Assignment 1: CS 754, Advanced Image Processing

Due: 8th Feb before 11:55 pm

Remember the honor code while submitting this (and every other) assignment. All members of the group should work on and *understand* all parts of the assignment. Exchange of answers between groups is not allowed. We will adopt a zero-tolerance policy against any violation, and we will expressly check for plagiarism.

Submission instructions: You should ideally type out all the answers in Latex or else in MS Word with the equation editor. In either case, prepare a pdf file. Create a single zip or rar file containing the report, code and sample outputs and name it as follows: A1-IdNumberOfFirstStudent-IdNumberOfSecondStudent.zip. (If you are doing the assignment alone, the name of the zip file is A1-IdNumber.zip). Upload the file on moodle BEFORE 11:55 pm on 8th Feb. Late assignments will be assessed a penalty of 50% per day late. Note that only one student per group should upload their work on moodle. Please preserve a copy of all your work until the end of the semester. *If you have difficulties, please do not hesitate to seek help from me.*

1. The first question tests your Google search skills :-). Your job is to look for a paper describing a compressed sensing architecture which is *different* from the ones we have seen in class. It may be a video/hyperspectral/color compressive camera but it cannot be any architecture done in class. Your job is to read the paper, and do as follows: (a) Summarize the architecture describe in the paper. You may refer to figures or tables in the paper, but write out a few key equations. Explain how it is different from the architecture studied in class that was most similar to it. (b) State the objective function presented in the paper for compressive reconstruction. Remember, that this is the only question for which a Google search is allowed. Apart from Google, good places to look for are archives of journals such as IEEE Transactions on Computational Imaging, Journal of the Optical Society of America, SIAM Journal of Imaging Sciences, and archives of various conferences such as ICIP, CVPR, ICCV, ECCV, ICASSP. [15 points]

2. What is the lower bound on the number of compressive measurements for exact reconstruction of a signal in $\mathbb{R}^n$ that is s in some orthonormal basis $\boldsymbol{\Psi}$ if (a) the basis pursuit algorithm is used for reconstruction, (b) a combinatorial algorithm (problem P0 as defined in the slides) is used for reconstruction? In each case, does this error improve or worsen if the signal were s-sparse in a different orthonormal basis? Explain. In both cases, what is the maximum bound allowed on the RIC of the matrix $\boldsymbol{\Phi\Psi}$? [6+4+5=15 points]

3. Prove the following relationship between the restricted isometry constant of order s of a matrix $\boldsymbol{A}$ (denoted as δ_S) and the mutual coherence μ of $\boldsymbol{A}$: $\delta_s \leq (s-1)\mu$. Assume all columns of $\boldsymbol{A}$ are unit-normalized. Show all steps very carefully. [15 points]

4. We will prove why the value of the coherence between $m \times n$ measurement matrix $\boldsymbol{\Phi}$ (with all rows normalized to unit magnitude) and $n \times n$ orthonormal representation matrix $\boldsymbol{\Psi}$ must lie within the range $(1, \sqrt{n})$. Recall that the coherence is given by the formula $\mu(\boldsymbol{\Phi}, \boldsymbol{\Psi}) = \sqrt{n} \max_{i,j \in \{0,1,\ldots,n-1\}} |\boldsymbol{\Phi^{i}}^t \boldsymbol{\Psi_j}|$. Proving the upper bound should be very easy for you. To prove the lower bound, proceed as follows. Consider a unit vector $\boldsymbol{g} \in \mathbb{R}^n$. We know that it can be expressed as $\boldsymbol{g} = \sum_{k=1}^{n} \alpha_k \boldsymbol{\Psi_k}$ as $\boldsymbol{\Psi}$ is an orthonormal *basis*. Now prove that $\mu(\boldsymbol{g}, \boldsymbol{\Psi}) = \sqrt{n} \max_{i \in \{0,1,\ldots,n-1\}} \dfrac{|\alpha_i|}{\sum_{j=1}^{n} \alpha_j^2}$. Exploiting the fact that $\boldsymbol{g}$ is a unit vector, prove that the minimal value of coherence is given by $\boldsymbol{g} = \sqrt{1/n} \sum_{k=1}^{n} \boldsymbol{\Psi_k}$ and hence the minimal value of coherence is 1. [15 points]

5. In class, we studied a video compressive sensing architecture from the paper 'Video from a single exposure coded snapshot' published in ICCV 2011 (See `http://www.cs.columbia.edu/CAVE/projects/single_`

`shot_video/`). Such a video camera acquires a 'coded snapshot' E_u in a single exposure time interval u. This coded snapshot is the superposition of the form $E_u = \sum_{t=1}^{T} C_t \cdot F_t$ where F_t is the image of the scene at instant t within the interval u and C_t is a randomly generated binary code at that time instant, which modulates F_t. Note that E_u, F_t and C_t are all 2D arrays. Also, the binary code generation as well as the final summation all occur within the hardware of the camera. Your task here is as follows:

(a) Read the video in the homework folder in MATLAB using the 'mmread' function which has been provided in the homework folder and convert it to grayscale. Extract the first $T = 3$ frames of the video.

(b) Generate a $H \times W \times T$ random code pattern whose elements lie in $\{0, 1\}$. Compute a coded snapshot using the formula mentioned and add zero mean Gaussian random noise of standard deviation 2 to it. Display the coded snapshot in your report.

(c) Given the coded snapshot and assuming full knowledge of C_t for all t from 1 to T, your task is to estimate the original video sequence F_t. For this you should rewrite the aforementioned equation in the form $\boldsymbol{Ax} = \boldsymbol{b}$ where $\boldsymbol{x}$ is an unknown vector (vectorized form of the video sequence). Mention clearly what $\boldsymbol{A}$ and $\boldsymbol{b}$ are, in your report.

(d) You should perform the reconstruction using Orthogonal Matching Pursuit (OMP), an algorithm we shall see in class very soon. For computational efficiency, we will do this reconstruction patchwise. Write an equation of the form $\boldsymbol{Ax} = \boldsymbol{b}$ where $\boldsymbol{x}$ represents the i^{th} patch from the video and having size (say) $8 \times 8 \times T$ and mention in your report what $\boldsymbol{A}$ and $\boldsymbol{b}$ stand for. For perform the reconstruction, assume that each 8×8 slice in the patch is sparse or compressible in the 2D-DCT basis. Carefully work out the error term in the OMP algorithm!

(e) Repeat the reconstruction for all overlapping patches and average across the overlapping pixels to yield the final reconstruction. Display the reconstruction and mention the relative mean squared error between reconstructed and original data, in your report as well as in the code.

(f) Repeat this exercise for $T = 5, T = 7$ and mention the mention the relative mean squared error between reconstructed and original data again.

(g) **Note: To save time, extract a portion of about 120×240 around the lowermost car in the cars video and work entirely with it. In fact, you can show all your results just on this part. Some sample results are included in the homework folder.** [40 points]