CS 6550 Computer Vision
Midterm Exam (Date: 11/30/2015)
There are totally 6 problems and totally 120 points.

1. (20pts) (a) What is the field of view of a camera system? (b) How is it related to the focal length and image sensor size? (c) Give the linear matrix equation for perspective projection by using homogeneous coordinate. Explain the notations used in the equation. Give the detailed projection matrix for perspective projection. (d) Give examples of the BRDF functions (Bidirectional Reflectance Distribution Function) for diffuse reflection and specular reflection. (e) What is the main difference between the diffuse reflection and specular reflection?

2. (20pts) Consider an N-by-N grayscale image I(x,y). (a) How do you apply the Sobel operators to compute its image gradient? (b) What is the gradient covariance matrix **H** (used in Harris corner detection) at pixel (u,v)? Give its mathematical definition. (c) For a pixel located on a straight line, what are the characteristics of the two eigenvalues for the corresponding matrix **H**? How about the corresponding eigenvectors? (d) Give an example of response function for corner detection based on **H**. Describe how you can detect the corners by using the response function and the non-maximum suppression strategy? What is the advantage of using the non-maximal suppression here? (e) Is the corner detector given in your answer in (d) scale invariant? If yes, why? If no, how do you make it scale invariant?

3. (20 pts) Given a set of 3D points in world coordinate $\{(X_i, Y_i, Z_i) \mid i=1, \ldots, n\}$ and their corresponding 2D image coordinates represented as (u_i, v_i), i= 1, …, n.
   (a) Write down the 3D-2D camera projection relationship with a camera projection matrix P by using the homogeneous coordinates.
   (b) Derive the linear equations used in the linear least-squares algorithm for computing the camera projection matrix P from these 3D-2D correspondence points.
   (c) Describe how to estimate the camera calibration matrix and the camera rotation and translation from the camera projection matrix P.
   (d) Given a camera projection matrix P and a 3D straight line represented by $(U_0 + c * U, V_0 + c * V, W_0 + c * W)$, where c is a real number, how do you compute the projected coordinate when the line goes to infinity?
   (e) What is a vanishing point? How do you compute a vanishing point from an image?

4. (20 pts) Consider the epipolar geometry between two views. Assume a set of 2D-2D image correspondence points between two images captured by two cameras is given. This set of n point correspondences is denoted by $\{(u_i, v_i), (u_i', v_i') \mid i= 1, \ldots, n\}$, where (u_i, v_i) is on the left image and (u_i', v_i') is the corresponding point on the right image.

(a) Describe the epipolar geometry between two cameras by using the Fundamental matrix **F**. How is the fundamental matrix F related to the two camera calibration matrices and their relative rotation matrix **R** and translation vector **t**?

(b) To estimate the fundamental matrix **F** from a set of 2D-2D image correspondence points between two images by using the eigenvector method, give the linear system and the associated least-squares error function for the estimation of the fundamental matrix.

(c) What is the minimal number of pairs of correspondence points required to obtain the least-squares solution? Explain why? Assume all the linear constraints associated with the corresponding points are independent.

(d) After the fundamental matrix F is estimated from the point correspondences, describe how do compute the epipolar lines for both images.

(e) How do you determine the epipoles for each of the two images from their fundamental matrix F?

5. (20 pts) Consider a stereo camera system. (a) Given the camera projection matrices for these two cameras P_1 and P_2, assume a pair of 2D correspondence points **x** and **x'** are detected from the two images. Derive the least-square solution to estimate its 3D location. (b) What is the purpose of the image rectification for stereo vision? What are the advantages of image rectification? (c) Give the basic idea and procedure to achieve image rectification for stereo vision. (d) What is the relationship between the depth, disparity and focal length for an ideal stereo system? (e) What is the straight line, ax+by+c=0 in an image I(x,y), represented in a homogeneous coordinate? If the image is transformed by a homography matrix H, what is the corresponding straight line representation in the transformed image?

6. (20 pts) Consider to partition a set of data points $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$, where each $\mathbf{x}_i$ is a 3-dimensional RGB color vector, into K clusters by using k-means algorithm. (a) Assume we are given a set of initial K mean vectors, $c_1, \ldots, c_K$. Give the main two-step procedure in the k-means clustering. (b) What is the objective function to be minimized in the k-means clustering? Will the k-means algorithm lead to the global minimum of this objective function? If no, what can you do to improve the clustering results by using k-means algorithm? (c) What is the main difference between the final results for k-means clustering and the probabilistic clustering? (d) What are the advantages of using mean-shift algorithm over k-means clustering? (e) What is a shape vector determined from a set of landmark points $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$ in a 2D image? Given a set of shape vectors $S_1, \ldots, S_N$, describe how to compute the Active Shape Model (ASM) from these shape vecors.