Assignment 5: CS 754, Advanced Image Processing

Due: 20th April 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. We will adopt a zero-tolerance policy against any violation.

Submission instructions: You should ideally type out all the answers in Word (with the equation editor) or using Latex. 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: A5-IdNumberOfFirstStudent-IdNumberOfSecondStudent.zip. (If you are doing the assignment alone, the name of the zip file is A5-IdNumber.zip). Upload the file on moodle BEFORE 11:55 pm on the due date. **There will be no late extensions for this assignments as the due date is the last day of classes**. 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. We have derived a model showing how the DCT coefficients of an image are Laplacian distributed. This model is derived by assuming a Gaussian distribution for the DCT coefficients of a patch assuming fixed variance, followed by imposing an exponential distribution on the patch variance values. Now suppose that the patch variance values were distributed as Uniform$(0, b)$, as is common in document images. Your job is to derive the resultant distribution of the DCT coefficients. You will not obtain a closed-form expression (unlike the case for natural images which yielded the Laplacian) so you will need to resort to numerical integration and plot the final distribution, as well as its closest Gaussian fit. Refer to the paper 'Analysis of the DCT Coefficient Distributions for Document Coding' by Edmund Lam, IEEE Signal Processing Letters, Feb 2004. You may work with images d5.jpg and d6.png in the homework folder. [20 points]

2. Refer to the paper 'User assisted separation of reflections from a single image using a sparsity prior' by Anat Levin, IEEE Transactions on Pattern Analysis and Machine Intelligence. Answer the following questions:
   - In Eqn. (7), explain what $A_{j\rightarrow}$ and b_j represent, for each of the four terms in Eqn. (6).
   - In Eqn. (6), which terms are obtained from the prior and which terms are obtained from the likelihood? What is the prior used in the paper? What is the likelihood used in the paper?
   - Why does the paper use a likelihood term that is different from the more commonly used Gaussian prior? [20 points]

3. We have studied Non-negative matrix factorization (NMF) and non-negative sparse coding (NNSC) in class. We have seen their applications in image denoising and in face recognition. Your job is to do a google search and find out a paper which explores any other application of either NMF or NNSC or some modification of these techniques. The application should be in the domain of image processing, machine learning, computer vision, audio processing or data mining. In your report, you should briefly describe the application, the main objective function being optimized in the paper, and a summary of the results. [20 points]

4. The following question refers to the paper 'Data Separation by Sparse Representations' by Gitta Kutyniok. The paper pdf can be found in the homework folder. Your job is to trace through the proofs of Lemma 1.1, Lemma 1.2 and Theorem 1.3 of the paper and answer the questions asked in the pdf of the paper. Also what are the advantages and disadvantages of the result Theorem 1.1 from the paper in comparison to a standard RIP-based result one commonly sees in compressed sensing? [7 x 2 + 6 = 20 points]

5. Consider compressive measurements of the form $\boldsymbol{y} = \boldsymbol{\Phi x} + \boldsymbol{\eta}$ under the usual notations with $\boldsymbol{y} \in \mathbb{R}^m, \boldsymbol{\Phi} \in \mathbb{R}^{m \times n}, m \ll n, \boldsymbol{x} \in \mathbb{R}^n, \boldsymbol{\eta} \sim \mathcal{N}(0, \sigma^2 \boldsymbol{I}_{m \times m})$. Instead of the usual model of assuming signal sparsity in an orthonormal basis, consider that $\boldsymbol{x}$ is a random draw from a zero-mean Gaussian distribution with known covariance matrix $\boldsymbol{\Sigma_x}$ (of size $n \times n$). Derive an expression for the maximum a posteriori (MAP) estimate of $\boldsymbol{x}$ given $\boldsymbol{y}, \boldsymbol{\Phi}, \boldsymbol{\Sigma_x}$ (see lecture slides for help). Also, run the following simulation: Generate $\boldsymbol{\Sigma_x} = \boldsymbol{U \Lambda U}^T$ of size 128×128 where $\boldsymbol{U}$ is a random orthonormal matrix, and $\boldsymbol{\Lambda}$ is a diagonal matrix of eigenvalues of the form $ci^{-\alpha}$ where $c = 1$ is a constant, i is an index for the eigenvalues with $1 \leq i \leq n$ and α is a decay factor for the eigenvalues. Generate 10 signals from $\mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma_x})$. For $m \in \{40, 50, 64, 80, 100, 120\}$, generate compressive measurements of the form $\boldsymbol{y} = \boldsymbol{\Phi x} + \boldsymbol{\eta}$ for each signal $\boldsymbol{x}$. In each case, $\boldsymbol{\Phi}$ should be a matrix of iid Gaussian entries with mean 0 and variance $1/m$, and $\sigma = 0.01\times$ the average absolute value in $\boldsymbol{\Phi x}$. Reconstruct $\boldsymbol{x}$ using the MAP formula, and plot the average RMSE versus m for the case $\alpha = 3$ and $\alpha = 0$. Comment on the results - is there any difference in the reconstruction performance when α is varied? If so, what could be the reason for the difference? [20 points]