University of Pennsylvania
CIS581: Computer Vision and Computational Photography
Midterm Examination
November 30th, 2017

You must take this exam independently, without assistance from anyone else. You may bring in one two-sided letter or A4 sheet of notes for reference (or two one-sided sheet of notes). Aside from these notes, you may not use a calculator and may not consult any outside references, such as a textbook or the Internet. Any suspected violations of Penn's Code of Academic Integrity will be reported to the Office of Student Conduct for investigation.

This exam consists of several problems. We recommend you to look at all the problems before starting to work. If you need clarifications on any of the questions, please feel free to ask the instructors. When you work out each problem, please show all the steps and box your answers (if applicable). The exam is worth a total of 100 points, and partial credit will be awarded for the correct approach even when you do not arrive at the correct answer.

Question 1	30 points
Question 2	20 points
Question 3	15 points
Question 4	20 points
Question 5	15 points
Total	100 points

There are two blank pages at the end of this exam. Please make use of them for rough work or to write answers if you've run out of the space provided below each question. If it's the latter, please ensure that you clearly write which question the answer corresponds to.

IMPORTANT NOTE: Certain questions have **(Optional)** written next to them. These are questions for extra credit and cannot be substituted for mandatory ones within a section. Please attempt them only if you have extra time. You MUST attempt questions that do not have **(Optional)** written next to them.

I agree to abide by the University of Pennsylvania Code of Academic Integrity during this exam. I pledge that all work is my own and has been completed without the use of unauthorized aid or materials.

- Name:
- PennKey:
- Signature:

1 Cameras, Convolutions and Edges - 30 Points

1.1 Convolutions

1. **(5 points)** Convolve an image, I and kernel, K to obtain an output image, F.

$$F = I \otimes K$$

$$I = \begin{bmatrix} 2 & 0 & 1 \\ 1 & -1 & 2 \end{bmatrix}, K = \begin{bmatrix} 1 & -1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

The size of the output should be 4×5.
Assume, **zero-padding** where necessary.
Important: Please show all necessary steps to receive full credit.

2. **(4 points)** In this question, we will count the *exact* number of multiplications (including the multiplications due to zero-padding) and additions involved in the previous question.
(i) What is the number of multiplications and additions to compute $F = I \otimes K$?
(ii) What is the number of multiplications and additions to compute $F = K \otimes I$?

1.2 Edge Detection

1. **(5 points)** Compute the first derivative of an image, I by convolving it with the kernel, K to obtain the output F.

$$F = I \otimes K$$

Assume, **zero-padding** where necessary.

$$I = \begin{bmatrix} 56 & 64 & 79 & 98 & 115 & 126 & 132 & 133 \end{bmatrix}$$
$$K = \begin{bmatrix} 1 & -1 \end{bmatrix}$$

Note: Do not compute a value for the first and last image pixel.
(i) Show the result of the convolution.
(ii) Indicate where an edge would be detected and why.
Important: Please show all necessary steps to receive full credit.

2. **(5 points)** For each of the following properties of edge detectors, indicate whether the Canny edge detector or the Laplacian-of-Gaussian detector is better with respect to the property, and explain briefly why:
(i) Fewer number of parameters that must be set

(ii) Closed chains of edges detected

(iii) Computes the edge orientation

(iv) Better localization of the true edge position

(v) Less likely to round corners where the boundary curvature is high.

3. **(Optional) (1 point)** What is the purpose of:
(i) non-maxima suppression and of
(ii) hysteresis that are done in the Canny edge detector?

4. **(Optional) (1.5 point)** What property of the coefficients of a kernel ensures that an appropriate output is obtained for regions of constant intensity in an image when:
(i) The kernel is approximating a first derivative.

(ii) The kernel is approximating a second derivative.

(iii)The kernel is approximating a Gaussian.

5. **(Optional) (1 point)** In Canny edge detection, we will get more discontinuous edges if we make the following change to the hysteresis thresholding:
(a) increase the high threshold
(b) decrease the high threshold
(c) increase the low threshold
(d) decrease the low threshold
(e) decrease both thresholds

1.3 Camera Models

The relationship between a 3D point at world coordinates (X,Y,Z) and its corresponding 2D pixel at image coordinates (u,v) can be defined as a projective transformation using a 3×4 camera projection matrix $\mathbf{P}$.

