CS4670 Spring 2016

HW1 - Filtering, Sampling, Features, Homographies, RANSAC

Due: Mar 14, 2016, 9am
Homeworks to be done alone. Remember the academic integrity policies.
Solutions will be released immediately after. Late homeworks will get 0 credit.
Consult piazza for clarifications on the HW1 FAQ.

1 Image Filtering

1. "To find edges in an image we compute the gradient magnitude on the image by finite difference approximation. Then we select candidate pixels which have a gradient magnitude greater than threshold, t. Edge pixels are those candidate pixels which have a larger gradient magnitude than its neighbors perpendicular to the edge orientation." The preceding proposal is missing one of the basic but critical steps of edge detection. Describe the missing step, explain why it is necessary, and suggest a fix.

   You need to do a blur before computing edges to suppress spurious edges.

   Also accepted: At the end you would want to extend and merge long contiguous edges and suppress short spurious edges.

2. Using a central finite differences approach, derive the second order approximation of $\frac{d^2 f}{dx^2}$. (Hint: start with deriving an approximation for the first derivative, and then using that, the second derivative.)

   All are acceptable answers:

0	0	0
1	-2	1
0	0	0

0	0	0	0	0
0	0	0	0	0
1	0	-2	0	1
0	0	0	0	0
0	0	0	0	0

0	0	0	0	0
0	0	0	0	0
-1	16	-30	16	-1
0	0	0	0	0
0	0	0	0	0

3. Let a contiguous region of a 2D image I be,

..	1	5	10	3	8	6	..
..	10	5	5	1	5	6	..
..	3	2	5	2	5	6	..

Let f be a 3×3 kernel,

r	p	r
p	q	p
r	p	r

The result of convolving I with f ($I \otimes f$) gives us the following row of values:

...	2	1	11	0	...

Solve for p, q, and r. Note that you should ignore any boundary conditions and assume we are very far from the boundary of the image.

(a) p =

(b) q =

(c) r =

p = 1, q = -4, r = 0;

4. What kind of kernel does f resemble? Describe in 1 sentence.

0	1	0
1	-4	1
0	1	0

It is the Laplacian/Difference of Gaussian filter, a kind of edge sharpening filter.

2 Fourier Transform

Look closely at this satellite image — it has banding artifacts (shown in inset). The log power spectrum, shown on the right, is a visualization of the fourier transform of this image. Note: this is just a 2D extension of the frequency spectra shown in lecture for a 1D signal.

(KB: 3/1/2016:) The code to create the image on the right is:

```
np.log(
  np.abs(np.fft.fftshift(np.fft.fft2(skimage.io.imread('banding.png')/255.0,axes=(0,1))))**2
)
```

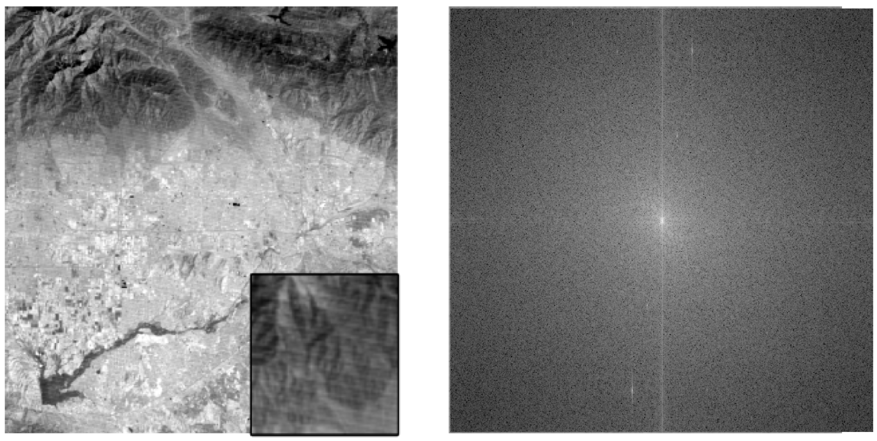
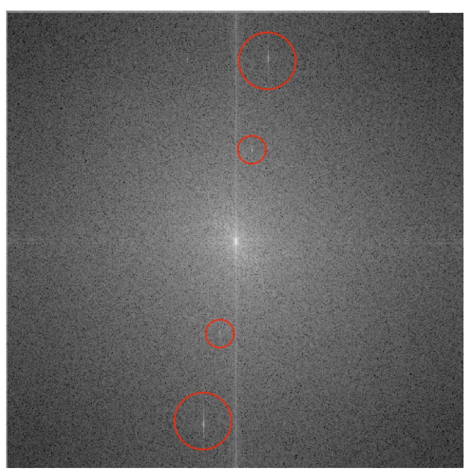


Image　　　　　　log Power Spectrum

1. Circle the parts of the image (on the right) which correspond to the banding artifacts.



Accept the 2 or 4 peaks, they must not circle the central peak or the vertical line.

