CSE 252B: Computer Vision II, Winter 2014 – Assignment 2

Instructor: Ben Ochoa

Due date: Tuesday, March 4, 2014, 11:59 PM

1. Given a line $\mathbf{l} = (a, b, c)^\top$, show that the point on $\mathbf{l}$ that is closest to the origin is the point $\mathbf{x} = (-ac, -bc, a^2 + b^2)^\top$ (hint: this calculation is needed in the two-view optimal triangulation method used below).

2. Automatic estimation of the fundamental matrix $\mathtt{F}$

   (a) Download input images
   `http://vision.ucsd.edu/~bochoa/cse252b-wi14/IMG_5030.JPG`
   and
   `http://vision.ucsd.edu/~bochoa/cse252b-wi14/IMG_5031.JPG`

   (b) Feature detection. For each input image, calculate an image where each pixel value is the minor eigenvalue of the gradient matrix

   $$\mathtt{N} = \begin{bmatrix} \sum\limits_w I_x^2 & \sum\limits_w I_x I_y \\ \sum\limits_w I_x I_y & \sum\limits_w I_y^2 \end{bmatrix}$$

   where w is the window about the pixel, and I_x and I_y are the gradient images in the x and y direction, respectively. Set resulting values that are below a specified threshold value to zero (hint: calculating the mean instead of the sum in $\mathtt{N}$ allows for adjusting the size of the window without changing the threshold value). Apply an operation that suppresses (sets to 0) local nonmaximum pixel values in the minor eigenvalue image. For resulting nonzero pixel values, determine the subpixel feature coordinate using the Förstner corner point operator.

   (c) Feature matching. Determine the set of putative feature correspondences by performing a brute-force search for the greatest normalized cross correlation value between features in each of the images. Only allow matches that are above a specified threshold value (hint: calculating the correlation coefficient (in the range [-1, 1]) allows for adjusting the size of the matching window without changing the threshold value). Consider constraining the search to coordinates within a proximity of the feature coordinates.

   (d) Outlier rejection. Determine the set of inlier point correspondences using the M-estimator Sample Consensus (MSAC) algorithm. For each trial, use the 7-point algorithm to estimate the fundamental matrix, resulting in 1 or 3 (real) solutions, and calculate the error and cost for each solution. Use the Sampson error as a first order approximation to the geometric error.

   (e) Linear estimation. For the set of inlier correspondences, estimate the fundamental matrix $\mathtt{F}_{\mathrm{DLT}}$ using the direct linear transformation (DLT) algorithm (with data normalization).

   (f) Nonlinear estimation. Use $\mathtt{F}_{\mathrm{DLT}}$ as an initial estimate to an iterative estimation method, specifically the *sparse* Levenberg-Marquardt algorithm, to determine the

Maximum Likelihood estimate of the fundamental matrix $\mathtt{F}$ that minimizes reprojection error in both images. Initialize the 3D scene points in the parameter vector using the two-view optimal triangulation method, as described in lecture (in the textbook, see algorithm 12.1 on page 318, but use the ray-plane intersection method for the final step instead of the homogeneous method). List the initial cost (i.e., the (weighted) reprojection error) and the cost at the end of each iteration.

(g) Figures. Produce a pair of figures for each of the above steps. For feature detection, draw a square about the detected feature where the size of the square is the size of the window in the gradient matrix $\mathtt{N}$. For feature correspondences, draw a square that is the size of the matching window and a line segment from the feature to the coordinates of the corresponding feature in the other image (see figure 11.4 on page 292 in the textbook as an example). Additionally, in another pair of images, choose three points in the first image that are not in the set of inlier correspondences, then draw a square around these points in the first image and the corresponding epipolar lines $\mathbf{l}'_i = \mathtt{F}\mathbf{x}_i$ in the second image.