Biomedical Imaging Informatics Practical #3

Instructions:

This group assignment is expected to **require 5-8 hours to complete.** Please plan accordingly.

Due Date: **Nov. 21st (Thursday) by 11:59pm EST**.
Submission Files: the **signed HONOR CODE page** (see HONOR CODE page), and **one zipped program folder file** (i.e., containing all code files)

Submission Protocol:
Upload the zipped file to your home directory on knn.bme.gatech.edu, and then send an email to 8813grading@lists.gatech.edu to inform us that you have submitted your assignment. For your email subject line, please use: BMED8813 last name 1, last name 2, group #, PRACTICAL #3.

Submission File Protocol:
(1) Your submission will be one zipped folder file. The folder should contain all of the necessary files (i.e., the main program file, possible additional function files, and necessary test data files) to run your code and show the results.
(2) The folder should be named: < last name 1>_<last name 2>_..._<group #>
(3) In your folder, code files should be named as: <last name 1>_<last name 2>_..._<group #>_<code name>.m (or .pl)
(4) Please write well-commented code, and include any explanations you think will help the instructor understand your program. If your program does not work completely, comments might help you get partial credit. However, you will get **NO CREDIT** if your program generates errors, generates warnings, or outputs "junk".

If you have problems, please contact the instructors at 8813grading@lists.gatech.edu with the proper email subject line as instructed in the syllabus.

HONOR CODE

The conditions of this assignment are subject to the Georgia Institute of Technology Academic Honor Code.

I pledge that the work in this assignment, including all written code, represents the original work of BMED8813 BHI Group ____ [write group #]. I have NOT communicated with anyone (other than my group members) about the contents of this assignment, nor participated in or observed any conduct prohibited by the Honor Code.

Student Signature _____

Student Signature _____

Student Signature _____

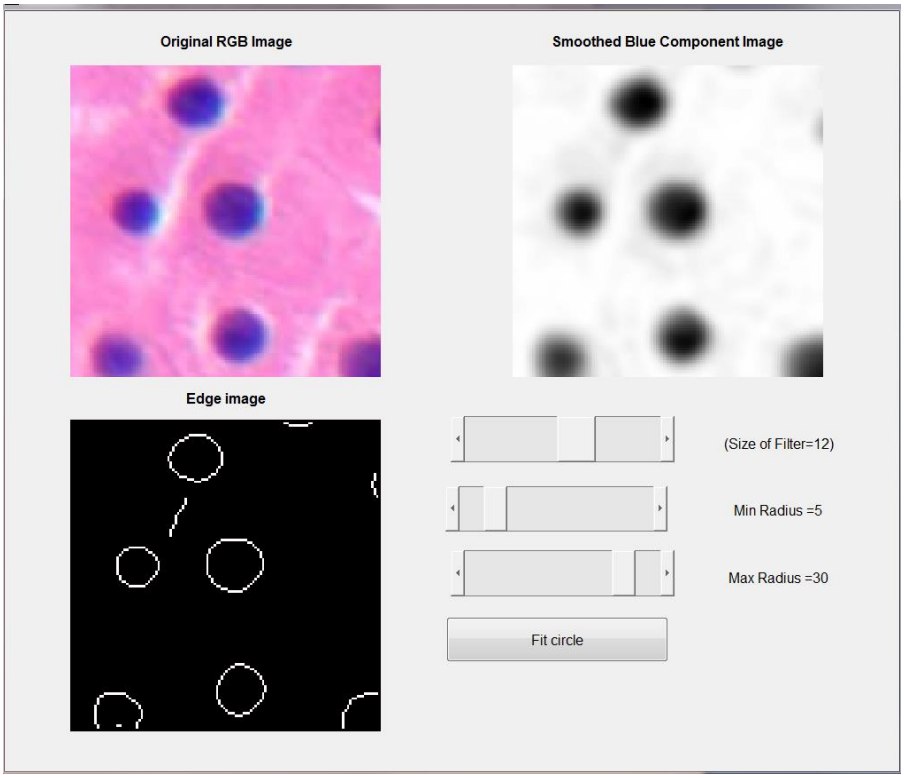
NOTE: Please print out this page, sign the page, scan the signed page, and submit with your programming code.