2. It is not easy to design a spatial filter to remove the artifacts in the image, but it is very easy to remove the artifacts in frequency domain. How would you remove the banding artifacts given the frequency spectrum? Answer in 1 or 2 sentences.

Acceptable answers:

Identify peak centers with non-max suppression, remove peaks and inverse fft.

Blur the peaks until they match the power spectrum.

Zero the peaks.

3 Feature Matching

Consider two images I_1 and I_2 with keypoints A and B in I_1 and keypoints C and D in I_2. The keypoints are vectors in $\mathbb{R}^3$ and have the following values:

A	=	(3, 2, 1)
B	=	(7, 4, -2)

C	=	(1, 2, 3)
D	=	(5, 2, -2)

1. Compute the best matches for the keypoints in I_1 to the keypoints in I_2 using the sum of squared distance (SSD) between two points. Show your work.

   The best match for A is ______
   The best match for B is ______

   $||A - C||^2 = 4 + 0 + 4 = 8$ (Best match for A)
   $||A - D||^2 = 4 + 0 + 9 = 13$
   $||B - C||^2 = 36 + 4 + 25 = 65$
   $||B - D||^2 = 4 + 4 + 0 = 8$ (Best match for B)

2. Based solely on the SSD value of the best match, what can you say about whether the match found for A is better or worse when compared to the match found for B. Explain briefly.

   Answer: We cannot pick one match over the other just using the SSD, because their SSD are the same, indicating that the matches are of the same quality.

3. Now, let us use the ratio test to find matches. Do the evaluations of the ratio test for the problem above (show your work). Will the ratio test say that one of the matches is better than the other? If so, which of the two matches would it select as better? Explain briefly.

   $$\frac{||A - C||^2}{||A - D||^2} = \frac{8}{13}$$
   $$\frac{||B - D||^2}{||B - C||^2} = \frac{8}{65}$$

   Clearly, the latter score is better (lower) than the former score, and hence $B - D$ match is stronger (better) than the $A - C$ match.

4 Transformations

Consider the following 3 matrices, where the elements a-g are non-zero:

$$A = \begin{bmatrix} a & b & 0 \\ d & e & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$B = \begin{bmatrix} a & b & c \\ d & e & f \\ 0 & 0 & 1 \end{bmatrix}$$

$$C = \begin{bmatrix} a & b & c \\ d & e & f \\ h & g & 1 \end{bmatrix}$$

1. Classify each as one of the following: similarity transformation, affine transformation, homography (choose the most specific class possible, given that a-g can have any value except 0, and that the matrix is invertible):

   (a) A affine

   (b) B affine

   (c) C homography

   (d) A*B affine

   (e) B*C homography

   (f) C*B homography

   (g) A*C homography

2. For each of the following statements, list all matrices (A, B, and C) which would preserve the indicated property (there could be multiple matrices per statement):

   (a) angles between lines? none

   (b) lengths of line segments? none

   (c) straight lines remain straight? A, B, C

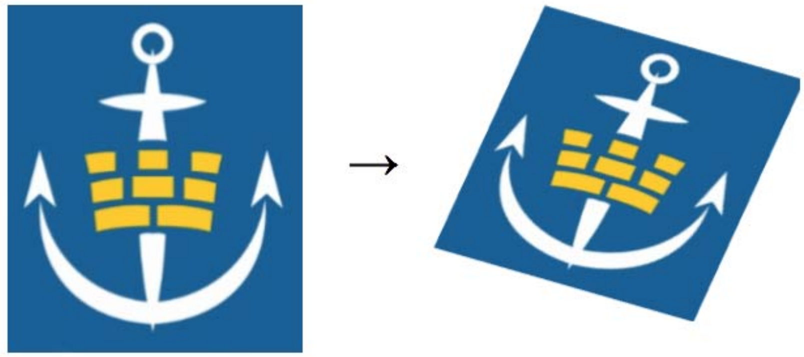
   (d) pairwise parallelism? A, B

   (e) the origin remains at (0,0)? A

3. Which matrix (or matrices) could potentially perform the image transformation (there could be multiple answers) shown in the Figure?