1. **(5 points)** A scene point at coordinates $(400,600,1200)$ is perspectively projected into an image at coordinates $(24,36)$, where both coordinates are given in millimeters in the camera coordinate frame and the camera's principal point is at coordinates $(0,0,f)$ (i.e., $u_0 = 0$ and $v_0 = 0$).
   Assuming the aspect ratio of the pixels in the camera is 1, **what is the focal length of the camera** f? (Note: the aspect ratio is defined as the ratio between the width and the height of a pixel; i.e., ku/kv)

2. **(4 points)** Given a 3×4 camera matrix, $P = \begin{bmatrix} 5 & -14 & 2 & 17 \\ -10 & -5 & -10 & 50 \\ 10 & 2 & -11 & 19 \end{bmatrix}$ and a 3D point in Cartesian coordinates $\mathbf{X} = \begin{bmatrix} 0 & 2 & 2 \end{bmatrix}^T$:
(i) What are the homogeneous coordinates of the point $\mathbf{X}$ in 3D?

(ii) What are the Cartesian image coordinates, (u, v), of the projection of $\mathbf{X}$?

3. **(2 points)** An ideal pinhole camera has focal length 5mm. Each pixel is 0.02 mm x 0.02 mm and the image principal point is at pixel $(500, 500)$. Pixel coordinates start at $(0, 0)$ in the upper-left corner of the image.
What is the 3 x 3 camera calibration matrix, **K**, for this camera configuration?

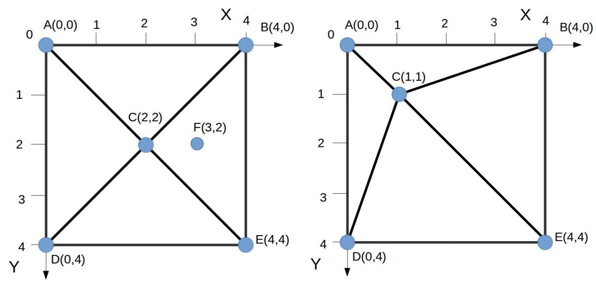
1.4 (Optional) (2 points)

Frame a question on convolutions or edges or camera models different from the ones given above. Solve it yourself in as much detail as you wish to.

2 Morphing and Blending - 20 Points

2.1 Image Morphing

1. Suppose you are given two images as in Figure 1. ABCED are the corresponding points clicked as below. For a point $F(3,2)$ in image 1 ($t = 0.5$), compute the coordinates of F' in source image 2 ($t = 0$) and image 3($t = 1$). (t controls the rate of morphing)



(a) Image 1, t=0.5 (b) Image 2, t= 0

Figure 1: Image Morphing

(a) **(5 points)** Compute the barycentric coordinates of F.

(b) **(5 points)** Compute the coordinates of the point F in image 2 and image 3.

2.2 Image Blending

1. **(10 points)** You are given two images, a source image and a target image. The mask is denoted by gray pixels on both the images. Compute the pixel values of blended image (**just the masked region**) using Gradient Domain Blending.

1	2	3	2	1
2	3	2	3	2
4	2	3	2	3

(a) Source Image

3	1	2	0	2
0	0	1	2	3
1	0	3	2	1

(b) Target Image

2.3 (Optional) (2 points)

Frame a question on Image Morphing or blending and should be different from the ones given above. Solve it yourself in as much detail as you wish to.

3 Image Transformations - 15 points

3.1 Affine Transforms

1. As shown in Figure 3, an Image A is warped to image B using the affine transform given below:

$$\begin{bmatrix} x_B \\ y_B \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0.25 & 1.5 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_A \\ y_A \\ 1 \end{bmatrix}$$

(i) **(5 points)** Compute the coordinates of image B and sketch it.

