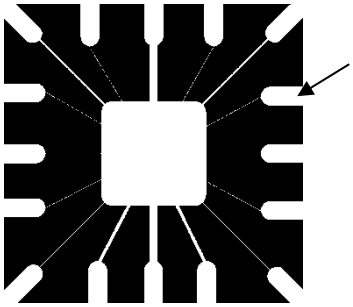
CS5680/CS6680 – Fall Semester 2021
Final Take-Home Exam
Due: 11:59 p.m., Monday, December 13, 2021
Total Points: 50 points

General Instructions:

Please ensure to have enough detail for each solution. You need to go through each step by your own computation. Do not cheat and refer to the cheating policy on the syllabus.

Problem 1 [Total: 5 points]: (Medium-Level Image Processing)

1.1. Given the original image shown below, **write down one morphological solution** to keep all 16 rounded thick sticks on the border (one sample rounded thick stick is pointed by the arrow) and remove the other components in the image. Make sure that you clearly indicate structuring element(s) and their origin(s) and **use the structuring elements with the minimum size** to solve the problem. There is no need to write Matlab code.



Problem 2 [Total: 10 points]: (Advanced Image Processing)

2.1. Seeded Region Growing [5 points] [Ch5.DIPSegmentation.pdf slides 40-45]

a. For the given data with the two atomic regions shaded (i.e., the intermediate result of the seeded region growing), go through the calculations to show the **next step** in the seeded region growing process. That is, determine which pixel will be picked out for merging into either of the two shaded regions and which pixels, if applicable, will be determined as boundary pixels to segment the two regions. (Use 8-connected neighborhood for your calculation).

b. Instead of using the two seeds as shown in the figure below, two different seeds are chosen for the seeded region growing algorithm. One seed is located at the upper right corner and the other seed is located at the lower-left corner. Discuss the difference(s) between the segmentation results obtained by the seeded region growing algorithm using these two sets of seeds.

3	2	2	1	1	1	1	1
6	5	2	2	1	1	1	1
7	7	7	7	7	1	1	1
7	7	6	7	6	7	1	1
8	7	7	8	6	7	1	1
8	8	7	3	2	1	1	1
8	8	8	4	2	1	2	1
8	8	8	5	1	1	1	1

Seed2

Seed1

2.2. Watershed Segmentation [5 points] [Ch5.DIPSegmentation.pdf slides 47-67]

a. Identify the minimums in the given data, which contains the gradient of an 8×8 image. Explain why they are minimums.

b. Determine the watershed line, which falls on the boundaries between catchment basins of the minimums. (Use 8-connected neighborhood for your calculation).

8	8	7	8	9	9	9	9
8	8	3	3	9	9	4	9
8	8	9	9	9	9	9	9
8	7	6	8	9	9	9	9
7	6	6	7	7	9	1	1
7	2	2	7	6	5	1	1
7	6	6	7	11	12	5	5
7	7	6	7	11	11	7	5

Problem 3 [Total 35 points]: (Pattern Recognition)
For all the sub-problems, you may treat the original data set as the training data set.

3.1. [6 points] Apply **the K-means algorithm [Ch7.2.PRDistanceFunctions.pdf slides 23-34]** to the data set {(0, 0)', (0, 1)', (6, 7)', (5, 5)', (4, 5)', (1, 2)'} to cluster the samples (i.e., patterns) into **two** appropriate pattern classes. Here, the initial cluster centers are selected as {(4, 5)'} and {(0, 1)'}.

3.2. [5 points] Apply the **Hierarchical Clustering algorithm [Ch7.2.PRDistanceFunctions.pdf slides 35 and 38]** on the same data set described in 3.1 to cluster the samples into **two** appropriate pattern classes. Make sure that the single-link clustering is used in step 3.

3.3. [10 points]

- Apply the **perceptron algorithm [Ch7.1.PRStatistics.pdf slides 3-13]** on the two pattern classes obtained from problem 3.1 to find a solution weight vector to linearly separate these two pattern classes. Initially, set **the weight vector w(1) as (1, 1, 2)**. The correction increment *c* is set as 1.
- Plot all your original data and sketch the found decision boundary on a graph.
- Decide which class an unknown pattern (2, 2) belongs to and explain how you made the decision.

3.4. [6 points]

- Apply the **minimum-distance classification technique [Ch7.2.PRDistanceFunctions.pdf slides 11-16]** to find the decision boundary to separate the two pattern classes obtained from problem 3.1.
- Plot all your original data and sketch the found decision boundary on a graph.
- Decide which class an unknown pattern (2, 2) belongs to and explain how you made the decision.

3.5. [8 points] Apply the **Principal Component Analysis (PCA) approach [Ch7.3.PRModel.pdf slides 3-22]** to all six training data in problem 3.1.

- Find the transformation matrix.
- Represent each data using its two most important principal components (i.e., the first two principal components), plot six represented (or transformed) data on a graph, and compute their variance.
- Represent each data using its most important principal component (i.e., the first principal component), plot six represented (or transformed) data on a graph, and compute their variance.
- Compare two variances computed in the previous two steps to discuss whether the two representations better describe the property of the original data set.