

# Practice Midterm 2

## CMSC 426

The midterm will cover material up to and including stereo geometry (but not stereo matching). The midterm will be cumulative, covering everything up to this point, including material covered in the first midterm. However, there may be an extra emphasis on material not covered in the midterm. The previous practice midterm can serve as a review for that material; here I'll only discuss new topics. But you should think of both practice midterms as providing preparation for the second midterm.

Here are some topics it will be helpful to master for the midterm:

- Interactive Segmentation with Min Cut
  - How you can represent image as a graph.
  - Understand how to create edge weights.
- Background subtraction and Kernel Density Estimation (KDE)
  - The main point here is to understand how to use samples to create a probability distribution. This is also relevant to image segmentation using mincut. It includes:
    - How to form a histogram, and use this to estimate probabilities.
    - How smoothing this with a Gaussian (or equivalently, placing a Gaussian at each sample point) creates a smoother distribution (this is KDE).
- Texture
  - Understand what is meant by a Markov process.
  - Understand how SSD can be used to find locations in a sample that match a template.
  - Understand the overall structure of the Efros and Leung synthesis algorithm.
- Mosaicing
  - Blob detection. What does the filter for a blob detector look like? How is non-maximum suppression done in blob detection (the filter output must be larger than at neighboring locations and scales).
  - SIFT descriptors. How are they constructed?
  - RANSAC. RANSAC for detecting lines, as an example. RANSAC as used in the mosaicking project.
- Geometry
  - 2D Transformations. How to represent rotation, translation, and scaling using matrices. How to represent affine transformations. How to solve for a similarity or affine transformation using a matrix inverse.
  - Perspective Projection. How 3D points turn into 2D image points. Vanishing points and the horizon.
  - The Epipolar Constraint. Depth from 2 images. Rectification.

**Practice:** The goal of this is to give you samples of the sorts of questions and topics that will come up in the midterm. Some of these questions may be a bit more involved or more vague than those that I would ask in a real midterm. It is very likely that there will be at least one question on the midterm that is quite similar to a practice question.

1. Kernel Density Estimation: Suppose that four pixels in an image are known to be part of the background. They have intensities of 3, 7, 3 and 12.
  - a. Build a histogram of these intensities. Using the histogram, what would you estimate is the probability that a new background pixel will have an intensity of 5? What about an intensity of 3?

The histogram just shows two values at 3, and one value at 7 and one at 12. The probability of getting a pixel with any other intensity will be zero. The probability of getting an intensity of 3 will be  $\frac{1}{2}$ .

- b. Use KDE, with a Gaussian that has a sigma of 1, to estimate the probability that a new background pixel will have an intensity of 5.

Let  $G(x; m, s)$  represent the value of a Gaussian distribution at point  $x$ , assuming that the Gaussian has a mean of  $m$  and a standard deviation of  $s$ . Then using KDE, we get:

$$2G(5; 3, 1) + G(5; 7, 1) + G(5; 12, 1).$$

This can be simplified as:

$$3G(2; 0, 1) + G(7; 0, 1)$$

2. Markov Models: Explain how you could use a sample of text, such as “the rain in spain falls mainly in the plane” to generate a new text string, using different order Markov models, following the example of Shannon.

With a 0<sup>th</sup> order Markov model we would choose letters randomly, but with the same frequency as in the text. We could do this as:

```
>> t1 = 'the rain in spain falls mainly in the plane'
```

```
t1 =
```

```
the rain in spain falls mainly in the plane
```

```
>> t1(randi(length(t1), [1, 50]))
```

```
ans =
```

```
hhlnlpi nla sll n t tplslnllnp tian ihsnnhnn
```

With a 1<sup>st</sup> order Markov model, similar to the texture synthesis project, we start with a random letter, then take the last letter of the generated text and randomly find a location in the sample text where this letter occurs, and take the next letter. For example:

```
>> txt = t1(randi(43)); for i = 1:50 last=txt(end); matches=find(t1(1:end-1)==last);ind=matches(randi(length(matches)));txt=[txt,t1(ind+1)]; end
>> txt
```

txt =

in main ine fain s splly the fan in in falainlain

3. RANSAC: Suppose an image contains  $n$  points. You wish to find a square that matches as many points as possible. We say that a square matches a point if the point is within 3 pixels of one of the sides of the square. Explain as precisely as you can how you would do this using RANSAC.

To do this, you have to randomly sample  $k$  points, and then find a square that has those  $k$  points on its boundary. Then count the number of other points that are within 3 pixels of the boundary of the square. Keep doing this, and pick the square that matches the most points.

So the real question is, what is  $k$ , and how do we find the square that matches  $k$  points.  $k=3$  in this case, since a square has three degrees of freedom (position and size). As an example, suppose the three points are: (3,2), (1,4) and (6,6). We can see that (3,2) can lie on the top of the square, and (1,4) on the left side of the square. The corner of the square will be at (1,2). Then, imagine expanding the square from this corner until it hits (6,6). This will happen when the square has a width of 5.

4. SIFT: Consider the 5x5 image below. Find the image gradient of the 3x3 inner square (the points in bold face). Construct a histogram of the direction of the image gradients.

