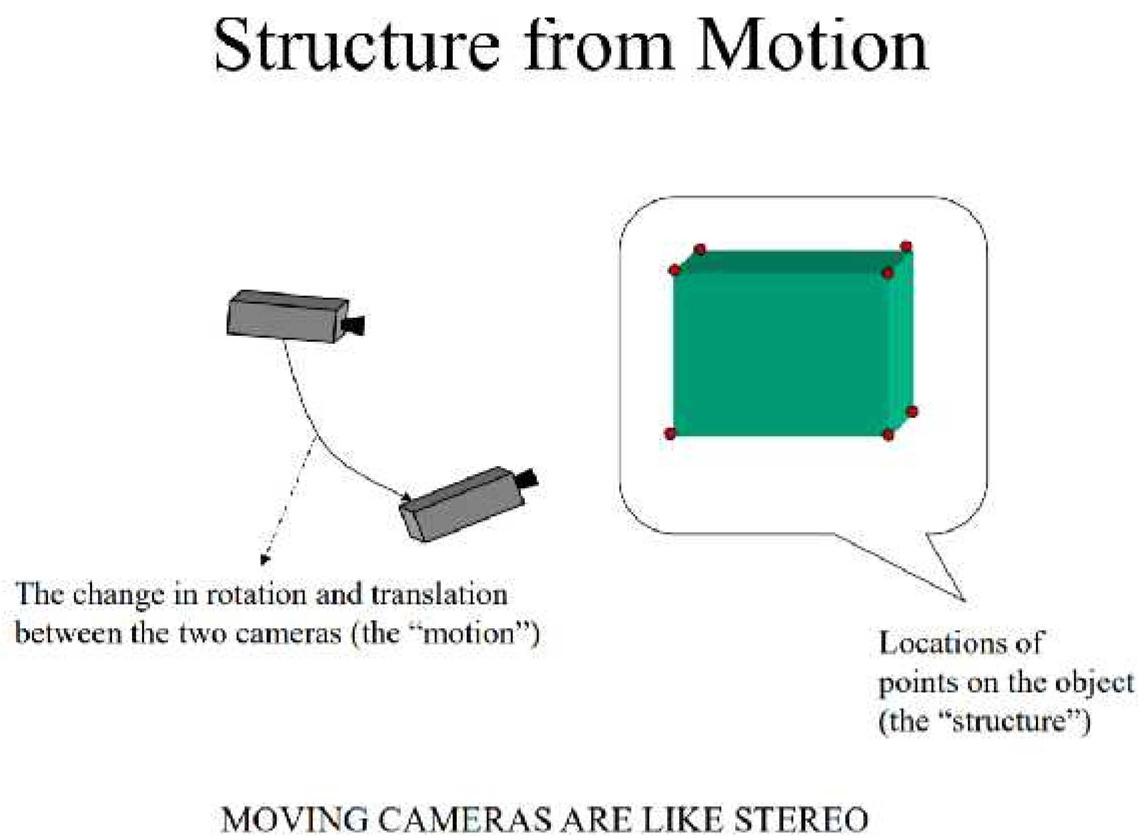
18CSE390T
Computer Vision

S1-SLO2-Two-frame structure from motion

Structure From Motion



Two-Frame Structure from Motion

- In 3D reconstruction we have always assumed that either 3D points position or the 3D camera poses are known in advance.
- Simultaneous recovery of 3D structure and pose from image correspondences

Two-Frame Structure from Motion (cont).

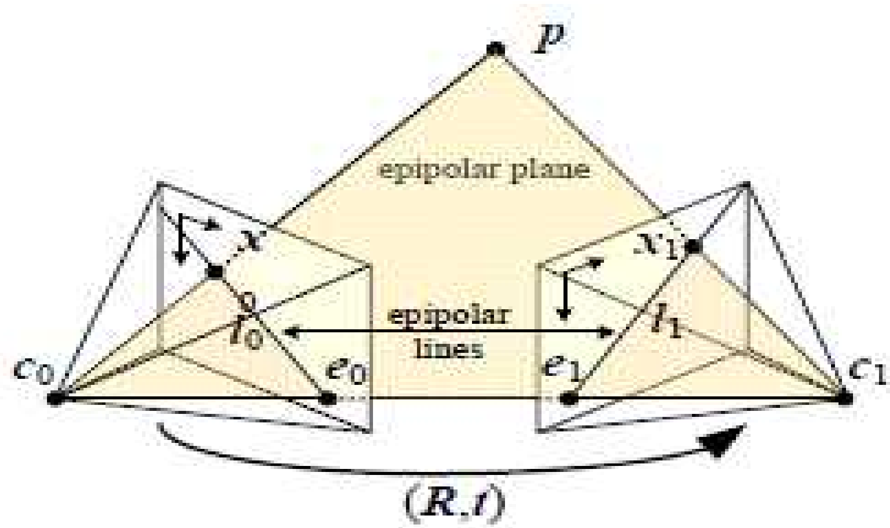


Figure: Epipolar geometry: The vectors $t=c_1 - c_0$, $p - c_0$ and p-$c_1$are co-planar and the basic epipolar constraint expressed in terms of the pixel measurement x_0 and x_1

Two-Frame Structure from Motion (cont).

- The observed location of point *p* in the first image, is mapped into the second image by the transformation

$$p_0 = d_0 \hat{x}_0$$

$$d_1 \hat{x}_1 = p_1 = R p_0 + t = R(d_0 \hat{x}_0) + t,$$

: t ray

$$\hat{x}_j = K_j^{-1} x_j$$

Two-Frame Structure from Motion (cont).

- Taking the cross product of both the sides with *t* in order to annihilate it on the right hand side yields

- Taking the dot product of both the sides with $d_1[\boldsymbol{t}]_\times \hat{\boldsymbol{x}}_1 = d_0[\boldsymbol{t}]_\times \boldsymbol{R}\hat{\boldsymbol{x}}_0.$ $\hat{\boldsymbol{x}}_1$

  yields $d_0\hat{\boldsymbol{x}}_1^T([\boldsymbol{t}]_\times \boldsymbol{R})\hat{\boldsymbol{x}}_0 = d_1\hat{\boldsymbol{x}}_1^T[\boldsymbol{t}]_\times \hat{\boldsymbol{x}}_1 = 0,$

Two-Frame Structure from Motion (cont).

- The right hand side is triple product with two identical entries
- We therefore arrive at the basic epipolar constraint

: ***essential matrix*** $$\hat{\boldsymbol{x}}_1^T \boldsymbol{E}\, \hat{\boldsymbol{x}}_0 = 0,$$

$$\boldsymbol{E} = [\boldsymbol{t}]_\times \boldsymbol{R}$$