

Assignment 7 Problem 2

Question :- Derive procedure for estimating the object's size in Problem 1 using uncalibrated stereo. Show all calculations & derivations.

Answer :-

For uncalibrated stereo, the intrinsic and extrinsic camera parameters are not known.

But using epipolar constraint, points on left image can be mapped to " " right ". Given point p_L from left image and corresponding point in right image p_R :-

$$p_L = \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \text{ and } p_R = \begin{bmatrix} u' \\ v' \\ 1 \end{bmatrix}. \text{ The relationship between the } p_L \text{ and } p_R$$

is defined using Fundamental Matrix F :-

$$p_R^T F p_L = 0 \text{ or } [u' \ v' \ 1] F \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = 0.$$

Left camera is called C_L and right camera is called C_R . Making the C_L centered at origin :-

$C_L = [I \ | \ 0]$ where I is identity matrix and 0 is zero vector. C_L is 3×4 matrix.

For deriving C_R , it is assumed that e' is epipole in the right image in such a way that $F^T e' = 0$. C_R is :-

$C_R = [e']_x F \ | \ e']$ where $[e']_x$ is skew-symmetric matrix representation for epipole e' . C_R is also a 3×4 matrix.

continued on next page →

continued from previous page

Given 3D point in space called \vec{x} represented by vector $\begin{bmatrix} X \\ Y \\ Z \\ W \end{bmatrix}$. The projection equations are $P_L(C_L x) = 0$ and $P_R(C_R x) = 0$

Linear system $A\vec{x} = 0$ is used to represent the two equations :-

$$A = \begin{bmatrix} U C_L^3 - C_L^1 \\ V C_L^3 - C_L^2 \\ U' C_R^3 - C_R^1 \\ V' C_R^3 - C_R^2 \end{bmatrix} \quad \begin{array}{l} C_L^i \text{ is } i\text{th row in } C_L \\ C_R^i \text{ is } i\text{th row in } C_R \end{array}$$

$A\vec{x} = 0$ is solved using SVD (singular value decomposition). The result $\vec{x} = \begin{bmatrix} X \\ Y \\ Z \\ W \end{bmatrix}$ is in the projective space because the system is uncalibrated stereo. We don't know the scalar factor λ separating \vec{x} from the correct actual point \vec{x}_{true} . For finding \vec{x} , it is assumed that one real world dimension of known object is known. (Width of object)

Now to finally calculate the size of an object, the dimension is defined by endpoints A and B . A_L is that point A from left camera, and A_R is A from right camera. Same for B .

Next we perform triangulation between A_L and A_R and X_A . Same for B_L , B_R , and X_B .

continued on next page →

Continued from previous page

The homogeneous coordinates x_A and x_B should be divided by w_A and w_B to get real world / Cartesian points to get R_A and R_B . (Only 3 values, x, y, z)

Finally we can apply standard distance formula for R_A and R_B :-

$$d = \sqrt{(x_A - x_B)^2 + (y_A - y_B)^2 + (z_A - z_B)^2}$$

Finally applying the scale factor from before, λ :-

$$\text{size} = \lambda \cdot d$$