1	1	1	1	0
2	<b>2</b>	<b>2</b>	<b>1</b>	0
3	<b>3</b>	<b>2</b>	<b>1</b>	0
4	<b>3</b>	<b>2</b>	<b>1</b>	0
4	3	2	1	0

The image gradients, starting with the first boldfaced 2 will be:

(0,1), (-1/2,1/2), (-1,0)  
 (-1/2, 1/2), (-1,0), (-1,0)  
 (-1,0) (-1,0), (-1,0)

If we discretize angles every 45 degrees, so we have 8 bins, we would get six entries at 180 degrees (corresponding to (-1,0)), 1 entry at 90 degrees, and 2 entries at 315 degrees (-1/2, 1/2). SIFT builds a histogram much like this to describe a region, but with a bit more sophistication. For example, in SIFT the entries in the histogram would be weighted by the gradient magnitude.

5. 2D Transformations:

- a. Give a matrix that rotates objects by 45 degrees counterclockwise and scales them by a factor of 2.

$$\begin{pmatrix} \frac{2 \cos \pi}{4} & -\frac{2 \sin \pi}{4} & 0 \\ \frac{2 \sin \pi}{4} & \frac{2 \cos \pi}{4} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

- b. Suppose S is a matrix that performs a similarity transformation. Applying S to the point (1,1) takes it to the point (4,2). Applying S to the point (2,2) takes it to the point (5,5). What is S?

We have

$$\begin{pmatrix} a & -b & t_x \\ b & a & t_y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \\ 1 \end{pmatrix}$$

This tells us  $a-b+t_x = 4$  and  $a+b+t_y=2$ . So  $2a + t_x + t_y = 6$ , and  $2b + t_y - t_x = -2$ . We also have:

$$\begin{pmatrix} a & -b & t_x \\ b & a & t_y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 5 \\ 5 \\ 1 \end{pmatrix}$$

So we get  $4a + t_x + t_y = 10$  and  $4b + t_y - t_x = 0$ .

So we get  $2a = 4$  and  $2b = 2$ .  $a = 2$ ,  $b = 1$ . We then get  $t_x = 3$ ,  $t_y = -1$ .

- c. Construct an affine transformation that will stretch objects by a factor of 2 in the diagonal direction. That is, it will take the point (1,1) to (2,2), for example.

We can do this by combining matrices that rotate clockwise 45 degrees, stretch in the x direction by a factor of 2, and then rotate back.

$$\begin{pmatrix} \frac{\cos \pi}{4} & -\frac{\sin \pi}{4} & 0 \\ \frac{\sin \pi}{4} & \frac{\cos \pi}{4} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{\cos \pi}{4} & \frac{\sin \pi}{4} & 0 \\ -\frac{\sin \pi}{4} & \frac{\cos \pi}{4} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

6. Perspective Projection

- a. Suppose we have a camera with a focal point at  $(0,0,0)$  and an image plane at  $z=1$ . If there is a point in the world at  $(16,8,4)$ , where will it appear in the image?

$(4,2,1)$

- b. Suppose we have a camera with a focal point at  $(0,0,0)$  and an image plane at  $z=2$ . If there is a point in the world at  $(16,8,4)$ , where will it appear in the image?

$(8,4,2)$

- c. Suppose we have a camera with a focal point at  $(0,0,0)$  and an image plane at  $x=1$ . If there is a point in the world at  $(16,8,4)$ , where will it appear in the image?

$(1, \frac{1}{2}, \frac{1}{4})$

- d. Suppose we have a camera with a focal point at  $(3,7,2)$  and an image plane at  $x+z = 6$ . If there is a point in the world at  $(16,8,4)$ , where will it appear in the image?

We can describe a line from the focal point to the world point as:

$$(3,7,2) + t(13,1,2) = (x,y,z)$$

Note that at the focal point,  $x+z = 5$ , and that for  $t>0$ ,  $x+z$  will increase. In fact, if  $t=1$ ,  $x+z$  increases by 15, compared to  $t=0$ . Since we only need to increase  $x+z$  by 1 to reach the image plane, we set  $t = 1/15$ . Then we get the point  $(3 \frac{13}{15}, 7 \frac{1}{15}, 2 \frac{2}{15})$ .

7. Vanishing points: Suppose we have a camera with a focal point at  $(0,0,0)$  and an image plane of  $z = 1$ .

- a. There is a line whose location is described by the equations,  $y = -4$ ,  $x+2z = 3$ . What is its vanishing point?

As  $z$  gets really big,  $y$  stays at  $-4$ , so the  $y$  coordinate of the image point will be  $-4/z$ , which tends towards 0.

We have  $x = 3-2z$ . So the  $x$  coordinate of the image point is  $(3-2z)/z$ . As  $z$  gets really big, this tends towards  $-2$ .

- b. There is a line whose location is described by the equation  $y + z = -1$ ,  $x + 2z = 3$ . What is its vanishing point?

Here, we have  $x = 3 - 2z$ , so the  $x$  coordinate in the image is  $(3 - 2z)/z$ , which goes to  $-2$  as  $z \rightarrow \infty$ .

We have  $y = -1 - z$ , so the  $y$  coordinate in the image is  $(-1 - z)/z$ , which goes to  $-1$  as  $z \rightarrow \infty$ .

8. Suppose we take two pictures. In the first, the camera has a focal point at  $(0,0,0)$  with an image plane of  $z=1$ . In the second, the focal point is at  $(0,1,1)$  and the image plane is  $z = 2$ . Give an example of conjugate epipolar lines in the two images. That is, give a pair of lines, one in each image, so that every point in line 1 matches some point in line 2.

First, let's find a plane that goes through the two focal points. The simplest plane is  $x = 0$ . Next, let's intersect this with the two image planes. To do this, we can just say that one line is the line  $x=0, z = 1$ , and the second line is  $x=0, z = 2$ .