Assignment 2: Stereo and 3D Reconstruction from Disparity

Theoretical Problems

Chuan Long

Problem 1: Epipolar Geometry with 3 Cameras (2 pts)

We discussed epipolar geometry between two cameras, which limits a search for correspondence to an epipolar line. The Java demo (http://www.ai.sri.com/luong/research/Meta3DViewer/EpipolarGeo.html) also illustrates the epipolar geometry for 3 camera views. Given the fact that for two cameras, the constraint for finding a corresponding point in the right camera for each point in the left camera is a line (i.e. still one degree of freedom), discuiss the situation with 3 cameras. Could it be that with 3 cameras, point to point correspondences are uniquely defined? Why or why not?

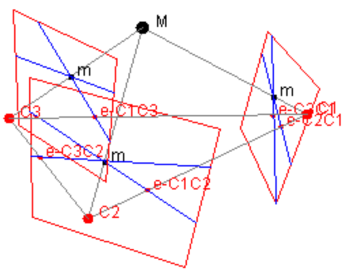


Figure 1: Epipolar geometry with 3 camera views.

Solution:
In this problem, we have to use the same kind of reasoning that we do with the two camera model but now extend the situation to the three cameras framework.

First of all, let us consider the epipoles, which are the red points that marked by e' in each image plane in the Figure 1. As we known in the two camera geometry setup, an epipole is the intersection points of the line run through the optical center of each camera with each image plane. In the stereo framework with two cameras we had two epipoles since we had two image plane and only one line to connect the two projection center. As for three cameras framework, we have 3 lines that connect the optical center of each camera. Now we have 6 epipoles which are represented in the figure 1.

Secondly, the three optical centers of each camera belong to only one single plane (three non-collinear points can define a plane) which is called the trifocal plane. Because of this set-up we can theoretically have a point wise relationship. In fact, in the standard two cameras stereo model, we have a dimensionality reduction, which is epipolar constraint. In this case, we are going to reduce the dimensionality of the corresponding match to only one single point resulting from the intersection of two epipolar lines (blue lines in the Figure 1) that coming from the epipoles of the other two cameras.

Now we discuss the intersection point in image plane of the third camera that intersected by the epipolar lines from the other two cameras. Because these two epipolar lines are the potential matching point, which is the projection of point M in the image plane of the third camera. ***But the intersection point is the only one choice point m (blue lines in each image intersect in one point m)that satisfy the epipolar constraint at the same time that presented by the other two cameras.*** So, with one more camera, we can define reduce the matching area to only one point other than a matching line that presented by epipolar constraint.

In a nut shell, if we have three cameras, the epipolar constraint can reduce from a line to a point that uniquely defined. So, we can define the point to the corresponding point other than corresponding line when we have three cameras.

Problem 2: Epiploar geometry and disparity forward translating camera (4pts)

In our course lecture, we discussed the epipolar geometry and fundamental matrix of a forward translating camera.

Given the definition of the epipolar plane as the plane defined by the two optical centers and a world point, explain why the geometry for a forward moving camera results in a set of radially oriented intersecting planes as shown in the following figure.

Given geometric considerations of a forward moving camera, develop a framework to calculate depth of world points from disparity as observed in consecutive frames. Make use of the definition of disparity as presented for the case of a stereo camera setup.

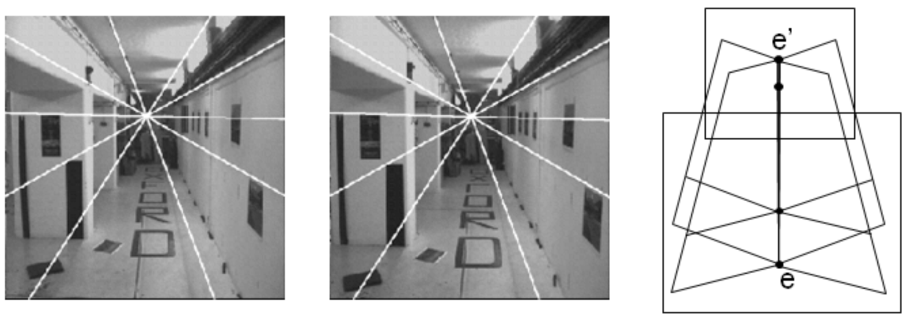


Figure 2: Epipolar geometry with Forward Camera

2.1 Solution:

To do so, we have to review the fundamental matrix that discussed in the class.