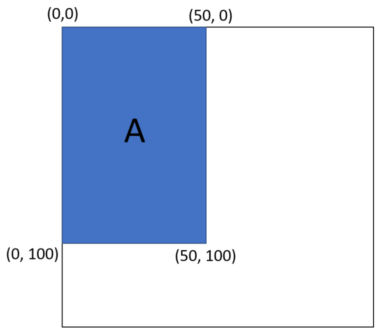


Figure 3: Image A

(ii) **(1 point)** How many degrees of freedom does an affine transformation have?

(iii) **(3 points)** Which of the following does an affine transformation preserve?

- Orientation
- Lengths
- Angles
- Parallelism
- Straight Lines

3.2 Homographies

1. **(3 points)** A homography H is used to transform image 1 into image 2. While warping, do we use $p = Hp'$ or $p' = Hp$? p and p' are homogenous coordinates in image 1 and image 2 respectively. Pick one and explain why so briefly.

2. **(3 points)** Can we estimate homography using more than four non-collinear points? If so, why would we do that?

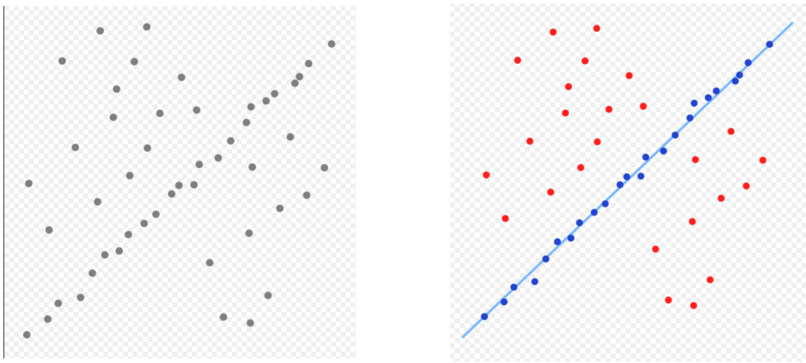
3.3 (Optional) (2 points)

Frame a question on Image Transformations different from the ones given above. Solve it yourself in as much detail as you wish to.

4 Random Sample Consensus - 20 points

4.1 RANSAC and Model Fitting

You are given a 2-D point cloud and are supposed to fit an optimal line model to it. Here, we decide to use RANSAC (Random Sample Consensus) to estimate the line and assign the points in the raw data (Figure (a)) to two different categories as in Figure (b).



(a) Original point cloud (b) Line model after RANSAC

1. **(15 points)** Describe how you will use RANSAC for the aforementioned task. It could either be a paragraph or in terms of pseudo-code.
   You can assume that you have a function to estimate a line, and a function to compute the distance of a point to the line.

2. **(5 points)** Assume that there are 60% inliers, compute the minimal number of RANSAC iterations needed to get, with probability 95%, at least one random sample that is free from outliers. (Feel free to use the fact: $\log_{10}^{2} = 0.30$, $\log_{10}^{3} = 0.48$, $\log_{10}^{5} = 0.70$)

4.2 (Optional) (2 points)

Frame a question on feature detection, or matching, or RANSAC or image stitching different from the ones given above.
Solve it yourself in as much detail as required.

5 Seam Carving - 15 points

5.1 (15 points)

In image carving, a vertical seam is removed from the image, by computing the minimum energy path using dynamic programming. The energy map E is shown below:

E =

6	9	0.5	2	2.5	8	3
2	1	2	0.5	5	6	4
3	0.5	4	7.25	2.5	5	7
2	2.5	3.5	8	6.5	4	3.5

Recall that dynamic programming constructs matrix M recursively, i.e.,

$$M(i,j) = E(i,j) + min(M(i-1,j-1), M(i-1,j), M(i-1,j+1))$$

and $T(i,j) \in \{1,2,3\}$ records the path choice of "left", "center", and "right" from previous row.

Finish the cumulative minimum table M and the backtrack table T, and find the vertical seam to be removed.

M=

6	9	0.5	2	2.5	8	3

T =

0	0	0	0	0	0	0

Note: Please stick to the conventions described in the question.

5.2 (Optional) (2 points)

Frame a question on seam carving or seam insertion or optical flow or SIFT features different from the ones given above.
Solve it yourself in as much detail as required.

Rough Page

Rough Page