Examination #1

CS 766: Computer Vision

October 20, 2005

Last (Family) Name: ______________________________

First Name: ______________________________

Problem	**Score**	**Out of**
1	_______	16
2	_______	20
3	_______	18
4	_______	10
5	_______	16
Total	_______	80

1. [16] **Image Projection**

(a) [4] Given the 3 · 4 camera matrix

$$P = \begin{bmatrix} 5 & -14 & 2 & 17 \\ -10 & -5 & -10 & 50 \\ 10 & 2 & -11 & 19 \end{bmatrix}$$

and a 3D point in homogeneous coordinates $\mathbf{X} = [0\ 2\ 2\ 1]^T$

(i) What are the Cartesian coordinates of the point **X** in 3D?

(ii) What are the Cartesian image coordinates, (u, v), of the projection of **X**?

(b) [11] An ideal pinhole camera has focal length 5mm. Each pixel is 0.02 mm · 0.02 mm and the image principal point is at pixel (500, 500). Pixel coordinates start at (0, 0) in the upper-left corner of the image.

(i) [6] What is the 3 · 3 camera calibration matrix, **K**, for this camera configuration?

(ii) [2] Assuming the world coordinate frame is aligned with the camera coordinate frame (i.e., their origins are the same and their axes are aligned), and the origins are at the camera's pinhole, what is the 3 x 4 matrix that represents the extrinsic, rigid-body transformation between the camera coordinate system and the world coordinate system?

(iii) [4] Combining your results from (i) and (ii), compute the projection of scene point (100, 150, 800) into image coordinates.

2. [20] **Camera Calibration**
A camera is rigidly mounted so that it views a planar table top. A projector is also rigidly mounted above the table and projects a narrow beam of light onto the table, which is visible as a point in the image of the table top. The height of the table top is precisely controllable but otherwise the positions of the camera, projector, and table are unknown. For each of the following table top heights, the point of light on the table is detected at the following image pixel coordinates:

Table height	Image coordinates of beam of light
50 mm	(100, 250)
100 mm	(140, 340)

(a) [4] Using a *projective camera model* specialized for this particular scenario, write a general formula that describes the relationship between world coordinates (x), specifying the height of the table top, and image coordinates (u, v), specifying the pixel coordinates where the point of light is detected. Give your answer using homogeneous coordinates and a projection matrix containing variables.

(b) [2] For the *first* table top position given above and using your answer in (a), write out the explicit equations that are generated by this one observation.

(c) [1] How many degrees of freedom does this transformation have?

(d) [1] How many table top positions and associated images are required to solve for all of the unknown parameters in the projective camera model?

(e) [3] Once the camera is calibrated, given a new unknown height of the table and an associated image, can the height of the table be uniquely solved for? If so, give the equation(s) that is/are used. If not, describe briefly why not.

(f) [3] If in each image we only measured the u pixel coordinate of the point of light, could the camera still be calibrated? If so, how many table top positions are required? If not, describe briefly why not.

(g) [3] If instead of assuming a projective camera for (a), we instead assume an *affine camera model* for this problem, write a general formula that describes the relationship between (x) and (u, v).

(h) [3] When would an affine camera model be appropriate instead of using a projective camera model? Give one advantage and one disadvantage of using an affine camera instead of a projective camera for this problem.

3. [18] **Edge Detection**
For each of the following properties of edge detectors, indicate whether the **Canny** edge detector or the **Marr-Hildreth** (also known as ^{2}G or Laplacian-of-Gaussian) detector is better with respect to this property, and explain briefly why.

(a) [2] Fewer number of parameters that must be set.

(b) [2] Closed chains of edges detected.

(c) [2] Computes the edge orientation.

(d) [2] Better localization of the true edge position.

(e) [2] Can be implemented more efficiently (i.e., faster execution time).

(f) [2] Requires non-maximum suppression to thin edges.

(g) [2] Is an isotropic operator.

(h) [2] Less likely to round corners where the boundary curvature is high.

(i) [2] Less sensitive to noise.

4. [10] **Distance Transform**

Consider an 8 · 8 image with two points at pixel coordinates $p = (3, 3)$ and $q = (7, 5)$, specifying the row and column coordinates, respectively. We want to compute the set of points that are *equidistant* from p and q. If we use the Euclidean (L_2) metric, the set of points is a line defining the perpendicular bisector between p and q. For each of the following two alternative metrics, compute the **distance transform** for the image and circle those pixels that are equidistant from p and q as defined by that metric. Each square should show one number, which is the distance to the closest of the two given points.

(a) City-block (L_1) metric.

		p					
				q			

(b) Chessboard (L_∞) metric.

		p					
				q			

5. [16] **Image Segmentation**

(a) [2] The Normalized-Cut segmentation algorithm computes the eigenvalues and eigenvectors of the normalized affinity matrix, $D^{-1/2}AD^{-1/2}$. Describe the meaning of the two eigenvectors corresponding to the two smallest eigenvalues.

(b) [2] Define an affinity measure, $A(i,j)$, which measures the similarity of pixels i and j and depends on the similarity of their intensities, $f(i)$ and $f(j)$, and inversely on their distance apart, $d(i,j)$. Briefly explain the rationale for your definition.

(c) [4] To minimize the computational burden of NCut, which requires computing a very large affinity matrix, suppose we first divide an input image into small sub-images, say each of size 50 · 50. Using these sub-images, describe the main steps of a procedure to use NCut in a two-stage fashion (first with the sub-images, and second with the results of the first stage) to segment an image.

(d) [8] Consider an image that is empty except for two sets of points. One is a set of points distributed roughly uniformly on a circle of radius *r* centered at point C1, which is near the center of the image. The other set of points is distributed on a circle of radius 2*r*, which is centered at point C2, which is located inside the other circle. Assume the points in each set are distributed densely enough so that the distances between points on the same circle are smaller than the distances between points on different circles.

(i) Describe what segmentation the **k-means clustering** algorithm would produce for this example, and briefly explain why.

(ii) Describe what segmentation the **normalized-cut** algorithm would produce for this example, and briefly explain why.