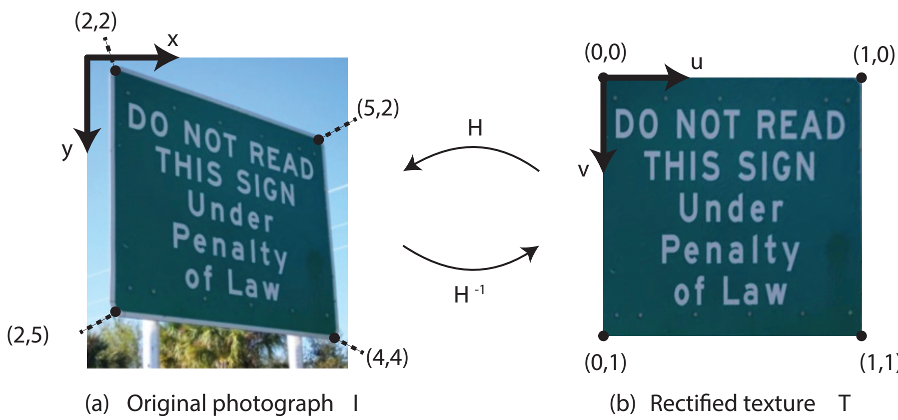
   Best answer: B

   A is acceptable if you state that you are assuming there is no translation component.



(coordinates not to scale)

5 Homography



(coordinates not to scale)

Given our input photograph I, we would like to "rectify" it to obtain a rectified square texture T. To do this, we need to know the mapping between points (u, v) in the rectified texture and points (x, y) in the original image. In general, this mapping is a homography H and can be written in the form:

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \lambda \underbrace{\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & 1 \end{bmatrix}}_{H} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix}$$

In this equation, (u, v) is a point in the rectified texture coordinates, and (x, y) is its corresponding point in the input image. Both points are expanded to homogenous coordinates by appending an extra coordinate. Recall that for homogenous coordinates, we need to introduce an arbitrary (nonzero) scalar λ that represents the arbitrary scale. The λ is different for different u, v values, and is necessary since homogenous coordinates can be scaled without changing the point that they represent.

We will solve for H by using correspondences between the two images.

1. Say that we know a single point correspondence, i.e., we know one point in I and its corresponding point in T. How many constraints does this give to solve for H?

   2

2. In general, how many total constraints are needed to solve for H?

   8

3. Here are most of the values in H. Using point correspondences, solve for the remaining entries a, b, d, and e:

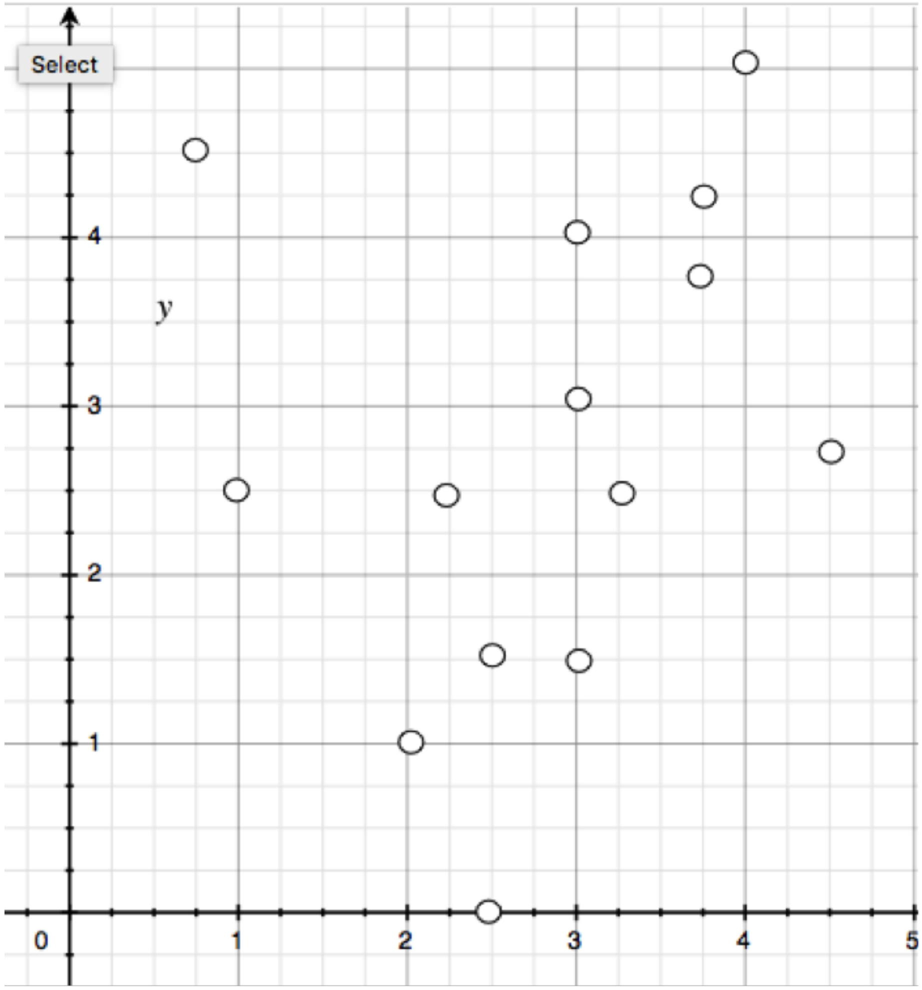
$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \lambda \underbrace{\begin{bmatrix} a & b & 2 \\ d & e & 2 \\ 1 & 1 & 1 \end{bmatrix}}_{H} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix}$$

