EE 569 Introduction to Digital Image Processing

Spring 2019

Midterm Exam 2

Name: _____

Student ID: _____

Exam Location (Please check one of the following):

- USC Campus
- Remote

Notes:

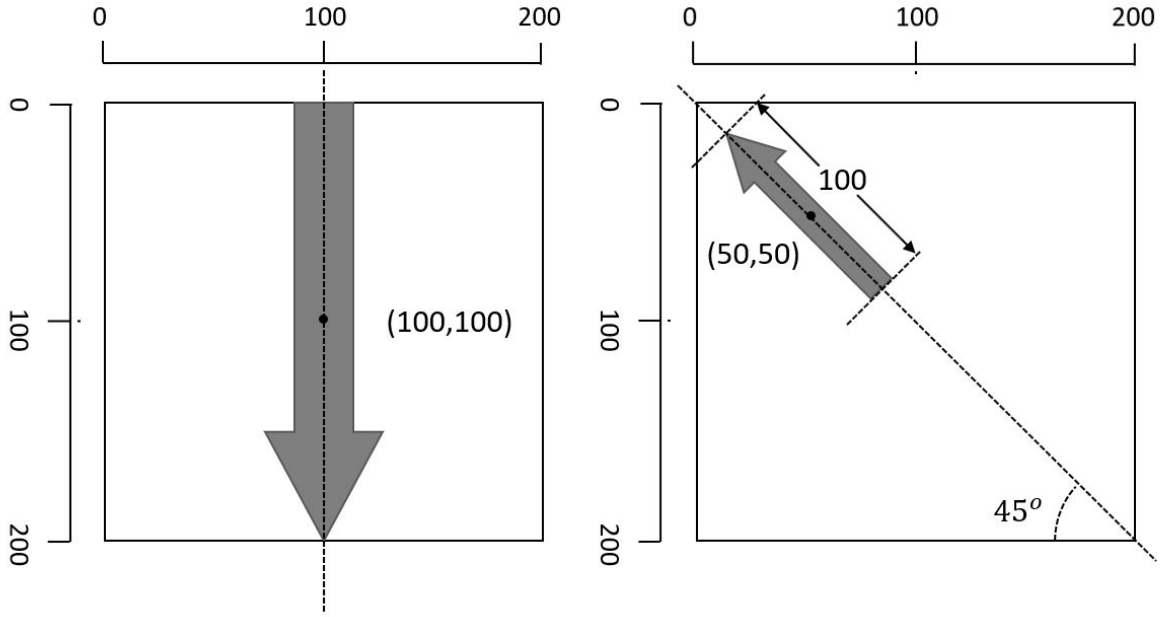
1. Permitted Time: 110 minutes
2. Open-book (textbooks, lecture notes, graded homework only)
3. Usage of computer, mobile phone, Internet is not allowed
4. Only one-line calculator is allowed
5. Grading is based on the answers written on sheets only

Problem	Point	Score
1	25	
2	25	
3	25	
4	25	
Total	100	

Problem 1: Geometric Transformation (25 pts.)

Given an input image of size 200x200 as shown in Fig. 1 (a), please apply the geometric modification technique (rotation, scaling and translation) to the input image to generate the desired output image as shown in Fig. 1 (b).

Please decompose the procedure into several steps. For each step, you should explain the required geometric manipulations and write all relevant equations. Please show all intermediate graphic results. To get full credits, you need to calculate the transformation matrices.