If you have a problem finding a scanner to scan in the page, please type in the above HONOR CODE statement in your email message body and sign with your name.

Any assignment without the HONOR CODE PAGE will be void to ZERO grade by default.

1) (50 Points) Cancer biopsy examination by a pathologist is a standard procedure for making diagnostic decisions. Pathologists often analyze nuclear properties while examining the tissue. In this problem, please implement a Hough transform for circle detection and use it for segmenting nuclei in histological images. Based on the segmentation result, please calculate two useful nuclear properties—average nuclear radius and number of nuclei in the image.

a) Design a GUI as shown below.



GUI for Hough transform

This GUI has three axes representing three images—

i) Original RGB image (*M* x *N* pixels)

ii) Smoothed blue component of the RGB image (*M* x *N* pixels)—to generate this image, take the blue component of the RGB image and smooth it using a Gaussian filter. The size of the Gaussian filter is controlled by a slider. If the slider value is *n*, then the filter is *n* x *n* and its standard deviation is *n*/5. (Hint: you may use the Matlab functions "fspecial" and "imfilter")

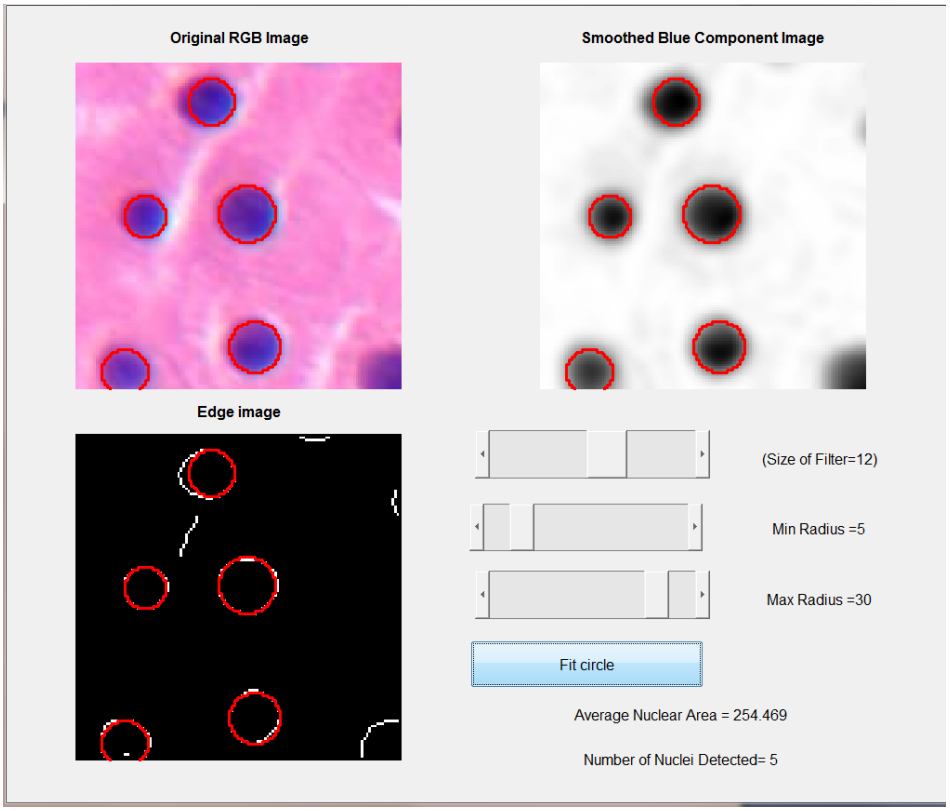
iii) Edge image (*M* x *N* pixels)—to generate this image (*Edge_image*) extract edges from the smoothed image using the "canny" edge detector. (Hint: you may use the Matlab function "edge")

This GUI also has three sliders—

i) Size of Filter—this slider controls the size of the Gaussian filter. Whenever user updates this slider, "smoothed image" and "edge image" axes are cleared and the images are updated.
ii) Min Radius—this slider controls the minimum radius, *Min_R*, for the Hough transform.
iii) Max Radius—this slider controls the maximum radius, *Max_R*, for the Hough transform.