   ANS: $a = e = 8$, $b = d = 2$

4. Given the point $(u, v) = (0.5, 0.5)$ in the rectified image, compute its corresponding point in the original image, using your homography from part 3.

   (3.5, 3.5)

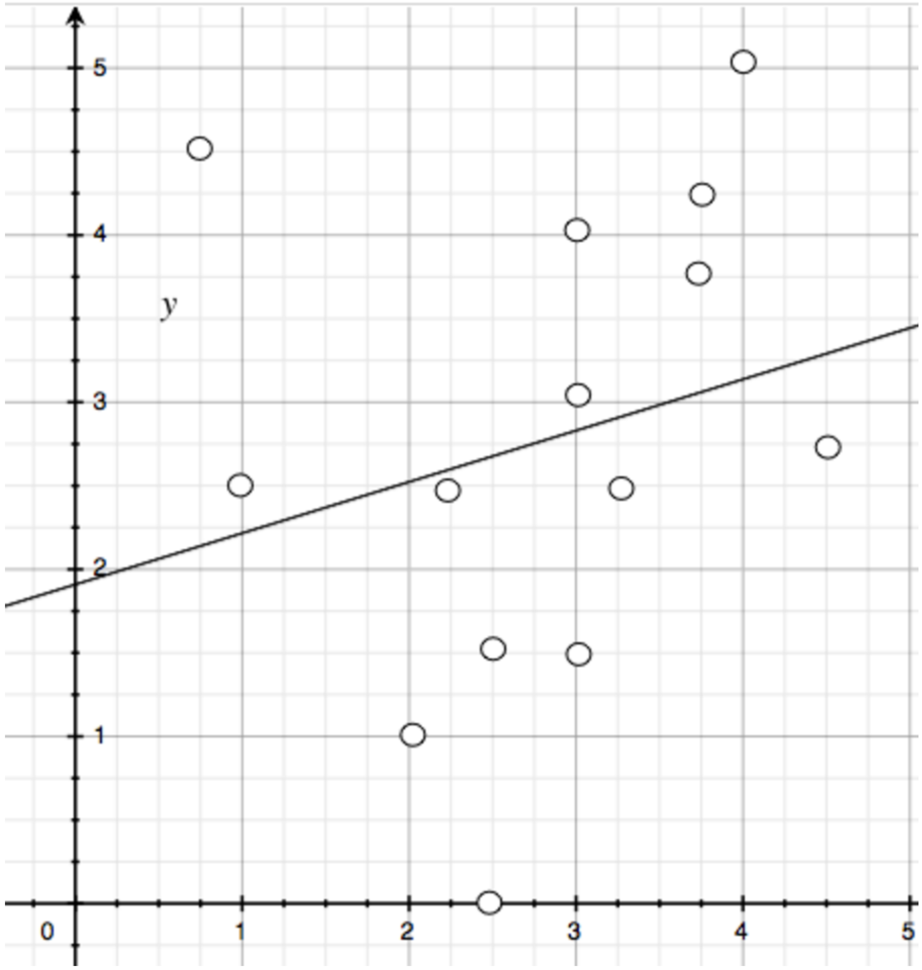
6 RANSAC



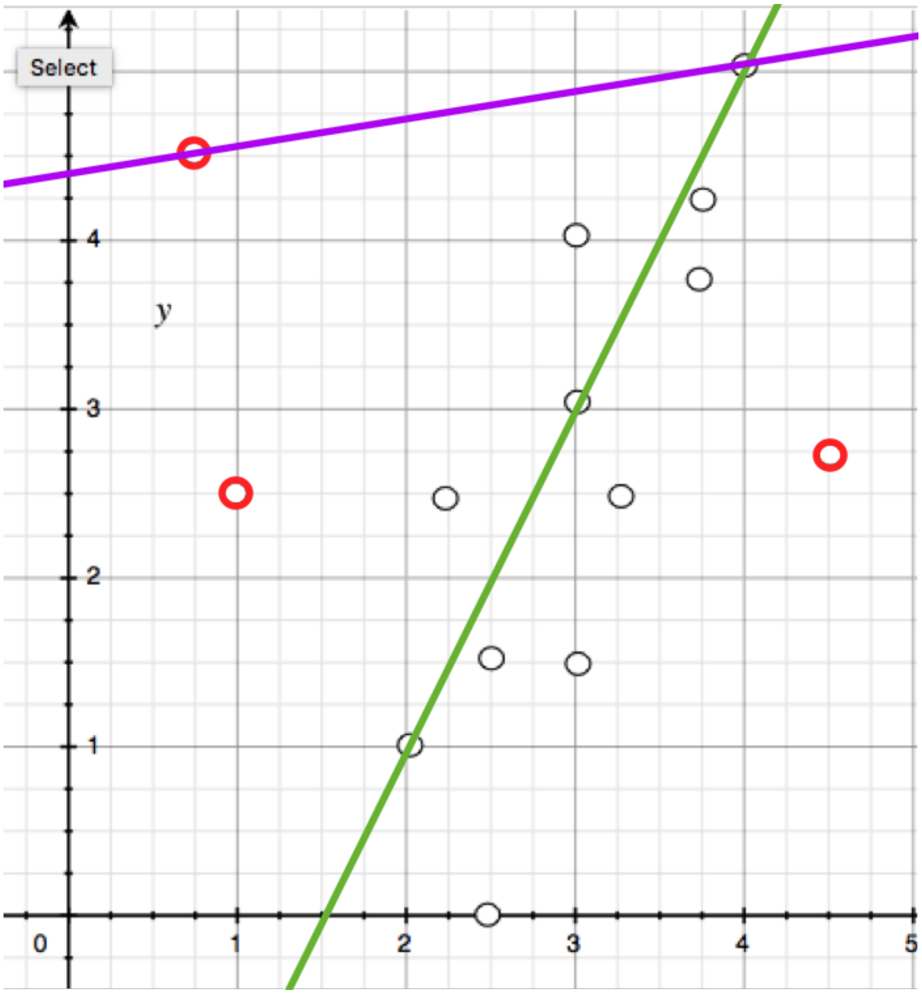
Consider a set of measurements as shown in the figure. (Note: you will probably have to write some code for this part.)

1. What line is produced by using a simple linear regression? Give the exact model in $y = mx + b$ form and draw it on the diagram.

   y=0.30698365527489x + 1.90720653789



2. Perform the RANSAC algorithm on this set of points to find the best fit line to the data. Assume the inlier threshold is 1 unit. Give an example of a model found by RANSAC with a high outlier percentage and one with a low outlier percentage. Draw these on the diagram as well. What is the best fit line you can find?



The outliers are circled in red, a line with a low outlier percentage line is in green, and a high outlier percentage is in purple.

The best fit line should be approximately the line y = 2x - 3.

Examples of good lines are the best fit line (shown in green), or anything that essentially gets the central mass of points.

A bad example is a line that uses any of the outliers together, or just anything that "looks bad"(example is shown in purple).

3. Given the number of inliers and outliers found from your model, what is the minimum number of trials necessary to produce a best fit line with 99% confidence? $(1-0.78^2)^n < 0.01 \rightarrow n * log(1-0.78^2) < log(0.01) \rightarrow n > log(0.01)/log(1-0.78^2) \rightarrow$ 5 trials