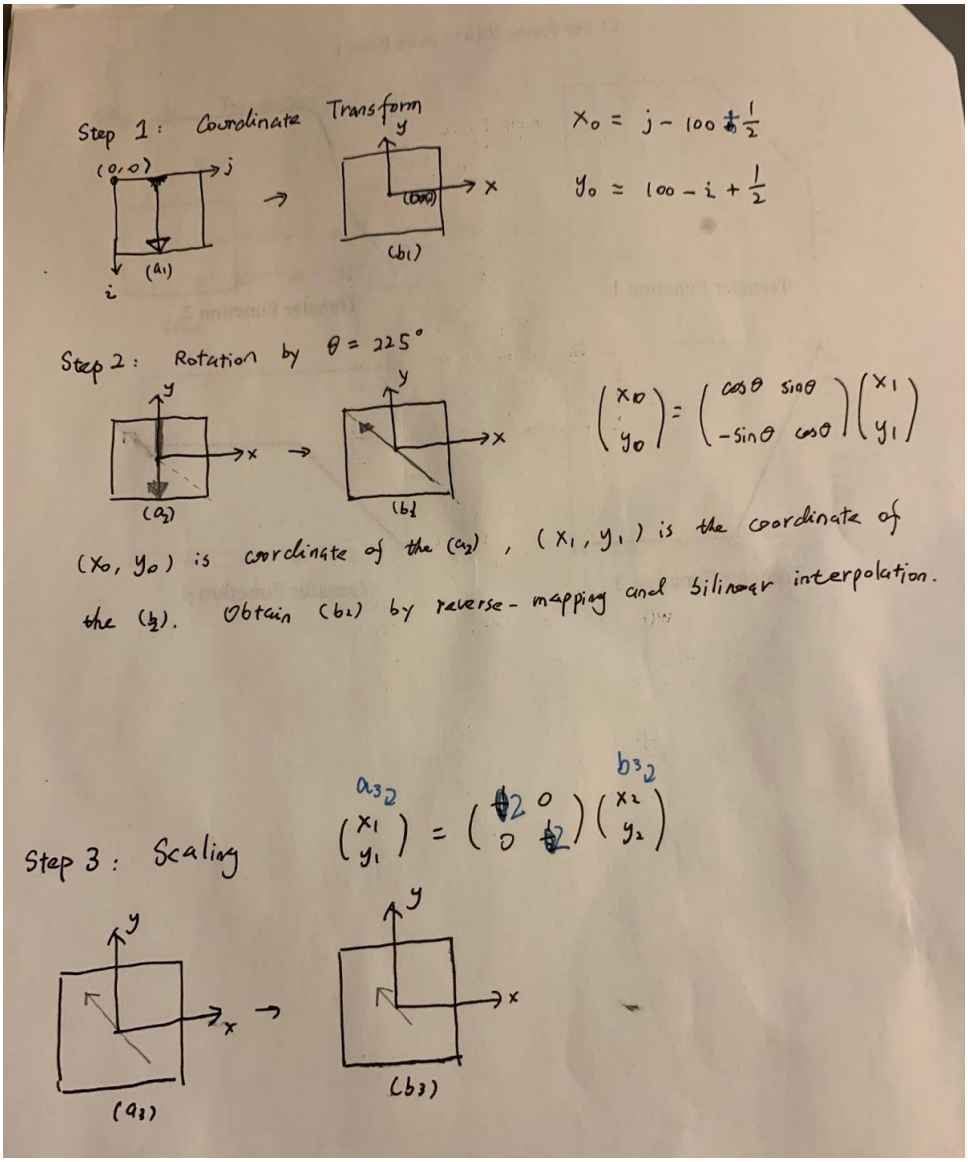
(a) (b)

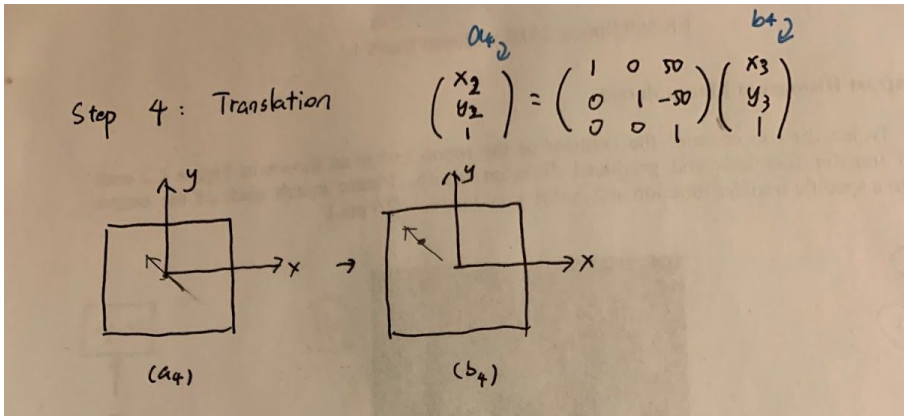
Figure 1

Grading criteria:

- We don't give 0.5 point for this problem.
- 7 points for rotation, translation scaling procedures, respectively.
  correct transformation matrix or equations: 2 points each
  correct graphic results: 2 points each
- 4 points for coordinate transform (might not necessary to change the origin location, but should have clear description of the new coordinate setting). To get full credits, you need to show correct transformation matrix or equations and clear descriptions.
- The value of transformation matrix is based on the coordinate settings. If you don't describe the coordinate setting (location of the origin, direction of the x-axis and y-axis) clearly, you can't get full credits for each procedure. Also If you don't describe the input and output of your transformation matrix clearly (forward/reverse mapping), you can't get full credits for each procedure.
- You can have different order of the transformations. But if you scale/rotate the arrow where the center of the arrow is not at (0,0), you might also translate the arrow. You will lose points, if you don't realize this.

- If you use the standard rotation matrix, the angle of rotation should be 225 degree or -135 degree.
- One possible solution:



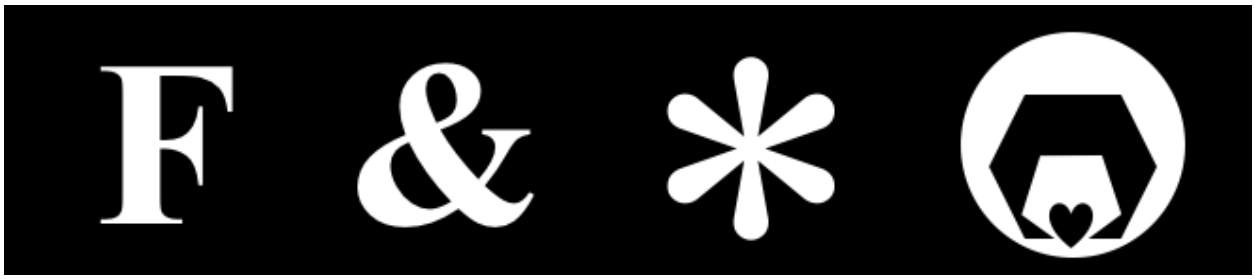


Problem 2: Morphological Processing (25 pts.)

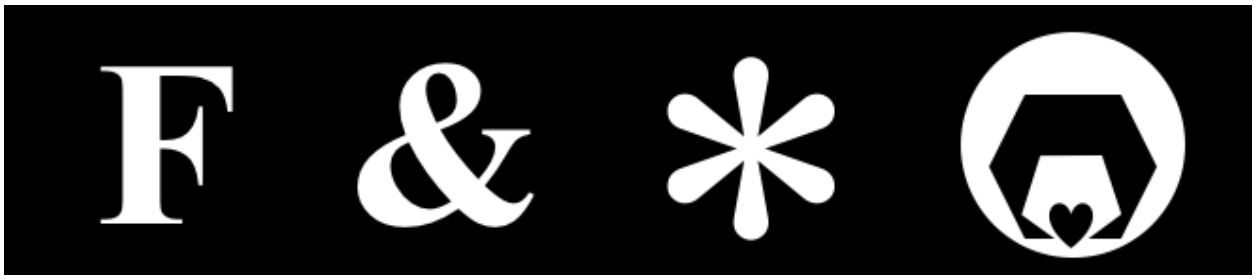
Show the final result of applying the shrinking, thinning, and skeletonizing morphological filters to the following 4 binary image patterns (white regions refer to object). Please show your results in the following 3 rows directly.

Note: If your result is a point (single pixel), please draw it visually clearly.

