Exam / Computer Vision Part at UJM

***Human Vision and Computer Vision* course**

Master CIMET - June 2013

(No document is authorized – calculates are authorized - duration: 2 hours)

(20 questions accounting for 21.5 points, the exam is scored for 20 points)

Camera Projection

Question 1: (*1 point*) True or False: The orthographic projection of two parallel lines in the world must be parallel in the image.

Question 2: (*1 point*) Under what conditions will a line viewed with a pinhole camera have its vanishing point at infinity?

Question 3: (*2 points*) 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., $u_0 \neq 0$ and $v_0 \neq 0$). 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.)

Camera Calibration

Question 1: (*2 points*) 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*.

Question 2: (*1 point*) Under what conditions is the use of an affine transformation appropriate when viewing a planar scene?

Question 3: (*1 point*) How many degrees of freedom are there to solve for in question 1, and what is the minimum

number of calibration points needed to estimate the calibration parameters?

Question 4: (*1 point*) What effects can a planar affine transformation have on parallel lines?

Image Projection

(a) Given the 3x4 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 $\mathbf{X} = [0\ 2\ 2\ 1]^T$

Question 1: (*0.5 point*) What are the Cartesian coordinates of the point **X** in 3D?

Question 2: (0.5 *point*) What are the Cartesian image coordinates, (u, v), of the projection of **X**?

(b) An ideal pinhole camera has focal length 5mm. Each pixel is 0.02 mmx0.02 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.

Question 3: (2 *points*) What is the 3x3 camera calibration matrix, **K**, for this camera configuration?

Question 4: (1 *point*) 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 3x4 matrix that represents the extrinsic, rigid body transformation between the camera coordinate system and the world coordinate system?

Question 5: (2 *points*) Combining your results from (a) and (b), compute the projection of scene point (100, 150, 800) into image coordinates.

Feature Detection and Description

Question 1: (*2 points*) 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 the Harris corner detector by giving the main steps of an algorithm, including a test to decide when a voxel is a corner point.

Stereo-Vision

Question 1: (*0.5 point*) Define what is an epipolar line in stereo-vision?

Question 2: (*0.5 point*) For which issues are the epipolar lines useful?

Question 3: (*1 point*) How do we estimate the parameters of the epipolar lines?

True or False?

Correct answer is 0.5 point per question; a false answer results in minus 0.5 point.

Answer:

TRUE FALSE An image may have no more than three vanishing points

TRUE FALSE The smaller the baseline of a stereo system, the lower the error rate for the correspondence step.

TRUE FALSE The epipolar plane is normal to the optical axis.

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

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