This GUI has one push-button, "Fit Circle", that will run the Hough transform and extract circular nuclei from "Edge image", with radius in the range of "Min Radius" to "Max Radius".

This GUI also has two text objects that display the average nuclear area and number of nuclei detected.



GUI with results

b) Please code the following function for the Hough transform accumulator:
Hough_Accum=Generate_Accum(*Edge_image*, *Min_R*, *Max_R*)

Where *Edge_image* is the edge image, *Min_R* and *Max_R* are minimum and maximum radii for the Hough transform respectively. The function returns a 3-D matrix of size *M* x *N* x length (*R_range*), where *R_range*=*Min_R*:1:*Max_R*.

Hence, *Hough_Accum*(:,:,1) and *Hough_Accum*(:,:,length(*R_range*)) correspond to *Min_R* and *Max_R*, respectively.

Please use the following algorithm for generating the accumulator.

1. Initialize accumulator, *Hough_Accum,* to zero.
2. For all radii, *R*, in *R_range*
    a. Generate circle coordinates [**x_c**, **y_c**] in Cartesian coordinates assuming center at origin.
    In polar coordinates, the circle is given by *R* and *theta*=0:*step*:2*pi. Assume *step* such that *theta* range is sampled 500 times. You may generate Cartesian coordinates from the Polar coordinates using the Matlab function "pol2cart".
    b. For all edge pixels:
        i. Move the origin of the circle to the edge point.
        ii. Round the coordinates to the nearest integer. Reject the coordinates outside the accumulator range.
        iii. Increment the accumulator at these coordinates and 3rd dimension corresponding to *R* by 1/*R*.

c) Please code the following function for the Hough transform peak detector
Hough_Peaks = Detect_Peaks (*Hough_Accum, R_range*)
Where *Hough_peaks* is a list of circles with their radius and center coordinates.

Please use the following rules while generating the peaks.

1. Define three parameters: *xy_margin*, *R_margin* and *Npeaks*. Both center coordinates of any two circles should be at least *xy_margin* distance apart and the radii should differ by at least *R_margin*. *Npeaks* defines the maximum number of peaks detected. Set all three parameters to 15.
2. Only detect a peak if its value is greater than or equal to half of maximum value in accumulator matrix and it is a maximum in local 3-D neighborhood defined by *xy_margin* and *R_margin.*

d) Plot the detected circles on "original", "smoothed" and "edge" images. Also, calculate the number of nuclei detected and their average area.

e) Use this GUI to find optimal parameters, including filter size, *n*, minimum radius, *Min_R* and maximum radius, *Max_R*, for the four sample images provided—**"**Image1.mat", **"**Image2.mat", **"**Image3.mat", and "Image4.mat". Report these parameters with snapshots of the GUI results.

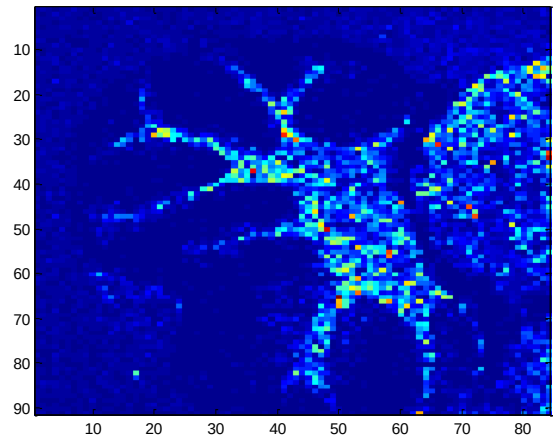
f) Comment on advantages and limitations of the Hough Transform and effect of smoothing on results.

2) (25 Points) Tissue imaging mass spectrometry (TIMS) is a technique for analyzing biological tissue samples. In this process, a mass spectrum is acquired at multiple locations on a tissue sample. The (91x84x50) TIMS dataset in this assignment contains 7644 mass spectra. The dataset can be interpreted as a set of 50 images with dimensions (91x84), each describing the spatial distribution of a biological component (for instance, a protein or other molecule) within the tissue. In this way, TIMS datasets provide a molecular description of tissue samples.

