Stereo Dense Reconstruction

1 Numerical Exercises

1. Consider a rectified stereo camera system (simplifed case) consisting of two equal cameras with an image sensor width of 300 pixels, height of 150 pixels, a focal length of 600 pixels and a baseline of 20 cm. Assume you have found the point projection of a 3D point $P = [X_p, Y_p, Z_p]$ in the left frame with pixel coordinates $p_l = [160, 50]$ and in right frame with pixel coordinates $p_r = [120, 50]$.

   (a) Compute the depth of point P.
   Solution:
   For a rectivied stereo system, we can compute the depth of point P using the following formula

   $$Z_P = \frac{bf}{u_l - u_r}$$

   In our case, we have a disparity $u_l - u_r = p_l - p_r = 160 - 120 = 40$ pixels, a baseline $b = 0.2m$ and a focal length $f = 600$ pixels. Plugging those values in the formula above gives us $Z_P = 3m$

   (b) Compute the X_p, Y_p coordinates of point P expressed in the coordinate system of the left camera.
   Solution:
   We can find X_p, Y_p using the image plane coordinates x, y in pixels and the perspective projection equations.

   $$\frac{x}{f} = \frac{X_p}{Z_p} \quad \Rightarrow \quad X_p = Z_p \frac{x}{f} \qquad\qquad \frac{y}{f} = \frac{Y_p}{Z_p} \quad \Rightarrow \quad Y_p = Z_p \frac{y}{f}$$

   To obtain the image plane coordinates x, y, we assume a simplifed camera model with the principal point exactly in the middle of the image sensor. Thus, we can simply subtract the center coordinates of the image sensor from the pixel coordinates p_l and p_r.

   $$\begin{bmatrix} x \\ y \end{bmatrix} = p_l - \begin{bmatrix} 150 \\ 75 \end{bmatrix} = \begin{bmatrix} 10 \\ -25 \end{bmatrix}$$

   Plugging in all the values in the perspective projection equations gives us the 3D point P expressed in meters

   $$P = [X_p, Y_p, Z_p] = [0.05, -0.125, 3]$$

   (c) What is the closest depth observable by this stereo camera system?
   Solution:
   The closest depth observable by a stereo camera system corresponds to the largest disparity detectable. In the given stereo system, it represents the pixel projection of a 3D point P in left camera with $p_l = 300$ and $p_r = 0$ leading to a disparity of $p_l - p_r = 300$. Thus, the closest observable depth can be computed as follows

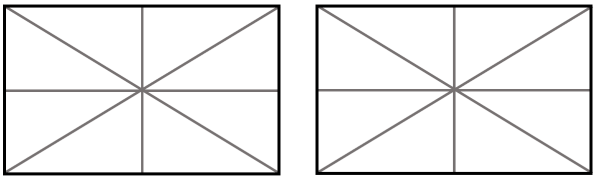
   $$Z_P = \frac{bf}{u_l - u_r} = 0.4m$$

(d) Which parameters of the camera system can be changed individually in order to measure 3D points even closer to the cameras?
Solution:
Based on the calculation in (c), the following parameters can be changed to enable closer depth measurments:

i. Decrease the focal length f
ii. Decrease the baseline
iii. Increase the width of the image sensor

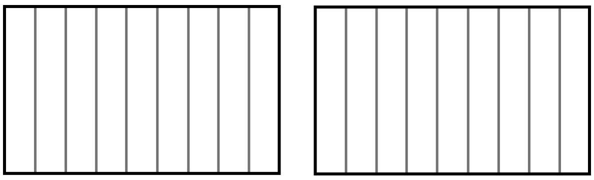
2. Draw the epipolar lines for a camera undergoing the following motions. Assume the standard coordinate system for cameras.

(a) Pure translation in Z direction.
Solution:
One easy way to draw epipolar lines is to think about where the camera center of the second camera lays in the camera frame, in which you want to draw the epipolar lines. After you have found this point, you can then just draw straight lines in all possible directions, which intersect at this point. In case of pure translation in Z direction, the camera center of the second camera is exactly in the middle of the frame.



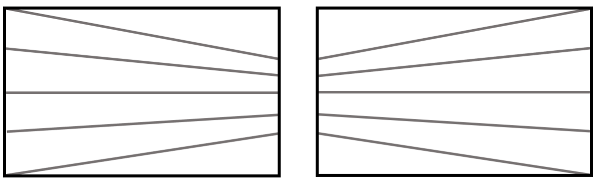
Frame 1 Frame 2

(b) Pure translation in Y direction.
Solution:



Frame 1 Frame 2

(c) 90-degree rotation around an axis, which represents the intersection of the image plane and the $Y - Z$ plane.
Solution:



Frame 1 Frame 2

3. For a rectified stereo camera system, derive mathematically the expression of depth uncertainty as a function of the disparity and as a function of the distance.
Solution:

To derive the uncertainty, we start with the formula used to compute the depth in the simplified stereo case

$$Z_P = \frac{bf}{d}$$

Next, we compute how the depth changes if we slightly perturb the distparity value. This is done by deriving for d.

$$\frac{\partial Z_P}{\partial d} = \frac{\partial}{\partial d}\frac{bf}{d} = -\frac{bf}{d^2}$$

Rearranging for the depth uncertainty ∂Z_P leads to the expression of depth uncertainty as a function of the disparity and the disparity error ∂d

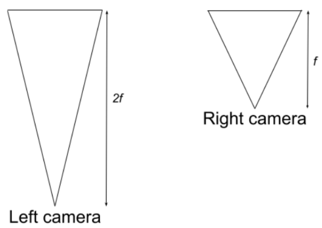
$$\partial Z_P = \frac{-bf}{d^2}\partial d$$

If we now substitue $d = \frac{bf}{Z_P}$ in the above equation, we get the expression of depth uncertainty as a function of the distance and the disparity error ∂d

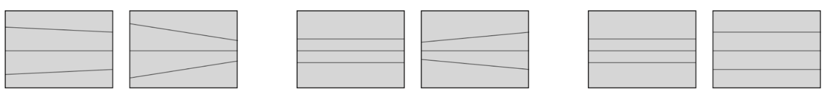
$$\partial Z_P = \frac{-{Z_P}^2}{bf}\partial d$$

Thus, it can be observed that points further away (larger depth) are more affected by disparity errors.

4. Which of the following sketches shows the epipolar lines corresponding to the setting of two cameras placed side-by-side? The focal length of the left camera is twice as large as the focal length of the right camera. Assume that the image planes of both cameras are coplanar as shown below.



Camera configuration



(a) (b) (c)

Solution: (a)