CS 231A Computer Vision (Winter 2015)
Problem Set 2

Due Feb 9^{th} 2015 11:59pm

1 Fundamental Matrix (20 points)

In this question, you will explore some properties of fundamental matrix and derive a minimal parameterization for it.

a Show that two 3×4 camera matrices M and M' can always be reduced to the following canonical forms by an appropriate projective transformation in 3D space, which is represented by a 4×4 matrix H. Here, we assume $e_3^T(-A'A^{-1}b+b') \neq 0$, where $e_3 = (0,0,1)^T$, $M = [A, b]$ and $M' = [A', b']$.
Note: You don't have to show the explicit form of H for the proof. [**10 points**]
Hint: The camera matrix has rank 3. Block matrix multiplication may be helpful. If you construct a projective transformation matrix H_0 that reduces M to $\hat{M}$, (i.e., $\hat{M} = MH_0$) can a H_1 be constructed such that not only does it not affect the reduction of M to $\hat{M}$ (i.e., $\hat{M} = MH_0H_1$), but it also reduces M' to $\hat{M}'$? (i.e., $\hat{M}' = M'H_0H_1$)

$$\hat{M} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \text{ and } \hat{M}' = \begin{bmatrix} a_{11} & a_{12} & a_{13} & b_1 \\ a_{21} & a_{22} & a_{23} & b_2 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

b Given a 4×4 matrix H representing a projective transformation in 3D space, prove that the fundamental matrices corresponding to the two pairs of camera matrices (M, M') and $(MH, M'H)$ are the same. [**5 points**]
(Hint: Think about point correspondences)

c Using the conclusions from (a) and (b), derive the fundamental matrix F of the camera pair (M, M') using a_{11}, a_{12}, a_{13}, a_{21}, a_{22}, a_{23}, b_1, b_2. Then use the fact that F is only defined up to a scale factor to construct a seven-parameter expression for F. *(Hint: The fundamental matrix corresponding to a pair of camera matrices $M = [I \mid 0]$ and $M' = [A \mid b]$ is equal to $[b]_\times A$.)* [**5 points**]

2 Fundamental Matrix Estimation From Point Correspondences (35 points)

This programming assignment is concerned with the estimation of the fundamental matrix from point correspondences. Implement both the linear least-squares version of the eight-point algorithm and its normalized version. In both cases, enforce the rank-two constraint for the fundamental matrix via singular value decomposition.

The data for this assignment can be found here: `http://cs231a.stanford.edu/hw/hw2/data.zip` The zipped file contains two datasets. For each dataset, you will find in the corresponding directory the files

```
pt_2d_1.txt
pt_2d_2.txt
image1.jpg
image2.jpg
```

The first two of these files contain matching image points (in the following format: number n of points followed by n pairs of floating-point coordinates). The two remaining files are the images where the points were found.

You must turn in:

a. An Algorithmic break-down of the process. [**15 points**]

b. a copy of your code, [**15 points**]

c. For both methods applied to both datasets, the average distance between the points and the corresponding epipolar lines for each image, as well as a drawing similar to Figure 1 (with the selected corresponding points and their epipolar lines). [**5 points**]

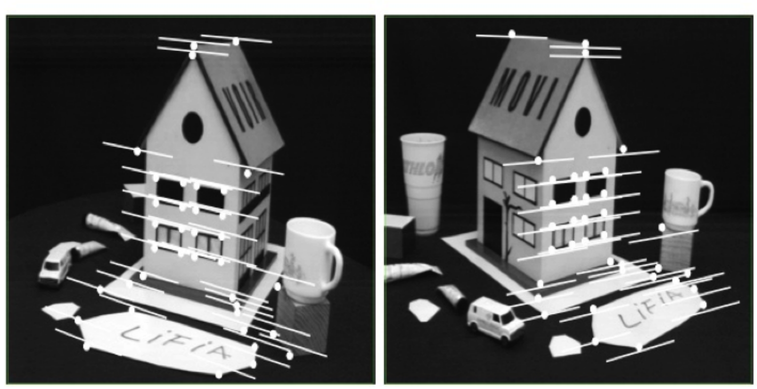

Figure 1: Example illustration, with epipolar lines shown in both images (Image courtesy Forsyth & Ponce)

3 Factorization Method (30 points)

In this question, you will explore the factorization method, initially presented by Tomasi and Kanade, for solving the affine structure from motion problem.