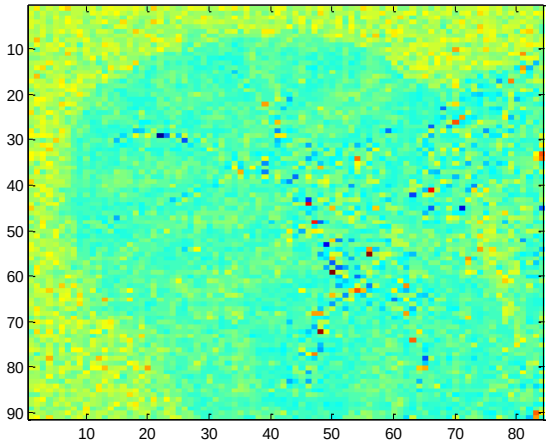
Principal component analysis (PCA) can be used to identify regions within the tissue sample which have distinctive molecular compositions (for instance, different tissue types, or diseased vs. healthy parts of the same tissue). The output of PCA can be interpreted as lists of molecules which are highly associated with areas of interest, such as diseased regions. This information can guide further experiments and may assist in identifying disease markers ('biomarkers').

To perform the PCA on TIMS:

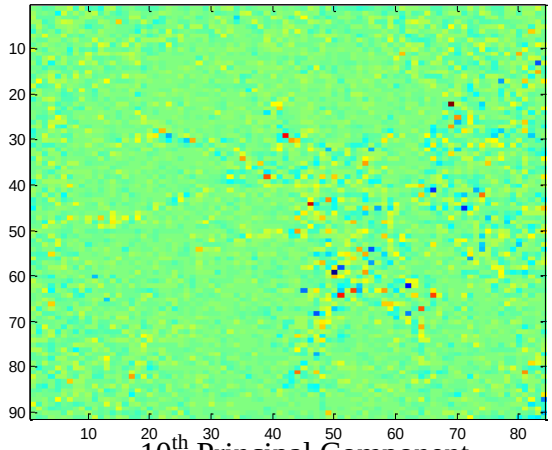
a) Read the input data 'TIMS_data.mat'.
b) Reshape each 2d matrix as a vector (would result in 50 x 7644 data matrix).
c) Find the mean vector of the transposed data (results in 50 x 1 vector).
d) Find the covariance matrix of the data (Matlab:cov())
$$\mathbf{C}_{\mathbf{x}} = E\left\{(\mathbf{x}-\mathbf{m}_{\mathbf{x}})(\mathbf{x}-\mathbf{m}_{\mathbf{x}})^{T}\right\}$$
e) Compute Eigen values and Eigen vectors. (Hint. Matlab: eig())
f) Sort the Eigen values and reorder the Eigen vectors matrix. (i.e., the first vector should correspond to the biggest Eigen value and so on).
g) Find the representation of data in principal component space. (equation 11.4-6, also called the Hotelling transform)
$$\mathbf{y} = \mathbf{A}(\mathbf{x}-\mathbf{m}_{\mathbf{x}})$$
h) Reconstruct the original data. (Equation 11.4-10.)
$$\mathbf{x} = \mathbf{A}^{T}\mathbf{y} + \mathbf{m}_{\mathbf{x}}$$
i) Is the reconstructed data exactly same as input data? Please explain.
j) Generate the images of the 1st, 5th, 10th, 15th, 20th and 50th components of the PCA space. (Hint: k^{th} row of y)
$$\mathbf{y} = \mathbf{A}(\mathbf{x}-\mathbf{m}_{\mathbf{x}})$$
k) What do these components capture, especially the 1st component?



