Stereo and Motion

Objects appear in different positions from different viewpoints - parallax

Disparity = Position difference, inversely proportional to depth

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

Epipolar Geometry:

Geometry depends on postion 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)

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

Simple Two-View Stereo:

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

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

General Two-View Stereo:

Coordinate Transformations:

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

T = 3D Camera Position Vector $\qquad R$ = 3D Camera Rotation MatrixVector defining P in camera coordinates

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