Stereo and Motion

Objects appear in different positions from different viewpoints - parallax

Disparity = Position difference, inversely proportional to depth

If we know disparity and viewpoints we can infer 3D scene structure.

Epipolar Geometry:

Geometry depends on position and orientation of cameras.

Camera Model:

Principal Point - Centre of the image plane (0, 0, f)

Focal Length - f

Optical/Principal Axis - Z

Point on image plane - P(x, y, f)

3D point on surface of object - P(X, Y, Z) projects to P(x, y, f)

$$x = \frac{fX}{Z}$$
 $y = \frac{fY}{Z}$ $p = \frac{fP}{Z}$

Simple Two-View Stereo:

Coplanar image planes, T= Baseline distance between the centre of projection (COP) of each camera

$$\frac{T}{Z} = \frac{T - x_L + x_R}{Z - f} \qquad x_L = \text{Point on left camera x} \qquad x_R = \text{Point on right camera x depth: } Z = \frac{fT}{x_L - x_R} = \frac{fT}{x_L$$

General Two-View Stereo:

Coordinate Transformations:

Rotate coordinate frame clockwise by θ : $v' = \begin{bmatrix} \cos(-\theta) & \sin(-\theta) \\ -\sin(-\theta) & \cos(-\theta) \end{bmatrix} v$ Rotation and Translation: $v' = \begin{bmatrix} \cos(-\theta) & \sin(-\theta) \\ -\sin(-\theta) & \cos(-\theta) \end{bmatrix} v$

 $T=3\mathrm{D}$ Camera Position Vector $R=3\mathrm{D}$ Camera Rotation Matrix Vector defining P in camera coordinates

Note: For rotation matrices: $R^T = R^{-1}$

Homogenous Coordinates:

We can define rotation and translation in a single 4x4 matrix

Camera To World =
$$H_{CW}P_W' = \begin{bmatrix} X_W \\ Y_W \\ Z_W \\ 1 \end{bmatrix} = \begin{bmatrix} P_W \\ 1 \end{bmatrix} = \begin{bmatrix} R^T & T \\ 0 & 1 \end{bmatrix} \begin{bmatrix} P_C \\ 1 \end{bmatrix} = H_{CW}P_c' \implies P_W = R^TP_C + T$$

$$H_{WC} = \begin{bmatrix} R & -RT \\ 0 & 1 \end{bmatrix} \implies P_C = R(P_W - T)$$

Stereo Coordinate Systems:

 $P_L = R^T P_R + T P_R = R(P_L - T)T = \text{Translation from left to right camera} R = \text{Rotation applied to right camera}$

Epipolar Geometry:

$$p_L = \begin{bmatrix} X_L \\ Y_L \\ f \end{bmatrix} = \frac{fP_L}{Z_L} \qquad p_R = \begin{bmatrix} X_R \\ Y_R \\ f \end{bmatrix} = \frac{fP_R}{Z_R}$$

Vectors P_L T and P_L - T all lie in an epipolar plane.

$$(P_L - T)^T (T * P_L) = 0 \text{Cross product can be defined by } S = \begin{bmatrix} 0 & -T_Z & T_Y \\ T_Z & 0 & -T_X \\ -T_Y & T_X & 0 \end{bmatrix} P_R = R(P_L - T)R_T P_R = (P_L - T)R_T P_R =$$

All points in the view plane satisfy the constraint above. RS is known as the essential matrix = E

$$E = RS$$
 $P_R^T E P_L = 0 P_L = \frac{Z_L p_L}{f}$ $P_R = \frac{Z_R p_R}{f} p_R^T E p_L = 0$ Therefore the two D points in the view plane