Princeton COS 429

Midterm

10/26/17

Name: _____

netID: _____

Pledge in full: _____

_____

_____

Signature: _____

Multiple choice (28 points)

1. **History (1 point).** The field of computer vision started:
    A. As an undergraduate summer project in the 1960s
    B. As an undergraduate summer project in the 1990s
    C. As a PhD student project in the 1960s
    D. As a PhD student project in the 1990s

2. **Aperture (2 points).** Decreasing the aperture of a camera:
    A. Increases the depth of field
    B. Decreases the depth of field
    C. Does not affect the depth of field

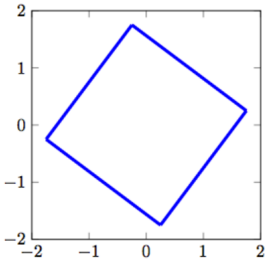
3. **Gaussian filter (2 points).** Consider a 1-D Gaussian with standard deviation of 2 pixels. Which of the following is the *most efficient* kernel that could capture at least 99% of this distribution?
    A. Filter kernel of size 1×3
    B. Filter kernel of size 1×7
    C. Filter kernel of size 1×13
    D. Filter kernel of size 1×23

4. **SIFT pipeline (2 points).** Which of the following is *not* a part of the SIFT calculation pipeline?
    A. Difference of Gaussians
    B. Histogram of Gradients
    C. Laplacian of Gaussians
    D. Orientation Histogram
    E. Image Pyramid

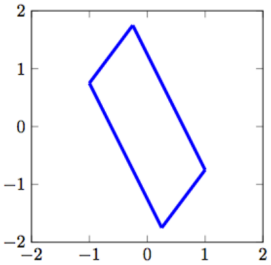
5. **SIFT robustness (2 points).** What is the SIFT feature least robust against?
    A. Affine illumination
    B. Rotation
    C. Scale
    D. Occlusion

6. **RANSAC vs Least squares (3 points).** What are the benefits of RANSAC compared to the least squares method? *Circle all that apply.*
    A. RANSAC is faster to compute in all cases
    B. RANSAC has a closed form solution
    C. RANSAC is more robust to outliers
    D. RANSAC handles measurements with small Gaussian noise better

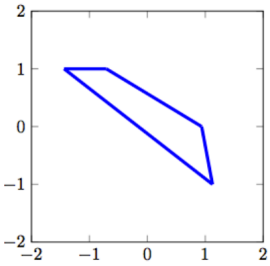
7. **Image Transformations (6 points).** The figures below show the outputs of applying one of the transformations on the right to a square with vertices at (1,1), (1, 0), (0, 0), and (0, 1). Circle the *most specific transformation* that could generate each output image.

(a) 

Projective Affine Similarity Rigid Body

(b) 

Projective Affine Similarity Rigid Body

(c) 

Projective Affine Similarity Rigid Body

8. **Affine transformation (3 points).** Which of the following always hold(s) under an affine transformation? *Circle all that apply.*
   A. Parallel lines will remain parallel .
   B. The ratio between two areas will remain the same .
   C. Perpendicular lines will remain perpendicular.
   D. The angle between two line segments will remain the same.

9. **Rotational invariance (3 points).** Which of the following representations of an image region are rotation invariant? *Circle all that apply.*
   A. Bag of words model with SIFT features
   B. Spatial Pyramid Model with SIFT features
   C. A HOG template model
   D. Deformable Parts Model

10. **Bag of words (3 points).** Which of the following are drawbacks of bag of words (BOW) models? *Circle all that apply.*
   A. They don't capture spatial information
   B. They are not ideal for solving detection problems
   C. Creating BOW features is time consuming
   D. They are hard to understand because their inner workings are opaque

11. **Object categories (1 point).** As mentioned in lecture, most researchers working on object recognition would be familiar with the PASCAL VOC dataset. How many object categories are in PASCAL VOC?

    A. 1

    B. 10

    C. 20

    D. 101

Short answer (72 points)

Instructions: for each question, write your entire answer within the blank box immediately below the question. Unless otherwise specified, anything outside of the corresponding box will not be graded. You may use the margins or the last page of the exam for scratch space.