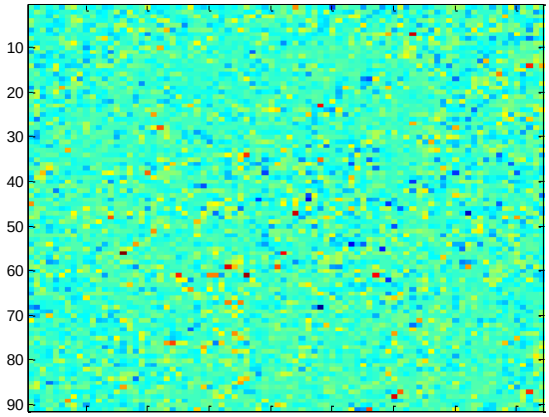
1st Principal Component



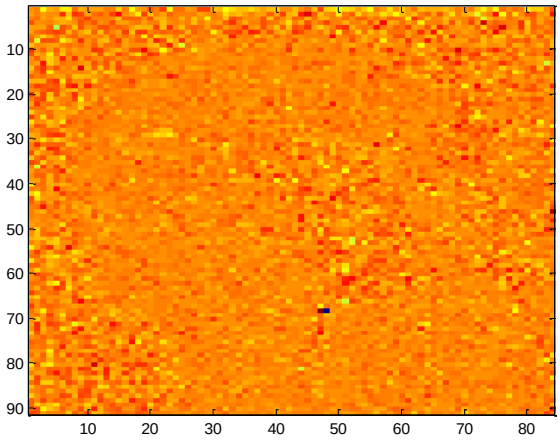
5th Principal Component



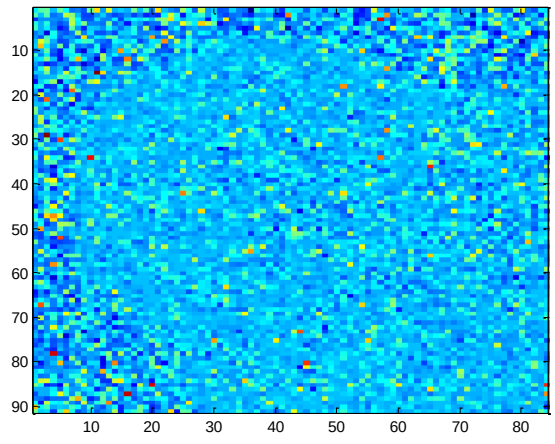
10th Principal Component



15th Principal Component



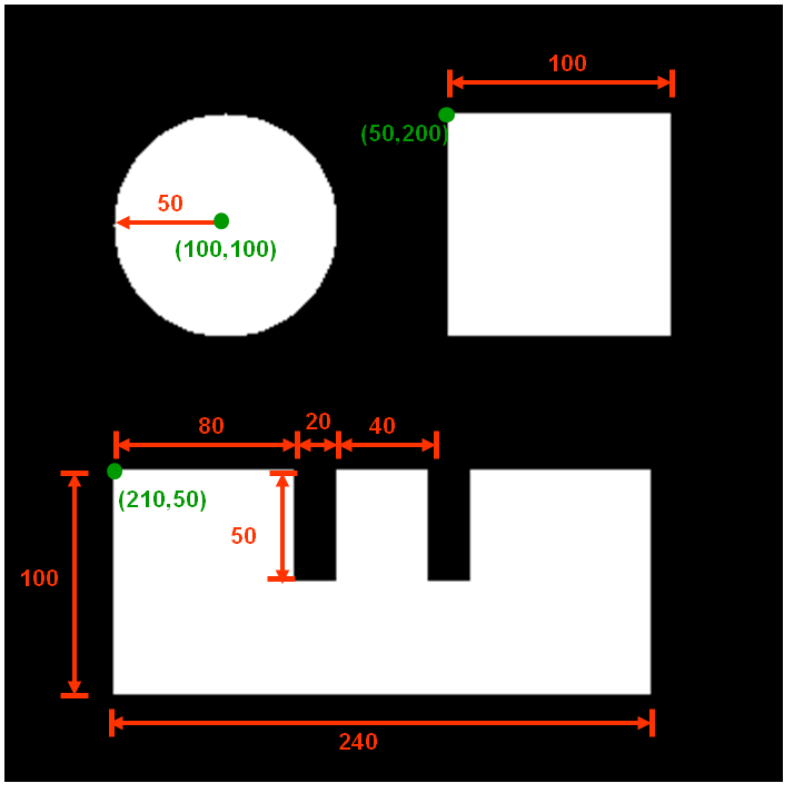
20th Principal Component



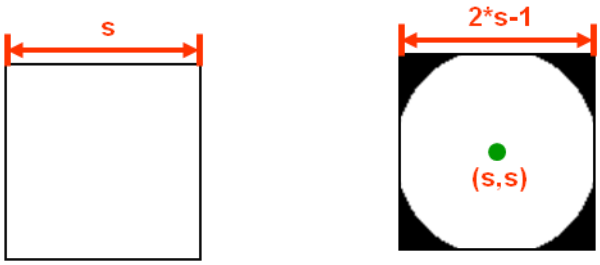
50th Principal Component

3) (25 Points) Morphological image processing is an important tool for image segmentation and image de-noising. In this problem, please implement different types of morphological operations and apply them to a synthetic image.

a) Please create an input image of size 350 x 350 as shown (dimensions and centers of objects are shown).



b) Create two structure elements, i) a square of size 11, and ii) a disk of radius 6.



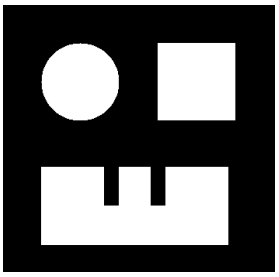