Begin by downloading the code templates and data from `http://cs231a.stanford.edu/hw/hw2/ps2_factor.zip` and running the function ps2q2.m. Observe the point correspondences that have been selected in the two images.

(a) Now implement the factorization method presented in the lecture slides. You only need to modify the file ps2q2.m, and you only need to enter code in the location that says "YOUR CODE HERE." [**7 points**]

(b) Run your code and visualize the result by plotting the resulting 3d points; the plot3 Matlab function may be helpful. For example:

```
plot3(x, y, z, '.');
```

where x, y, z are vectors. We also recommend calling "axis equal" after making your plots so that the axes will be scaled appropriately.

Compare your result to the ground truth provided. The results should look identical, except for a scaling and rotation. Explain why this occurs. [**5 points**]

(c) Report the 4 singular values from the SVD decomposition. Why are there 4 non-zero singular values? How many non-zero singular values would you expect to get in the idealized version of the method, and why? [**6 points**]

(d) Change the file readTextFiles.m by setting usesubset to true, and re-run the code. The file will now only load a subset of the original correspondences. Compare your new results to the ground truth, and explain why they no longer appear similar. [**6 points**]

(e) Report the new singular values, and compare them to the singular values that you found previously. Explain any major changes. [**6 points**]

4 Stereo Reconstruction (15 points)

(a) The figure above shows a rectified stereo rig with camera centers O and O', focal length f and baseline B. x and x' are the projected point locations on the virtual image planes by the point P; note that since x' is to the left of O', it is negative. Give an expression for the depth of the point P, shown in the diagram as Z. Also give an expression for the X coordinate of the point P in world coordinates, assuming an origin of O. **You can assume that the two are pinhole cameras for the rest of this question**. [**3 points**]

(b) Now assume that the camera system can't perfectly capture the projected points location on the image planes, so there is now some uncertainty about the point's location since a real digital camera's image plane is discretized. Assume that the original x and x' positions now have an uncertainty of $\pm e$, which is related to discretization of the image plane.

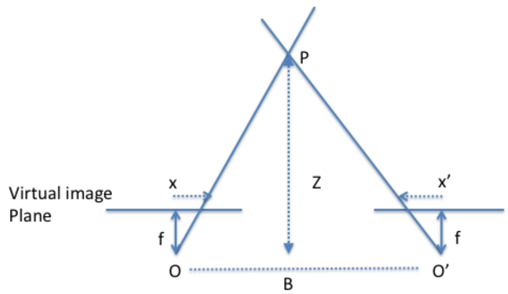

Figure 2: Rectified Stereo Rig
(B is not the mid point of OO', the figure isn't symmetric)

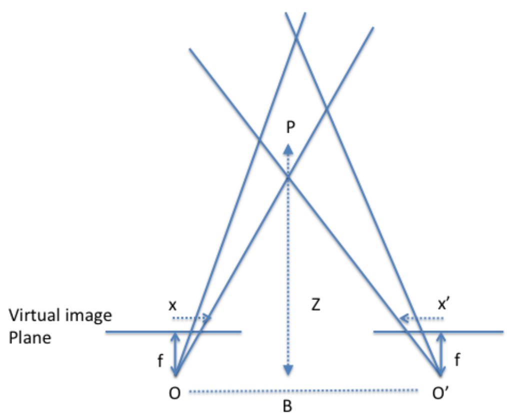

Figure 3: Rectified Stereo Rig with image plane error

(i) Give an expression of the X, Z locations of the 4 intersection points resulting from the virtual image plane uncertainty. Your answer should be in terms of f, B, x, x', and e. **[3 points]**

(ii) Give an expression for the maximum uncertainty in the X and Z directions of the point P's location in world coordinates. Your answer should still be in terms of f, B, x, x', and e. **[3 points]**

All expressions should be in terms of image coordinates only, you can assume that x is always positive and x' is always negative.

(c) Assume the X coordinate of the point P is fixed.

(i) Give an expression for the uncertainty in the reconstruction of Z, in terms of the actual value of Z and the other parameters of the stereo rig: f, B, and e (your answer should not include terms for x and x'). **[3 points]**

(ii) Find the depth when the uncertainty is at its maximum and give a physical interpretation and a drawing to explain. **[3 points]**