CS223b Midterm Exam, Computer Vision

Monday February 25th, Winter 2008, Prof. Jana Kosecka

Your name _____

email _____

- This exam is 8 pages long including cover page. Make sure your exam is not missing any pages. The exam has maximum score of 100 points. You have 75 minutes to take the exam.
- The exam is open book, open notes, but no electronic devices that can communicate with the outside world.
- Write your answers in the space provided. If you need extra space, use the back of the preceding sheet.
- Write clearly and been concise.
- SCPD students: If you are taking this exam off campus, you have to fax it (650)-725-1449 exactly 75 minutes after receipt. Alternatively you can e-mail the results to kosecka@ai.stanford.edu.

Questions	Points
1 (15)	
2 (15)	
3 (15)	
4 (20)	
5 (10)	
6 (25)	
total	

1. (15) An important parameter of the imaging system is the *field of view* (FOV). Field of view is twice the angle between the optical axis (z-axis) and the end of the retinal plane (CCD array). Imagine that you have a camera system with focal length 16mm, and retinal plane (CCD array) is $(16mm \times 12mm)$ and that imaging surface is sampled 640×480 pixels in each dimension.

   a) Compute the FOV (horizontal and vertical)

   b) Write down the relationship between the image coordinate and a point in 3D world expressed in the camera coordinate system.

   d) Describe how is the size of FOV related to the focal length and how it affects the resolution in the image.

   e) Given the horizontal FOV you computed, how many images do you need to create 360 degree panorama, assuming that you will need 50% overlap between neighboring views.

2. (15) Show the resulting image obtained after convolution of the original image with the following approximation of the derivative filter $[-1, 0, 1]$ in the horizontal direction.

0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	1^*	1	1	1	0	0	0
0	0	0	1	1	1	1	0	0	0
0	0	0	1^*	1	1	1^*	0	0	0
0	0	0	1	1	1	1	0	0	0
0	0	0	1	1	1	1	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0

(a) Compute gradient magnitude at pixels (4,4), (6,4) and (6,7) (marked with * in the image).

(b) Compute gradient direction at those same points.

(c) Describe in words what does the non-maximum suppression step on the gradient magnitude in edge detection process accomplish ?

3. (15) **Motion recovery.** Consider a set of corresponding points $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ in retinal coordinates in two views, which are related by pure translation T.

   (a) Write down a simplified version of the epipolar constraint in case the motion is pure translation.

   (b) Describe a linear least squares algorithm for estimation of translation T. What is the minimal number of corresponding points needed to solve for T ?

   (c) i) Suppose now that the camera is not calibrated and you can measure only pixel coordinates points $\boldsymbol{x}'_1$ and $\boldsymbol{x}'_2$. Can you still recover the translation between the two views ?

   ii) Suppose that you know all intrinsic camera parameters expect the focal length. Can you now recover the translation between the two views ?

4. (20) A camera in a low flying plane captures an image of a 100x100m planar wheat field. The four corners of the field have the following image coordinates (in pixels) (-30,-50), (120,-30), (100,50), (-30,60), where the origin of the image coordinate system is assumed to be in the center of the image. You can assume that field is perfectly planar and that you know the focal length of the camera. How would you compute the relative orientation of the camera and the field at the instance the image was captured ? What can you say about the distance between the camera and the plane ? Write down the geometry of the problem and describe the individual steps of the algorithm (you don't have to compute the actual values).

5. (10) Suppose you observe 3 vanishing points in the image and assume that you know all the intrinsic parameters except the focal length. How would you determine the rotation of the camera with respect to the origin of the world coordinate frame ? Write down the geometry of the problem and describe the individual steps of the algorithm.

6. (25) Provide short answers to the following questions ?

   - (3) What is the aperture problem ?

   - (3) What does it mean for the 2D convolution kernel (filter) to be separable ?

   - (3) When is a purely translational model suitable for image stitching ?

   - (3) Given two lines in the image denoted by their projective coordinates l_1 and l_2, how would you compute an intersection point of these two lines ?

   - (3) The rank of the multiview measurement (point correspondences) matrix M of dimensions (2m x n) under orthographic projection is 3. Explain why ?

   - (3) How is the rotational invariance enforced in SIFT features ?

   - (3) What is an epipole ?

- (1) Does cylindrical projection preserve straight lines YES/NO

- (1) Image thresholding is not a linear operation YES/ NO

- (1) Fundamental matrix maps points to points in respective views YES/NO

- (1) Essential matrix is not invertible YES/NO