray-scale images. The future work could be done with color-ed images of higher resolution, which may yield some interesting findings.

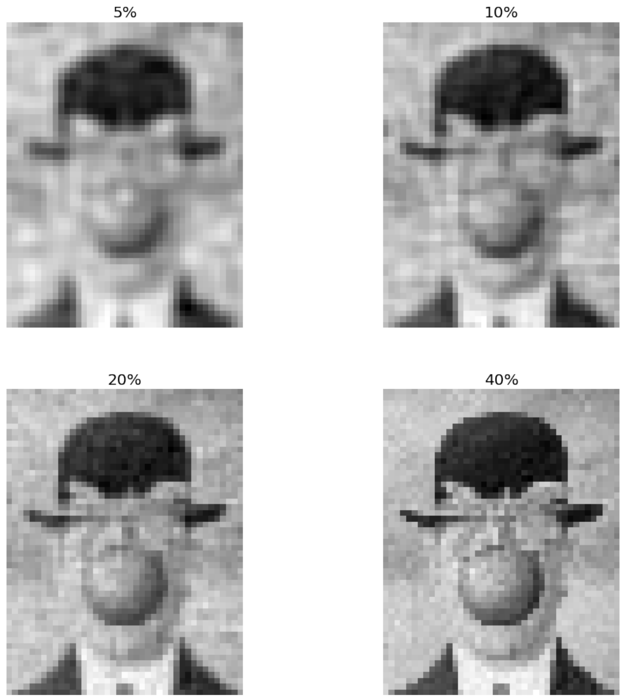

FIGURE 2. Compressed Son of Man at various thresholds

Acknowledgements

The author is sincerely thankful to Professor Bamdad Hosseini and TAs for AMATH 482 over this past quarter. The assignments are designed with interesting background and the author found joy in taking the class. Happy spring break!

References

[1] J. Makhoul. A fast cosine transform in one and two dimensions. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 28(1):27–34, 1980.

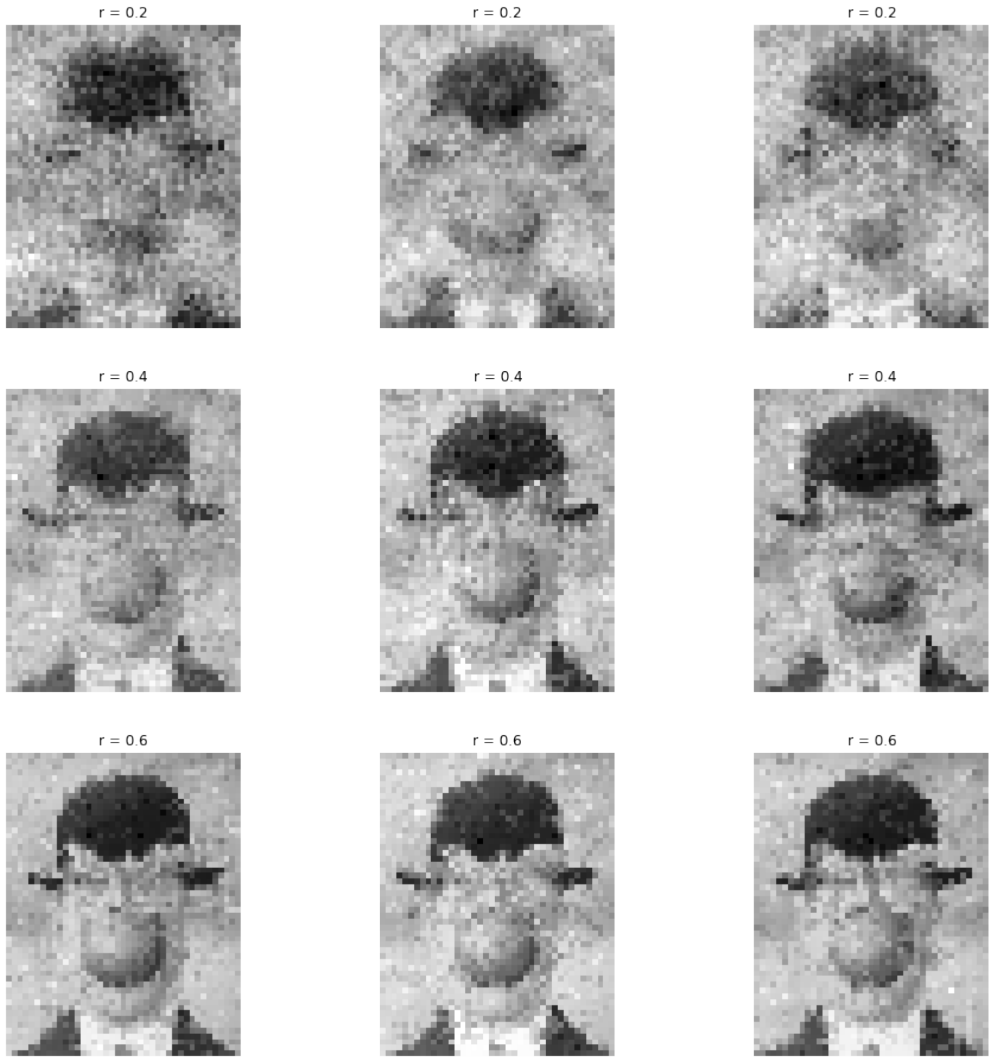

Figure 3. Recovered images at different information ratio

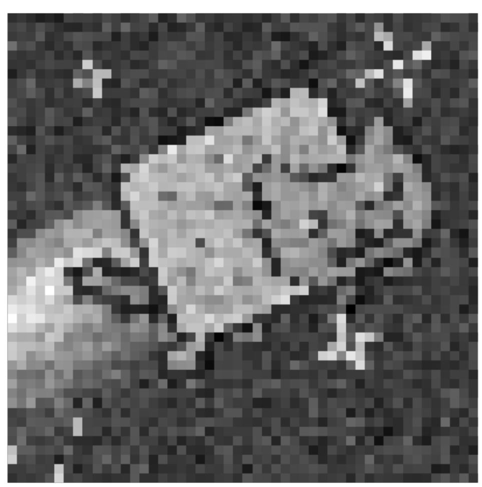

Figure 4. Recovered images at different information ratio