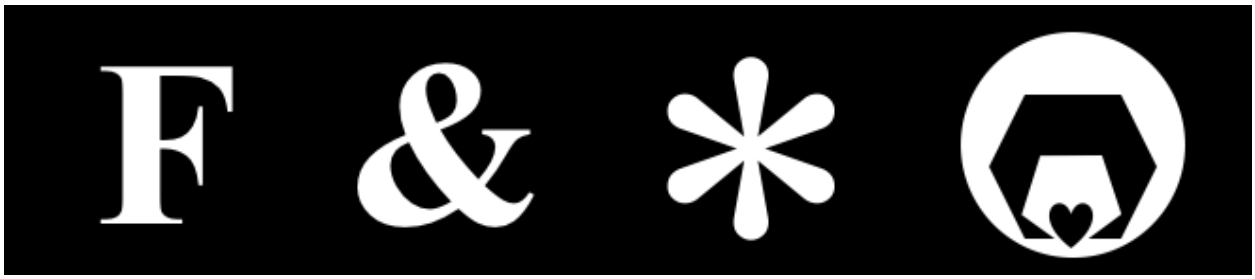
Origin



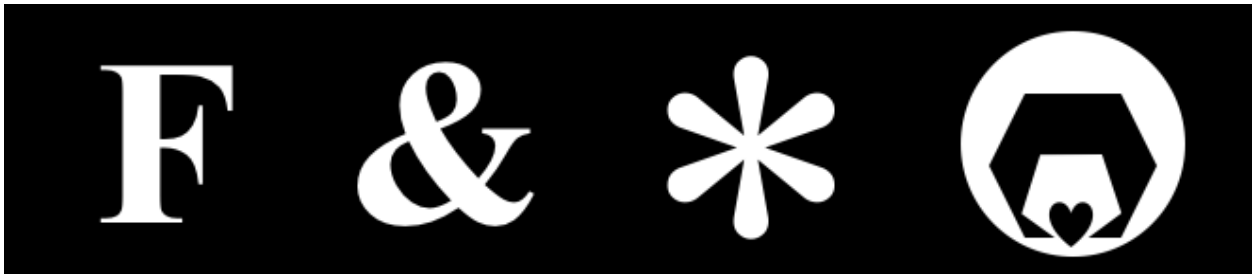
1) Shrinking



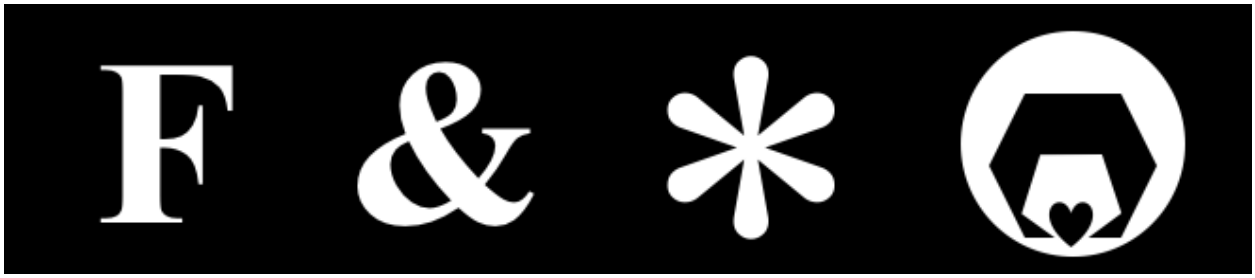
2) Thinning



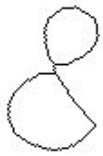
3) Skeletonizing

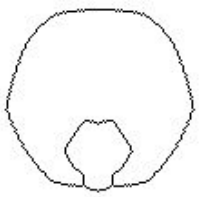


Answer:
Each correct drawing/2pts. Final score is by deduction.
Origin

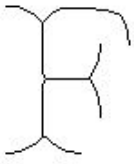


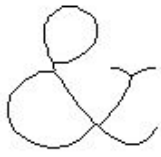
Shrinking

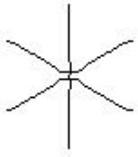


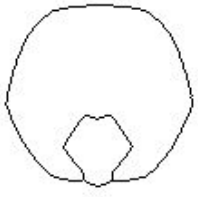


Thinning

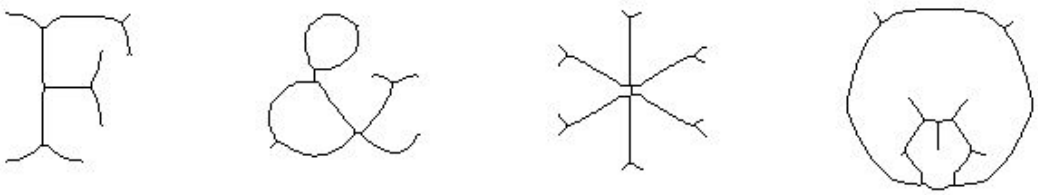








Skeletonizing



Problem 3: Texture (25 pts.)

For your reference:

Table 1: 1D Kernel for 5x5 Laws Filters

Name	Kernel
L5 (Level)	[1 4 6 4 1]
E5 (Edge)	[-1 -2 0 2 1]
S5 (Spot)	[-1 0 2 0 -1]
W5(Wave)	[-1 2 0 -2 1]
R5 (Ripple)	[1 -4 6 -4 1]

1) True-or-False questions. Write "True" or "False" beside each statement below. No justification required. (10 points, 2 points each)
    a. Results of K-means clustering can be affected by initialization and choice of distance metrics. (True)
    b. Among the 5x5 Laws' masks, E5L5 gives strongest response to vertical edges. (False, L5E5 should give higher response)
    c. L5L5 is generally expected to have little discriminant power to separate the texture images of similar luminance. (True)
    d. Both L5 and R5 are high pass filters. (False, L5 is low pass)
    e. In texture segmentation, using a larger window would generally do a better job at texture boundary than using a smaller window. (False, using smaller window size is better at boundary, larger window would generally blur the boundary)
2) Short answer questions. (15 points)
    a. Among the 5x5 Laws masks, which will give strongest response to sand images? Why? (3 points)

This question was not carefully designed. I gave full marks to all papers this time.

