18CSE390T
Computer Vision

S1-SLO1-Triangulation

Structure from Motion

- Study of visual perception.
- Process of finding the three-dimensional structure of an object by analyzing local motion signals over time.
- Method for creating 3D models from 2D pictures of an object.

Triangulation

- A problem of estimating a point's 3D location when it is seen from multiple cameras is known as *triangulation.*
- It is a converse of pose estimation problem.
- Given projection matrices, 3D points can be computed from their measured image positions in two or more views.

Triangulation (cont).

- Find the 3D point *p* that lies closest to all of the 3D rays corresponding to the 2D matching feature locations $\{x_j\}$ observed by cameras

  $\{P_j = K_j [R_j \mid t_j]\}$

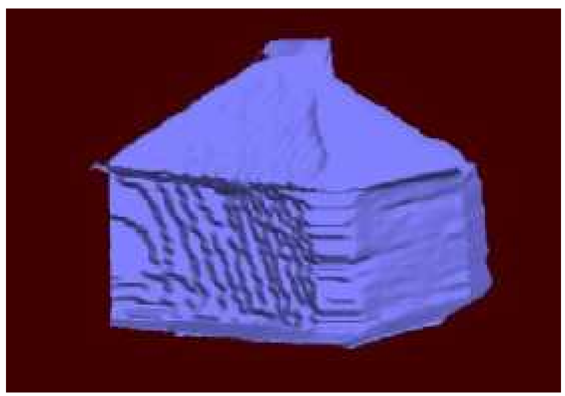
$$t_j = -R_j c_j$$

c_j is the jth camera center

Example

Picture 1

Picture 2



3D model created from the two images

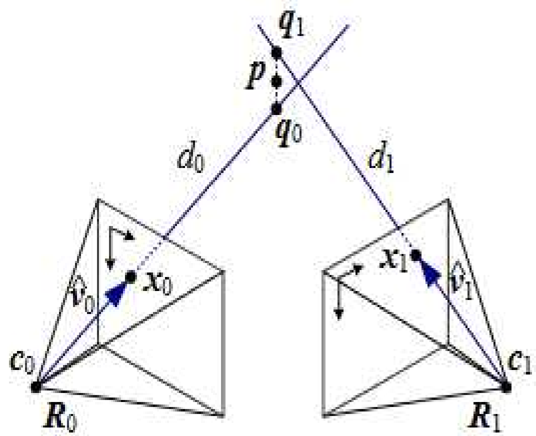


Figure 7.2 3D point triangulation by finding the point $\boldsymbol{p}$ that lies nearest to all of the optical rays $\boldsymbol{c}_j + d_j\hat{\boldsymbol{v}}_j$.

Triangulation (cont).

- The rays originate at c_j in a direction

$$\hat{v}_j = \mathcal{N}(R_j^{-1} K_j^{-1} x_j).$$

- The nearest point to p on this ray, which is denoted as q_j, minimizes the distance.

$$\|c_j + d_j \hat{v}_j - p\|^2$$

which has a minimum at

$$d_j = \hat{v}_j \cdot (p - c_j).$$

Hence,

$$q_j = c_j + (\hat{v}_j \hat{v}_j^T)(p - c_j) = c_j + (p - c_j)_{\|},$$

Triangulation (cont).

- Alternative formulation which is optimal and can produce better estimates if some of the cameras are closer to the 3D points than the others, it minimizes the residual in the measurement equations.

$$x_j = \frac{p_{00}^{(j)} X + p_{01}^{(j)} Y + p_{02}^{(j)} Z + p_{03}^{(j)} W}{p_{20}^{(j)} X + p_{21}^{(j)} Y + p_{22}^{(j)} Z + p_{23}^{(j)} W}$$

Triangulation (cont).

$$y_j = \frac{p_{10}^{(j)}X + p_{11}^{(j)}Y + p_{12}^{(j)}Z + p_{13}^{(j)}W}{p_{20}^{(j)}X + p_{21}^{(j)}Y + p_{22}^{(j)}Z + p_{23}^{(j)}W},$$

(x_j, y_j): the measured 2D feature location

$\{p_{00}{}^{(j)} \ldots . p_{23}{}^{(j)}\}$: the known entries in camera matrix p_j.

Triangulation (cont).

- The squared distance between p and q_j is

$$r_j^2 = \|(\boldsymbol{I} - \hat{\boldsymbol{v}}_j \hat{\boldsymbol{v}}_j^T)(\boldsymbol{p} - \boldsymbol{c}_j)\|^2 = \|(\boldsymbol{p} - \boldsymbol{c}_j)_\perp\|^2.$$

- The optimal value for p, which lies closest to all of the rays, can be computed as a regular least square problem by summing over all the r_j^2 and finding the optimal value of p,

$$\boldsymbol{p} = \left[\sum_j (\boldsymbol{I} - \hat{\boldsymbol{v}}_j \hat{\boldsymbol{v}}_j^T)\right]^{-1} \left[\sum_j (\boldsymbol{I} - \hat{\boldsymbol{v}}_j \hat{\boldsymbol{v}}_j^T)\boldsymbol{c}_j\right].$$