Midterm Exam CS223B

Stanford CS223B Introduction to Computer Vison, Winter 2006

Full Name: _____

Email: _____

Welcome to the CS223B Midterm Exam!

- The exam is 6 pages long. Make sure your exam is not missing any sheets. The exam has a maximum score of 120 points. You have 60 minutes.
- The exam is closed book, closed notes, closed cell phones, etc.
- Write your answers in the space provided. If you need extra space, use the back of the preceding sheet.
- Write clearly and be concise. If your answer is more than a few sentences, you're missing something.
- SCPD students: If you are taking this exam off campus, you must fax your completed exam to (650) 725-1449 within 70 minutes of receipt. Alternatively, you can email your answers to cs223b+midterm@gmail.com.

Question	Points
1 (15 max)	
2 (30 max)	
3 (15 max)	
4 (20 max)	
5 (40 max)	
total	

1 SIFT Features 15pts

1. We learned that SIFT features have a number of invariances, but they are not invariant to arbitrary changes in the scene. Name three different invariances that are not addressed by SIFT features. Put differently, what changes to a scene would make a SIFT feature detector fail?

2. Suppose you are given range images in addition to camera images. Could you imagine a modification of the basic SIFT feature that would introduce a new invariance, which SIFT features currently don't offer? What would that modification be, and what invariant would it provide?

2 Stereo 30pts

1. Most stereo rigs have 2 cameras. Name two distinct advantages and two distinct disadvantages of adding a third camera to a stereo rig. How would you arrange the three cameras?

2. Consider a single scan line in a rectified stereo rig. Suppose the left scan line observes the following features

   a a c c a a c

   and the right scan line observes

   a c c a a b c

   for the different features 'a', 'b', and 'c'. Assume that the left and right side of the scan lines align perfectly. Using dynamic programming, compute the most likely correspondence, assuming a constant cost of 2 for a feature mismatch, and a cost of 1 for occlusion and dis-occlusion. The following table might be of help.

   (Set up the table and fill in the exact numerical values, as discussed in class. Then read off the optimal matching.)

	a	c	c	a	a	b	c
a							
a							
c							
c							
a							
a							
c							

Specify correspondences here:

left:	
right:	

3 Fundamental Matrix 15pts

1. If F is the fundamental matrix of the camera-pair (P, P'), then what is the fundamental matrix for (P', P)?

2. If for a point x in the first image, the corresponding epipolar line is Fx, then what is the epipolar line corresponding to x' in the second image?

3. What are the left and right null-spaces of F?

4 Structure From Motion 20pts

1. Name two reasons for and two reasons against using the affine method for SFM by Tomasi/Kanade.

2. Suppose we would like to determine the size of a cube from a set of K calibrated cameras whose extrinsics are unknown (but whose intrinsics are known). Suppose each of the cameras can see the same M corners of the cube, and suppose there is no correspondence problem. How many cameras K and how many corners M do we need to determine the size of the cube? (Notice: a cube has only 8 corners, hence $M \leq 8$). If mulitple solutions exist, give them all. If no solution exists, explain why.

5 True or False? 40pts

Correct answers are worth 2 points per question; a false answer results in minus 2 points.

TRUE FALSE The stronger the curvature of a thin lens, the shorter its focal length.

TRUE FALSE Template matching via convolution is invariant to rotation.

TRUE FALSE There exist linear filters for which the result is rotationally invariant.

TRUE FALSE The Harris corner detector is a linear filter.

TRUE FALSE A camera can be calibrated from a single image of a planar calibration pad (with 64 point features) if we are given the camera extrinsics.

TRUE FALSE Knowing the extrinsics of a two-camera system is sufficient to estimate the depth of a point feature observed by both cameras.

TRUE FALSE For stereo vision we need to know the intrinsics of the cameras.

TRUE FALSE Epipole is synonym for vanishing point in a stereo rig.

TRUE FALSE The inverse of the essential matrix is symmetric.

TRUE FALSE The computational complexity of a Hough transform increases exponentially with the number of unknown parameters in the geometric shape that is being detected.

TRUE FALSE It is possible that an image does not even have a single vanishing point.

TRUE FALSE The aperture effect in optical flow exists only for 2D images. For 3D range images (where we know the depth of each pixel), we do not have an aperture effect.

TRUE FALSE Image panoramas are more easily constructed from a rotating than a translating camera.

TRUE FALSE SFM problems can be solved using stereo vision techniques when the camera intrinsics and extrinsics are known.

TRUE FALSE The fundamental matrix can be computed from the camera intrinsics and extrinsics.

TRUE FALSE A pair of points in physical space and their corresponding image points always lie in a plane.

TRUE FALSE After you change the zoom on your computer vision rig, it's okay to continue using the camera's intrinsic parameters from an earlier calibration.

TRUE FALSE The epipolar constraint in stereo vision is useful because it reduces a 3-dimensional search for corresponding points to a 2-dimensional search.

TRUE FALSE The bigger the baseline in stereo, the bigger the parallax.

TRUE FALSE The cornerstone assumption in dense optical flow calculations is that the intensity of light reflected from objects is constant from frame to frame.