Midterm Examination

CS 766: Computer Vision

November 9, 2006

LAST NAME: ____

FIRST NAME: ______________________________________

Problem	**Score**	**Max Score**
1	________	10
2	________	14
3	________	12
4	________	13
5	________	16
6	________	15
Total	________	80

1. [10] **Camera Projection**

(a) [2] True or False: The orthographic projection of two parallel lines in the world must be parallel in the image.

(b) [3] Under what conditions will a line viewed with a pinhole camera have its vanishing point at infinity?

(c) [5] 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? (Note: the aspect ratio is defined as the ratio between the width and the height of a pixel; i.e., k_u/k_v.)

2. [14] **Camera Calibration**

(a) [5] Show how the projection of a point in a planar scene at world coordinates (X, Y) to pixel coordinates (u, v) in an image plane can be represented using a *planar affine camera model*.

(b) [3] Under what conditions is the use of an affine transformation appropriate when viewing a planar scene?

(c) [3] How many degrees of freedom are there to solve for in (a), and what is the minimum number of calibration points needed to estimate the calibration parameters?

(d) [3] What effects can a planar affine transformation have on parallel lines?

3. [12] **Edge Detection**
Compare the Canny edge detector and the Laplacian-of-Gaussian (LoG) edge detector for each of the following questions.

(a) [3] Which of these operators is/are isotropic and which is/are non-isotropic?

(b) [3] Describe each operator in terms of the order of the derivatives that it computes.

(c) [3] What parameters must be defined by the user for each operator?

(e) [3] Which detector is more likely to produce long, thin contours? Briefly explain.

4. [13] **Feature Detection and Description**

(a) [5] We want a method for *corner detection* for use with 3D images, i.e., there is an intensity value for each (x,y,z) voxel. Describe a generalization of either the Harris corner detector or the Tomasi-Kanade corner detector by giving the main steps of an algorithm, including a test to decide when a voxel is a corner point.

(b) [8] The *SIFT descriptor* is a popular method for describing selected feature points based on local neighborhood properties so that they can be matched reliably across images. Assuming feature points have been previously detected using the SIFT feature detector, (i) briefly describe the main steps of creating the SIFT feature *descriptor* at a given feature point, and (ii) name three (3) scene or image changes that the SIFT descriptor is invariant to (i.e., relatively insensitive to).

5. [16] **Hough Transform and RANSAC**
After running your favorite stereo algorithm assume you have produced a dense depth map such that for each pixel in the input image you have its associated scene point's (X, Y, Z) coordinates in the camera coordinate frame. Assume the image is of a scene that contains a single dominant plane (e.g., the front wall of a building) at unknown orientation, plus smaller numbers of other scene points (e.g., from trees, poles and a street) that are not part of this plane. As you know, the plane equation is given by $ax + by + cz + d = 0$.

(a) [8] Define a *Hough transform* based algorithm for detecting the orientation of the plane in the scene. That is, define the dimensions of your Hough space, a procedure for mapping the scene points (i.e., the (X, Y, Z) coordinates for each pixel) into this space, and how the plane's orientation is determined.

(b) [8] Describe how the *RANSAC algorithm* could be used to detect the orientation of the plane in the scene from the scene points.

6. [15] **Epipolar Geometry and Stereo**
(a) [5] Consider the following top view of a stereo rig:

Draw the front views of the two 2D images and show (and clearly label) the approximate positions of the epipoles and epipolar lines for this configuration of cameras.

(b) [3] Given a conjugate pair of points, at pixel coordinates $p = (x_l, y_l)$ in the left image and at pixel coordinates $q = (x_r, y_r)$ in the right image of a stereo pair, give the equation that describes the relationship between these points when neither the intrinsic nor extrinsic parameters of the cameras are known. Also specify how many conjugate pairs of points are needed to solve for all of the unknowns in your equation.

(c) [3] What is the *disparity gradient constraint* that is often used in solving the stereo correspondence problem.

(d) [4] Yes or No: Can scene depth be recovered from a stereo pair of images taken under each of the following circumstances:

(i) [2] Two images produced by weak perspective projection.

(ii) [2] Two images taken by a single perspective projection camera that has been rotated about its optical center.