$$\vec{a} \times \vec{b} = \begin{bmatrix} 0 & -a_z & a_y \\ a_z & 0 & -a_x \\ -a_y & a_x & 0 \end{bmatrix} \begin{bmatrix} b_x \\ b_y \\ b_z \end{bmatrix} = \vec{c} \qquad \begin{matrix} \vec{a} \cdot \vec{c} = 0 \\ \vec{b} \cdot \vec{c} = 0 \end{matrix}$$

$$[a_x] = \begin{bmatrix} 0 & -a_z & a_y \\ a_z & 0 & -a_x \\ -a_y & a_x & 0 \end{bmatrix} \qquad \boxed{\vec{a} \times \vec{b} = [a_x]\vec{b}}$$

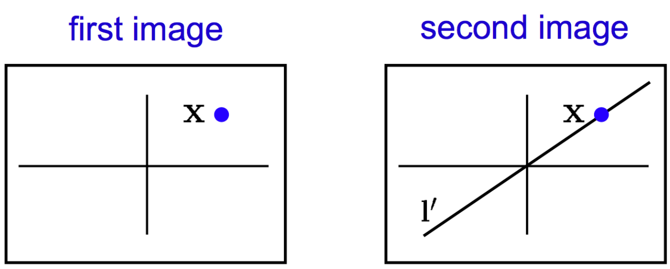
Because this situation is forward translation, we do not have to consider the rotation matrix. Assume, the direction of forward translation is along the Z-axis positively. Now we have the fundamental matrix.

$$\begin{aligned} \mathrm{F} &= \mathrm{K}'^{-\top}[\mathrm{t}]_\times \mathrm{R}\mathrm{K}^{-1} \\ &= \begin{bmatrix} 1/f & 0 & 0 \\ 0 & 1/f & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & -t_z & 0 \\ t_z & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1/f & 0 & 0 \\ 0 & 1/f & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \end{aligned}$$

So, we assume the forward distance in pixel unit is 1(tz = 1). So, the epipolar line for a point (x, y, 1) is:

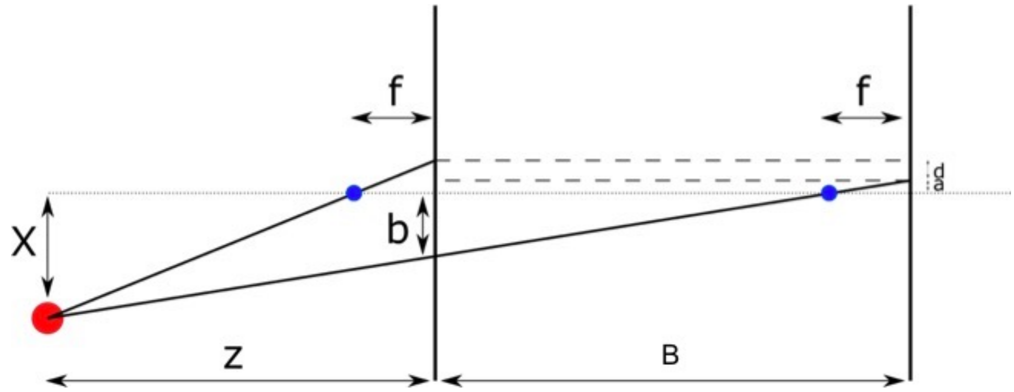
$$\mathrm{l}' = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} -y \\ x \\ 0 \end{pmatrix}$$

And we can present like the following image:



Because the epipolar lines will finally intersected in the same point (the original point of the image plane), which results in a set of radially oriented intersecting planes as shown in the above figure.

2.2 Solution:



We can set up our equation system like the above image. B defines the baseline between two cameras, Z defines the depth that point X correspond to the forward camera. **Used the principal of similar triangles**, we can find that:

$$\frac{b}{X} = \frac{Z}{(Z + B - f)}$$

For the backward camera settings, we have:

$$\frac{a}{b} = \frac{f}{(B - f)}$$

With these two equations, we replace b so we have:

$$a = \frac{f\,X\,Z}{(Z + B - f)(B - f)}$$

Finally, with the forward camera setting, we have:

$$\frac{(d + a)}{X} = \frac{f}{Z}$$

With the above equations, we can finally get the calculation framework of disparity(we can also calculate the Z based on this equation):

$$d = \frac{f\,X}{Z} - \frac{f\,X\,Z}{(Z - B - f)(B - f)}$$