CSE 252B: Computer Vision II, Winter 2016 – Assignment 2

Instructor: Ben Ochoa

Due: Tuesday, February 2, 2016, 11:59 PM

Instructions

- Review the academic integrity and collaboration policies on the course website.
- This assignment only contains programming problems that must be completed using MATLAB.
- You must prepare a report describing the problems, and your solutions and results. The report must contain enough information for a reader to understand the problems and replicate your work without also having the assignment (i.e., this document).
- Your report will be a pdf file named `CSE_252B_hw2_lastname_studentid.pdf`, where `lastname` is your last name and `studentid` is your student ID number. The report must be prepared using LaTeX.
- All of your MATLAB source code must be included in an appendix of your report. You may find the `listings` package useful for this.
- You must create a zip file named `CSE_252B_hw2_lastname_studentid.zip`, where `lastname` is your last name and `studentid` is your student ID number. This zip file will contain the pdf file and a directory named `code` that contains all of your MATLAB source code.
- Submit your completed assignment by email to `bochoa@ucsd.edu` **and** `nkinkade@eng.ucsd.edu`. The subject of the email message must be `CSE 252B Assignment 2`. Attach the zip file to the message.
- It is highly recommended that you begin working on this assignment early to ensure that you have sufficient time to correctly implement the algorithms and prepare a report.

Problems

1. **Programming: Estimation of the camera projection matrix (45 points)**

   (a) **Linear estimation (15 points)**

   Download input data from the course website. The file `hw2_points3D.txt` contains the coordinates of 50 scene points in 3D (each line of the file gives the $\tilde{X}_i$, $\tilde{Y}_i$, and $\tilde{Z}_i$ inhomogeneous coordinates of a point). The file `hw2_points2D.txt` contains the coordinates of the 50 corresponding image points in 2D (each line of the file gives the $\tilde{x}_i$ and $\tilde{y}_i$ inhomogeneous coordinates of a point). The scene points have been randomly generated and projected to image points under a camera projection matrix (i.e., $\mathbf{x}_i = \mathtt{P}\mathbf{X}_i$), then noise has been added to the image point coordinates.

Estimate the camera projection matrix $\mathtt{P}_{\text{DLT}}$ using the direct linear transformation (DLT) algorithm (with data normalization). You must express $\mathbf{x}_i = \mathtt{P}\mathbf{X}_i$ as $[\mathbf{x}_i]^{\perp}\mathtt{P}\mathbf{X}_i = \mathbf{0}$ (not $\mathbf{x}_i \times \mathtt{P}\mathbf{X}_i = \mathbf{0}$), where $[\mathbf{x}_i]^{\perp}\mathbf{x}_i = \mathbf{0}$, when forming the solution. Include the numerical values of the resulting $\mathtt{P}_{\text{DLT}}$, scaled such that $\|\mathtt{P}_{\text{DLT}}\|_{\text{Fro}} = 1$, in your report with sufficient precision such that it can be evaluated (hint: use `format shortg` in MATLAB prior to displaying your results). Following is an example of the expected precision.

$$\mathtt{P}_{\text{DLT}} = \begin{bmatrix} 0.0061695 & -0.00467 & 0.0088168 & 0.85179 \\ 0.0089068 & -0.0024049 & -0.0062232 & 0.52363 \\ 4.8924 \times 10^{-6} & 4.3809 \times 10^{-6} & 2.4664 \times 10^{-6} & 0.0012341 \end{bmatrix}$$

(b) **Nonlinear estimation (30 points)**

Use $\mathtt{P}_{\text{DLT}}$ as an initial estimate to an iterative estimation method, specifically the Levenberg-Marquardt algorithm, to determine the Maximum Likelihood estimate of the camera projection matrix that minimizes the projection error. You must parameterize the camera projection matrix as a parameterization of the homogeneous vector $\mathbf{p} = \operatorname{vec}(\mathtt{P}^{\top})$. It is highly recommended to implement a parameterization of homogeneous vector method where the homogeneous vector is of arbitrary length, as this will be used in following assignments (see section A6.9.2 (page 624) of the textbook, and the corrections and errata).

In your report, show the initial cost (i.e., the cost at iteration 0) and the cost at the end of each successive iteration. Show the numerical values for the final estimate of the camera projection matrix $\mathtt{P}_{\text{LM}}$, scaled such that $\|\mathtt{P}_{\text{LM}}\|_{\text{Fro}} = 1$, in your report with sufficient precision such that it can be evaluated.