1. **Filters (8 points).** Give a 3x3 example of each of the following types of image filters:

    a. A brightening filter (only increases the image intensity)

    b. A smoothing filter

    c. A sharpening filter

    d. A filter for detecting a vertical edge

2. **Convolution and Filtering (7 points).**

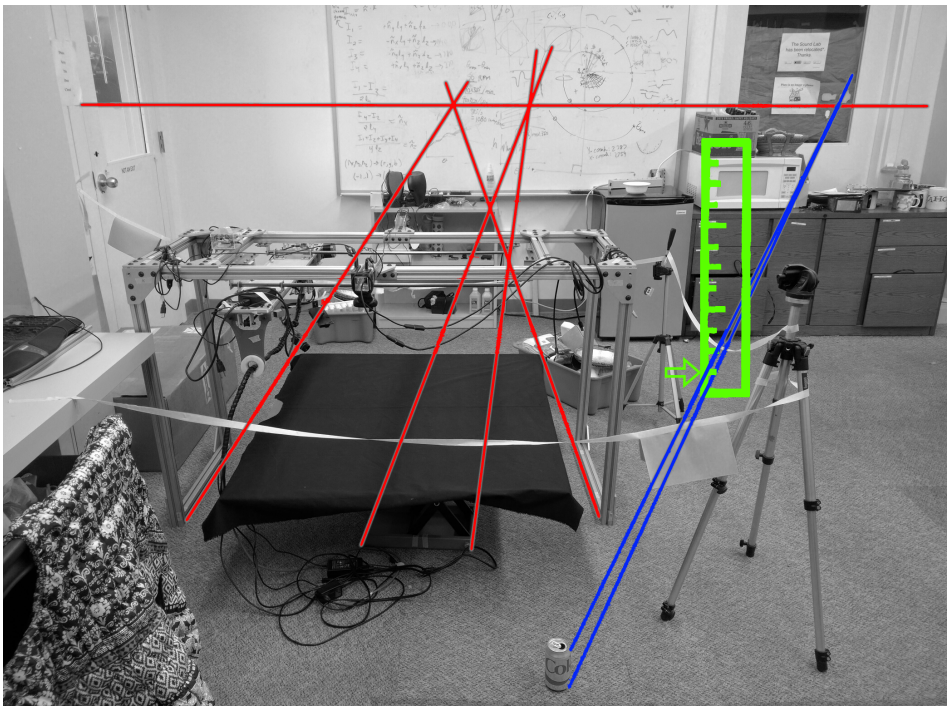
Recall the definition of a convolution of an image f with filter g is:

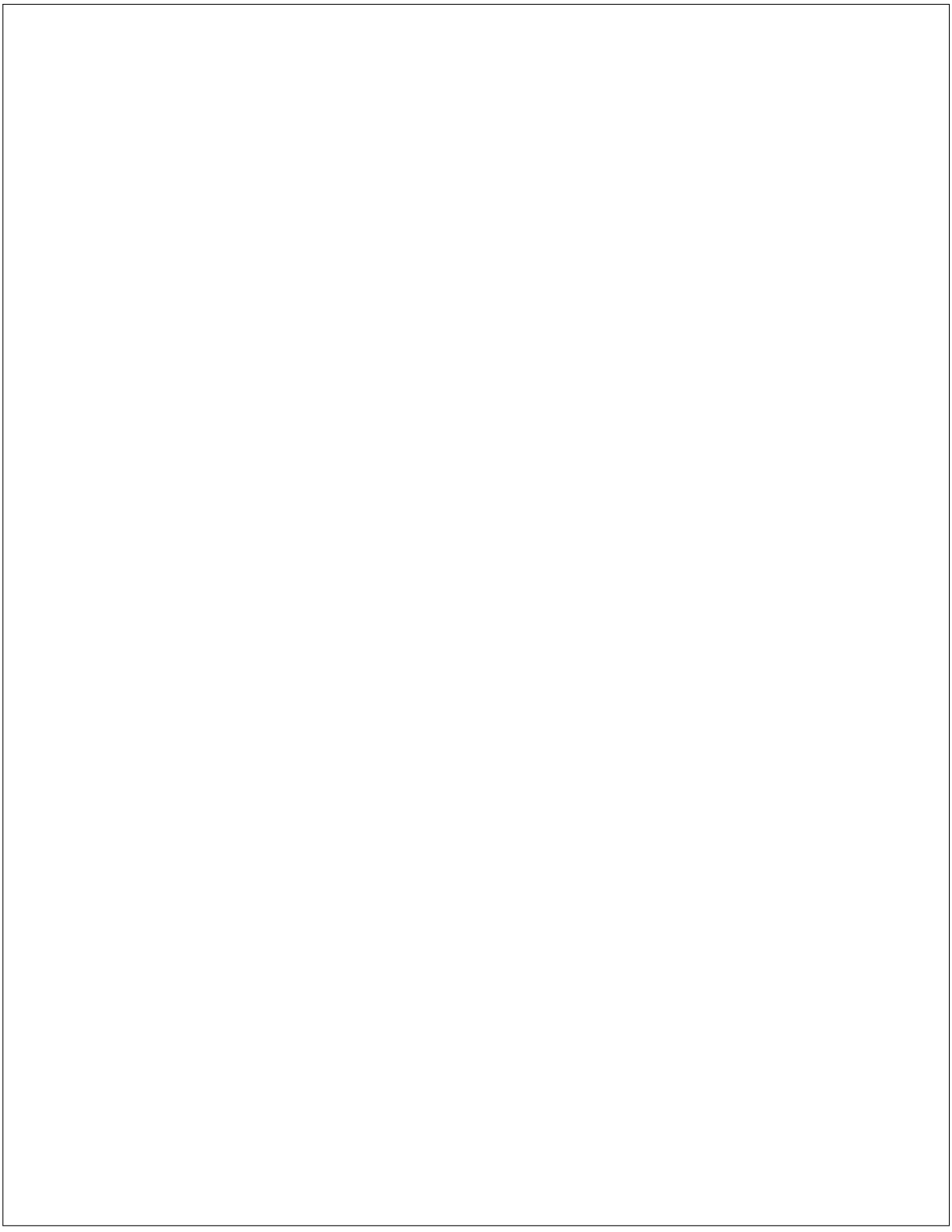
$$(f * g)(u) = \sum_x f(u - x)g(x)$$

Prove that convolution with the unit impulse filter will leave the image unchanged.

3. **Image Formation (8 points).** Your friend shows you the following image, hoping to find out exactly how tall the 3D scanner is (the scanner is the large metal frame). The scanner is custom-built, so you can't just look the answer up. Remembering your COS429 training, a quick internet search reveals that a coke can is 5 inches tall. How tall is the scanner? Show your work for full credit.

   Instructions: for this question, there are two boxes (one on the image, one on the next page) which can be used for your answer.





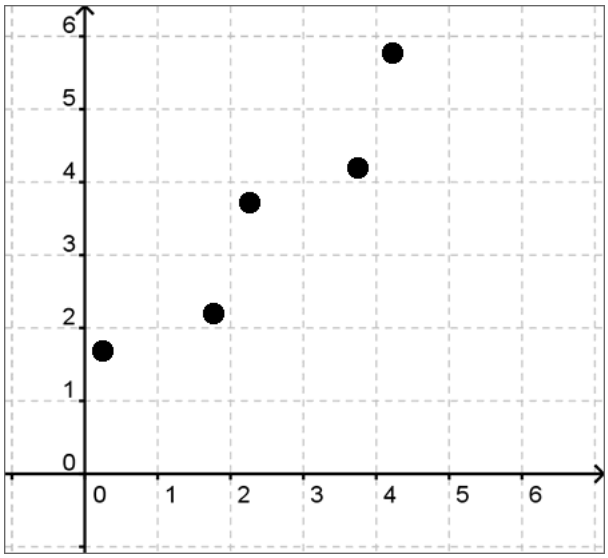
4. **Feature Detectors and Descriptors (6 points).** For each of the image locations A, B, and C, (they are marked with small blobs) what will the eigenvalues of the second moment matrix be? Describe in words how large each of the values will be and their relationship to one another.

A

B

C

5. **Fitting, Hough Transforms, and RANSAC (20 points).** Suppose that we have a set of 2D point observations, and we want to fit a line to them. However, we already know the right slope: 1. Our observations are below.



a. What is the equation of our model? (2 points)

b. How many observations do we need to fit a model? (2 points)

c. Suppose (for the remaining parts of this question) we decide to apply the Hough transform to find the right model. How many models does each observation vote for? What shape does this take on in Hough transform space? (3 points)

d. Suppose we use bins of size 2, with bins $[0, 2), [2, 4)$, and so on. What line will we fit? How many inliers and outliers will there be? (4 points)

e. Suppose instead our bins are of size 1, with bins $[0, 1), [1, 2)$, and so on. What line will we fit? How many inliers and outliers will there be? (2 points)

f. Discuss one way to make the Hough transform method more robust against rounding issues that arise with smaller bin sizes. (Hint: recall *splatting* from the image alignment lecture). (7 points)

6. **Recognition (23 points)**

a. Briefly describe 3 reasons why template matching of RGB values does not work for object recognition in general. (6 points)

b. Given an example of a car, why does matching SIFT keypoints to this example not work for recognizing other cars? Briefly describe 2 reasons. (4 points)

c. Give 1 example when a color histogram works well as an object descriptor. Briefly describe 3 situations when it does not work well. For each of the 3 situations, describe an approach discussed in class that was tried to improve on this weakness. (13 points)

END OF EXAM

PAGE INTENTIONALLY BLANK
Use for scratch work or for extra space