c) Dilate and erode the image using these structuring elements. (do not use "imerode" or "imdilate")
(Hint: do the edge detection on the input image to detect the boundary of the objects. Then center the structuring element on the boundary of the objects and perform erosion or dilation.)

d) Open and close the image.

e) Perform step 2 and 3 for a square of size 17, and 21, and a disk of radius 8, and 10.

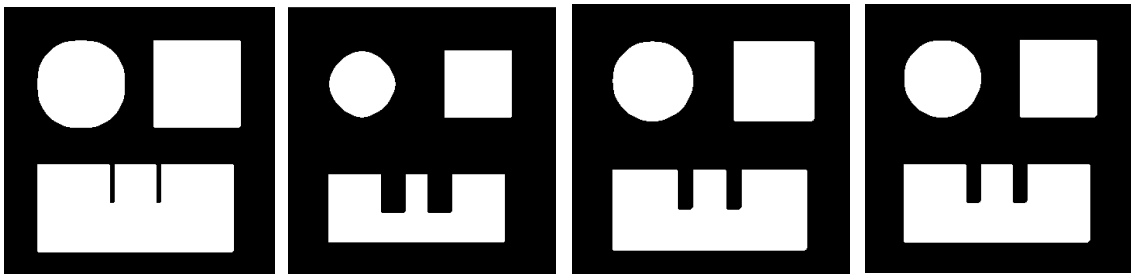
f) What is the difference between the closing and opening operations?

g) Find the size of the structural elements for which the third shape is divided into two parts for closing and opening operations. Please also provide the result in your report.



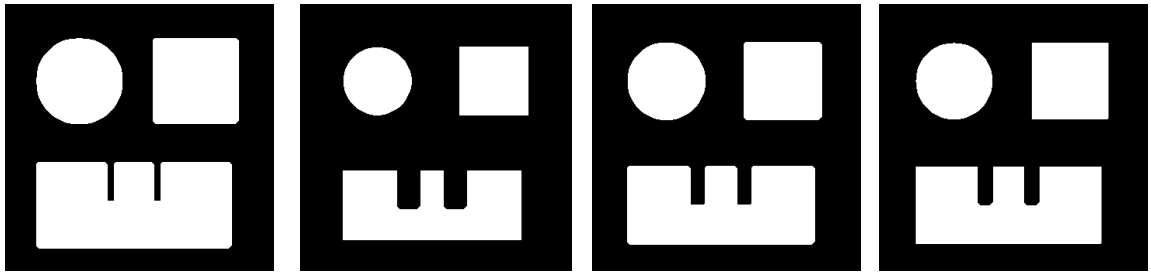
Input Image

Morhpological operations with a square of size, s, 11.



Dilate Erode Open Close

Morhpological operations with a disc of size, s, 6.



Dilate Erode Open Close