Assignment 5: CS 754, Advanced Image Processing

Due: 18th 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 <u>understand</u> 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. The cutoff is 10 am on 19th April after which no assignments will be accepted. 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. <u>If you have difficulties, please do not hesitate to seek help from me.</u>

1. Consider that you learned a dictionary D to sparsely represent a certain class $\mathcal{S}$ of images - say handwritten alphabet or digit images. How will you convert D to another dictionary which will sparsely represent the following classes of images? Note that you are not allowed to learn the dictionary all over again, as it is time-consuming.

   (a) Class $\mathcal{S}_1$ which consists of images obtained by applying a known derivative filter to the images in $\mathcal{S}$.

   (b) Class $\mathcal{S}_2$ which consists of images obtained by rotating a subset of the images in class $\mathcal{S}$ by a known fixed angle α, and the other subset by another known fixed angle β.

   (c) Class $\mathcal{S}_3$ which consists of images obtained by applying an intensity transformation $I^i_{new}(x, y) = \alpha(I^i_{old}(x,y))^2 + \beta(I^i_{old}(x,y)) + \gamma$ to the images in $\mathcal{S}$, where α, β, γ are known.

   (d) Class $\mathcal{S}_4$ which consists of images obtained by applying a known blur kernel to the images in $\mathcal{S}$.

   (e) Class $\mathcal{S}_5$ which consists of images obtained by applying a blur kernel which is known to be a linear combination of blur kernels belonging to a known set $\mathcal{B}$, to the images in $\mathcal{S}$.

   (f) Class $\mathcal{S}_6$ which consists of 1D signals obtained by applying a Radon transform in a known angle θ to the images in $\mathcal{S}$.

   (g) Class $\mathcal{S}_7$ which consists of images obtained by translating a subset of the images in class $\mathcal{S}$ by a known fixed offset (x_1, y_1), and the other subset by another known fixed offset (x_2, y_2). Assume appropriate zero-padding and increase in the size of the image canvas owing to the translation. [4+4+4+4+4+6+4=30 points]

2. How will you solve for the minimum of the following objective functions: (1) $J(\boldsymbol{A_r}) = \|\boldsymbol{A} - \boldsymbol{A_r}\|^2_F$, where $\boldsymbol{A}$ is a known $m \times n$ matrix of rank greater than r, and $\boldsymbol{A_r}$ is a rank-r matrix, where $r < m, r < n$. (2) $J(\boldsymbol{R}) = \|\boldsymbol{A} - \boldsymbol{RB}\|^2_F$, where $\boldsymbol{A} \in \mathbb{R}^{n \times m}, \boldsymbol{B} \in \mathbb{R}^{n \times m}, \boldsymbol{R} \in \mathbb{R}^{n \times n}, m > n$ and $\boldsymbol{R}$ is constrained to be orthonormal. Note that $\boldsymbol{A}$ and $\boldsymbol{B}$ are both known.
   In both cases, explain briefly any one situation in image processing where the solution to such an optimization problem is required. [5+5+5+5=20 points]

3. We have studied the non-negative matrix factorization (NMF) technique in our course and examined applications in face recognition. I also described the application to hyperspectral unmixing. Your job is to find a research paper which explores an application of NMF in any task apart from these. You may look up the

wikipedia article on this topic. Other interesting applications include stain normalization in pathology. Your job is to answer the following: (1) Mention the title, author list, venue and year of publication of the paper and include a link to it. (2) Which task does the paper apply NMF to? (3) How exactly does the paper solve the problem using NMF? What is the significance of the dictionary and the dictionary coefficients in solving the problem at hand? [15 points]

4. In parallel bean computed tomography, the projection measurements are represented as a single vector $\boldsymbol{y} \sim \text{Poisson}(I_o \exp(-\boldsymbol{Rf}))$, where $\boldsymbol{y} \in \mathbb{R}^m$ with $m =$ number of projection angles $\times$ number of bins per angle; I_o is the power of the incident X-Ray beam; $\boldsymbol{R}$ represents the Radon operator (effectively a $m \times n$ matrix) that computes the projections at the pre-specified known projection angles; and $\boldsymbol{f}$ represents the unknown signal (actually tissue density values) in $\mathbb{R}^n$. If $m < n$, write down a suitable objective function whose minimum would be a good estimate of $\boldsymbol{f}$ given $\boldsymbol{y}$ and $\boldsymbol{R}$ and which accounts for the Poisson noise in $\boldsymbol{y}$. State the motivation for each term in the objective function. Recall that if $z \sim \text{Poisson}(\lambda)$, then $P(z = k) = \lambda^k e^{-\lambda}/k!$ where k is a non-negative integer. Now suppose that apart from Poisson noise, there was also iid additive Gaussian noise with mean 0 and known standard deviation σ, in $\boldsymbol{y}$. How would you solve this problem (eg: appropriate preprocessing or suitable change of objective function)? [6+ 4 = 10 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 $\boldsymbol{y}, \boldsymbol{\Phi}, \boldsymbol{\Sigma_x}$. 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? [25 points]