CSE 252A Computer Vision I Fall 2019 - Homework 4

Instructor: Ben Ochoa

Assignment published on: Tuesday, November 05, 2019

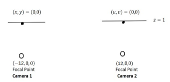
Due on: Tuesday, November 19, 2019 11:59 pm

Instructions

- Review the academic integrity and collaboration policies on the course website.
  - This assignment must be completed individually.
- All solutions must be written in this notebook.
  - This includes the theoretical problems, for which you **must** write your answers in Markdown cells (using LaTeX when appropriate).
  - Programming aspects of the assignment must be completed using Python in this notebook.
- If you want to modify the skeleton code, you may do so. It has only been provided as a framework for your solution.
- You may use Python packages (such as NumPy and SciPy) for basic linear algebra, but you may not use packages that directly solve the problem.
  - If you are unsure about using a specific package or function, then ask the instructor and/or teaching assistants for clarification.
- You must submit this notebook exported as a PDF. You must also submit this notebook as `.ipynb` file.
  - Submit both files (`.pdf` and `.ipynb`) on Gradescope.
  - **You must mark the PDF pages associated with each question in Gradescope. If you fail to do so, we may dock points.**
- It is highly recommended that you begin working on this assignment early.
- **Late policy: assignments submitted late will receive a 15% grade reduction for each 12 hours late (i.e., 30% per day). Assignments will not be accepted 72 hours after the due date. If you require an extension (for personal reasons only) to a due date, you must request one as far in advance as possible. Extensions requested close to or after the due date will only be granted for clear emergencies or clearly unforeseeable circumstances.**

Problem 1: Epipolar Geometry [4 pts]

Consider two cameras whose image planes are the z=1 plane, and whose focal points are at (-12, 0, 0) and (12, 0, 0). We''ll call a point in the first camera (x, y), and a point in the second camera (u, v). Points in each camera are relative to the camera center. So, for example if (x, y) = (0, 0), this is really the point (-12, 0, 1) in world coordinates, while if (u, v) = (0, 0) this is the point (12, 0, 1).



a) Suppose the points (x, y) = (8, 7) is matched to the point (u, v) = (2, 7). What is the 3D location of this point?

b) Compute the Essential Matrix.

c) Consider points that lie on the line x + z = 2, y = 0. Use the same stereo set up as before. Write an analytic expression giving the disparity of a point on this line after it projects onto the two images, as a function of its position in the right image. So your expression should only involve the variables u and d (for disparity). Your expression only needs to be valid for points on the line that are in front of the cameras, i.e. with z > 1.

a) d = 24, XL = 8, XR = 2
$X = \frac{24 \cdot 8}{(8-2)} = 32$
$Z = \frac{24 \cdot 1}{(8-2)} = 4$
The 3D location is $[32,7,4]$

b) The essential matrix would be $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -t_{x} \\ 0 & t_{x} & 0\end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1\end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -t_{x} \\ 0 & t_{x} & 0\end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -24 \\ 0 & 24 & 0\end{bmatrix}$

c) By observing this graphically - the disparity will be equal to $d = 2 - z - u$

In [163]:

```
import numpy as np

#### Calculating Essential Matrix ####
u1 = np.array([[8], [7], [1]])
u2 = np.array([[2], [7], [1]])



E = np.array([[0,0,0],[0,0,-24],[0,24,0]])

print("Essential Matrix\n")
print(E)

print("\nCheck p.T * E * p' = 0")
print(np.matmul(np.matmul(u1.T,E),u2))
```

```
Essential Matrix

[[  0   0   0]
 [  0   0 -24]
 [  0  24   0]]

Check p.T * E * p' = 0
[[0]]
```

Problem 2: Epipolar Rectification [4 pts]

In stereo vision, image rectification is a common preprocessing step to simplify the problem of finding matching points between images. The goal is to warp image views such that the epipolar lines are horizontal scan lines of the input images. Suppose that we have captured two images I_A and I_B from identical calibrated cameras separated by a rigid transformation

$_{A}^{B}\textrm{T}= \begin{bmatrix} \boldsymbol{R} & \boldsymbol{t} \\ 0^T & 1 \end{bmatrix}$ and $\boldsymbol{t}=[t_x,t_y,t_z]$

Without loss of generality assume that camera A's optical center is positioned at the origin and that its optical axis is in the direction of the z-axis.

From the lecture, a rectifying transform for each image should map the epipole to a point infinitely far away in the horizontal direction $ H_{A}e_{A} = H_{B}e_{B} = [1, 0, 0]^T$. Consider the following special cases:

a) Pure horizontal translation $\boldsymbol{t} = [t_{x}, 0, 0]^T$, $\boldsymbol{R} = \boldsymbol{I}$

The essential matrix would be $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -t_{x} \\ 0 & t_{x} & 0\end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1\end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -t_{x} \\ 0 & t_{x} & 0\end{bmatrix}$

$[1, 0, 0]$ is in the right null space of the essential matrix, which means that epipole 2 can be mapped to infinity. $[1, 0, 0]^T$ is in the left null space of the essential matrix so epipole 1 can be mapped to infinity.

Using the equation: $ l_{B} = E^{T} \cdot p_{A}$ The epipolar lines in image B have the equation: $[x_{B}, y_{B}, 1] \cdot [0, t_{x}, -t_{x} \cdot y_{A}]^T = 0$

b) Pure translation orthogonal to the optical axis $\boldsymbol{t} = [t_{x}, t_{y}, 0]^T$, $\boldsymbol{R} = \boldsymbol{I}$