b. What is the difference between supervised learning and unsupervised learning? Name one supervised learning algorithm and one unsupervised learning algorithm. (3 points)

Requires labels or not (1pt)

K-means, PCA etc. (1pt)

Random Forest, SVM etc (1pt)

c. In texture segmentation, why do we subtract the local mean from each pixel before applying Laws filter bank? (2 points)

Because there can be illumination variation over the entire image, which can distract the classification.

d. What is the motivation to perform dimensionality reduction? Name one dimensionality reduction algorithm. Do we always perform better on classification using the dimension-reduced features than the original features? Why? (5 points)

(2pt) Any reasonable motivation: for example,
Data points sparse with respect to dimensionality;
Unreliable modeling;
Computationally expensive.

Method: PCA, or any other dimensionality reduction method (1pt)

No , we can lose useful information when performing dimension reduction (2pt)

e. Why do we normalize the range of all energy response features to [0,1] before K-means clustering? (2 points)

So that each feature contributes approximately proportionately when computing Euclidean distance.

Problem 4: SIFT and Bag-of-visual-words (25 pts.)

Suppose you found a way to improve SIFT and called it super-SIFT (SSIFT). The SSIFT can represent each key-point with a vector size of 2. To show that SSIFT can be equally effective as SIFT for image classification with a much shorter vector size, you may apply a k-mean clustering algorithm to the key-points generated by SSIFT with training images. For the training data, you have 5 key-points. They are represented by a SSIFT feature vector of 2 as follows:

SSIFT Feature Vectors

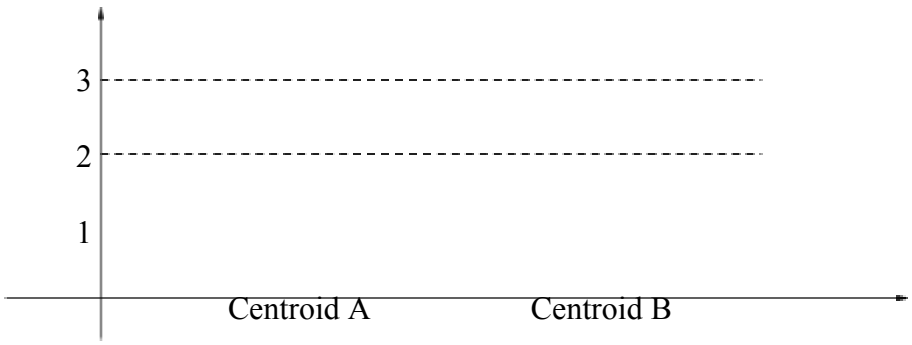
Image 1	key-point 1	[0,0]
	key-point 2	[0,0.5]
	key-point 3	[0.5,0]
Image 2	key-point 4	[1,0]
	key-point 5	[1,1]

