CS-512 – Assignment 5 (4%)

Stereo vision

Due by: November 23, 2017

Review questions

Answer the following questions. Make sure that your answers are concise. In questions requiring explanations, make sure your explanations are brief.

1. Stereo
   (a) Explain the difference between sparse and dense stereo matching. What are the advantages/disadvantage of each approach?
   (b) Explain how normalized cross correlation (NCC) and sum of square distances (SSD) can be used for point matching. What is the isk in allowing the search space to be the entire image? How can the search space be reduced to a line?
   (c) Given an axis aligned stereo pair with corresponding points $(100, 200)$ and $(103, 200)$ in the left and right images respectively, compute the depth (z-coordinate) of the 3D point that produced this projection. Assume that the focal length of both cameras is 10, that the baseline is 100, and that we are working in camera coordinates.
   (d) Explain the ambiguity problem in stereo matching.
   (e) Given R_l, T_l and R_r, T_r (rotation and translation of left/right cameras with respect to world), Write the expression for the rotation and translation of the right camera with respect to the left camera.
2. Epipolar geometry
   (a) Explain what is the epipole. How are epipolar lines "formed" in the image?
   (b) Write the expression of the essential matrix E. Given a set of corresponding points p_l, p_r (in camera coordinates), write the epipolar constraint equation using the essential matrix E.
   (c) Write the expression of the fundamental matrix F. Given a set of corresponding points p_l, p_r (in image coordinates), write the epipolar constraint equation using the fundamental matrix F.
   (d) What is the rank of the essential and fundamental matrices? Why?
   (e) Given the point p_l in the left image and the fundamental matrix F, what is the corresponding right epipolar line?
   (f) Given the point p_r in the right image and the fundamental matrix F, what is the corresponding left epipolar line?
   (g) Explain what is a weak calibration of a stereo pair.
   (h) Given corresponding points $(100, 200), (50, 100)$ in the left/right images respectively, write the first line in the matrix that has to be formed to solve for the unknown fundamental matrix using the 8-point algorithm. Do not normalize the point.

(i) Explain how to normalize the points in the 8-point algorithm for estimating the fundamental matrix. Why is this necessary? Explain how the fundamental matrix of the original points can be recovered from the fundamental matrix of the normalized points.

(j) Explain how can the epipoles be recovered from the fundamental matrix.

3. Reconstruction

   (a) Explain how a stereo pair can be rectified. What happens after the pair is rectified?

   (b) Explain the different approaches for reconstruction depending on what is known about the cameras.

   (c) Let the right image be rotated by R and translated by T with respect to the left one. Let p_l and p_r be left and right corresponding points in the two images. Write the matrix that has to be formed to solve for the coefficients (a, b, c) of the corresponding triangulated 3D point.

   (d) Using the coefficients (a, b, c) of the triangulated point, write the formula to compute it.

   (e) Explain why there is an unknown scale in Euclidean reconstruction. How can this unknown scale be removed?

   (f) Given the estimated essential matrix E, explain how this matrix can be normalized to have a baseline of 1.

   (g) Explain how the unknown signs of rotation and translation can be determined when using Euclidean reconstruction.

Programming questions

In this assignment you need to implement ONE of several stereo vision algorithms.

1. The input image should always be converted to grayscale before processing it.
2. The main parameters of each algorithm should be made available for interactive manipulation through keyboard/mouse/trackbar interaction.
3. You may NOT use the OpenCV function which directly implements the algorithm you choose. You may, however, use it for verification purposes. You may use other OpenCV functions as necessary.
4. Your program must include a help key describing its functionality.
5. You need to evaluate the performance of the algorithm you choose using test data. The results of your evaluation should be included in your report. Try to determine the strengths and weaknesses of the algorithm.
6. For submission instructions and the necessary format of the report please refer to assignment 1.

Epipolar lines estimation

1. Given a stereo pair as input to the program, load and the display the images next to each other.
2. The user is then allowed to specify a set of corresponding points in the images using the mouse.
3. Using the set of corresponding points specified by the user, estimate the fundamental matrix relating the two views and display the estimated matrix. Normalize the points before computing the fundamental matrix and make sure that the estimated fundamental matrix is a rank 2 matrix.

4. Using the estimated fundamental matrix, let the user select a point in one image, and display the corresponding epipolar line in the second image. The user should be able to select points in both the left and right images.

5. Using the estimated fundamental matrix, compute and display (numerically) the coordinates of the left and right epipoles.

6. Test stereo images are available on the cs512 website. In addition, test the algorithm on images you take yourself.

Model reconstruction

1. Given a set of two views in which the intrinsic camera parameters are KNOWN (the extrinsic camera parameters are assumed UNKNOWN), and a set of corresponding points in the two views, reconstruct 3D points on the surface of the unknown object.

2. Using a subset of the provided corresponding points, estimate the fundamental matrix relating the two views and display the estimated matrix. Make sure that the estimated fundamental matrix has a rank of 2. Using the knowledge of the intrinsic parameters compute and display the essential matrix.

3. Using the essential matrix, estimate up to an unknown sign the rotation and translation between the views and display them. Note that the true extrinsic parameters are provided so that you can verify your computations.

4. Reconstruct the 3D points using the obtained rotation and translation and choose a consistent solution.

5. The corresponding points are described in a text file. The first line in this file contains the number of matched points. Subsequent lines contain the description of pairs of corresponding points. Each pair of corresponding points is described by 4 coordinates: x_r, y_r, x_l, y_l, where (x_r, y_r) and (x_l, y_l) are corresponding points.

6. The reconstructed points should be saved in a VRML file as a 3D point cloud and rendered using an external VRML viewer (alternatively, write a simple OpenGl program to display the points which allows for rotating the points and zooming in/out). Generation surfaces between the points is not required.

7. A calibrated image pair with a dense correspondence and known intrinsic/extrinsic parameters is available on the cs512 website. In addition, test the algorithm on images you take yourself.