1) You want to generate a codebook with 2 elements (classes), and your k-mean initial centroids are **CENTROID A**: [0.25, 0.25] and **CENTROID B**: [0.75, 0.75]. Use Euclidean distance (sum of squared difference), **DISREGARD** the key-point if it is equal distance to both centroids, please carry out steps of k-mean and fill in the following tables (15 pts):

STEP 1	
CENTROID A	[0.25, 0.25]
CENTROID B	[0.75, 0.75]
Closer to A	key-point:
Closer to B	key-point:
STEP 2	
CENTROID A	
CENTROID B	
Closer to A	key-point:

Closer to B	key-point:
… (steps in-between)	
STEP FINAL	
CENTROID A	
CENTROID B	
Closer to A	key-point:
Closer to B	key-point:

2) Now, you want to test an image with three SSIFT key-points: [0.5,0.5], [0,0.25], [1,1.5]. Plot the histogram. The image with more counts for centroid B is assigned to class I. Otherwise, it is assigned to class II. What class is this testing image? (5 pts)



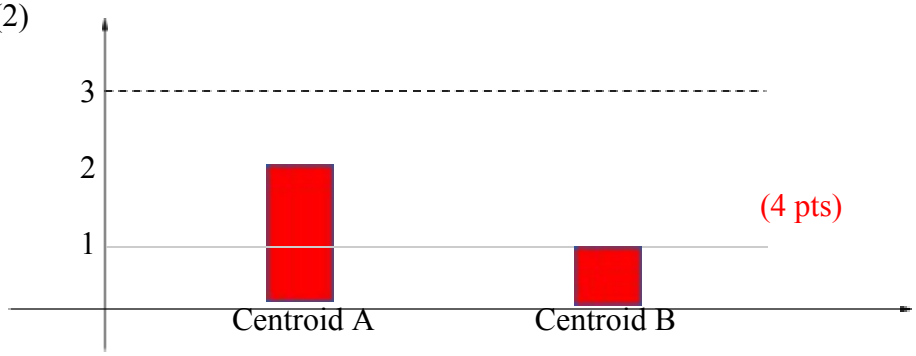
3) Suppose SSIFT has all the other properties of SIFT. If we do simple geometric modification (scaling, rotation, and translation) on the training images 1 and 2, what would be the class of the testing image in (2)? Why? Although SSIFT may save a lot of computation by having a much smaller vector size, you may still want to use the difference of Gaussian (DoG) instead of the Laplacian of Gaussian (LoG), what would be the potential advantages? (5 pts)

Sol:
(1)

STEP 1

CENTROID A	[0.25, 0.25]
CENTROID B	[0.75, 0.75]
Closer to A	key-point: 1, 2,3 (2 pts)
Closer to B	key-point: 5 (1 pts)
STEP 2	
CENTROID A	[1/6, 1/6] (2 pts)
CENTROID B	[1,1] (1 pts)
Closer to A	key-point:1,2,3,4 (2 pts)
Closer to B	key-point:5 (1 pts)
… (steps in-between)	
STEP FINAL	
CENTROID A	[3/8, 1/8] (2 pts)
CENTROID B	[1,1] (1 pts)
Closer to A	key-point: 1,2,3,4 (2 pts)
Closer to B	key-point: 5 (1 pts)

(2)



(4 pts)

ClassII (1 pts)

(3) ClassII (1 pts). Because SIFT is invariant to those geometric modifications. (2 pts). You can still have computational speed up (1 pts), DoG is scale invariant.(1 pts)