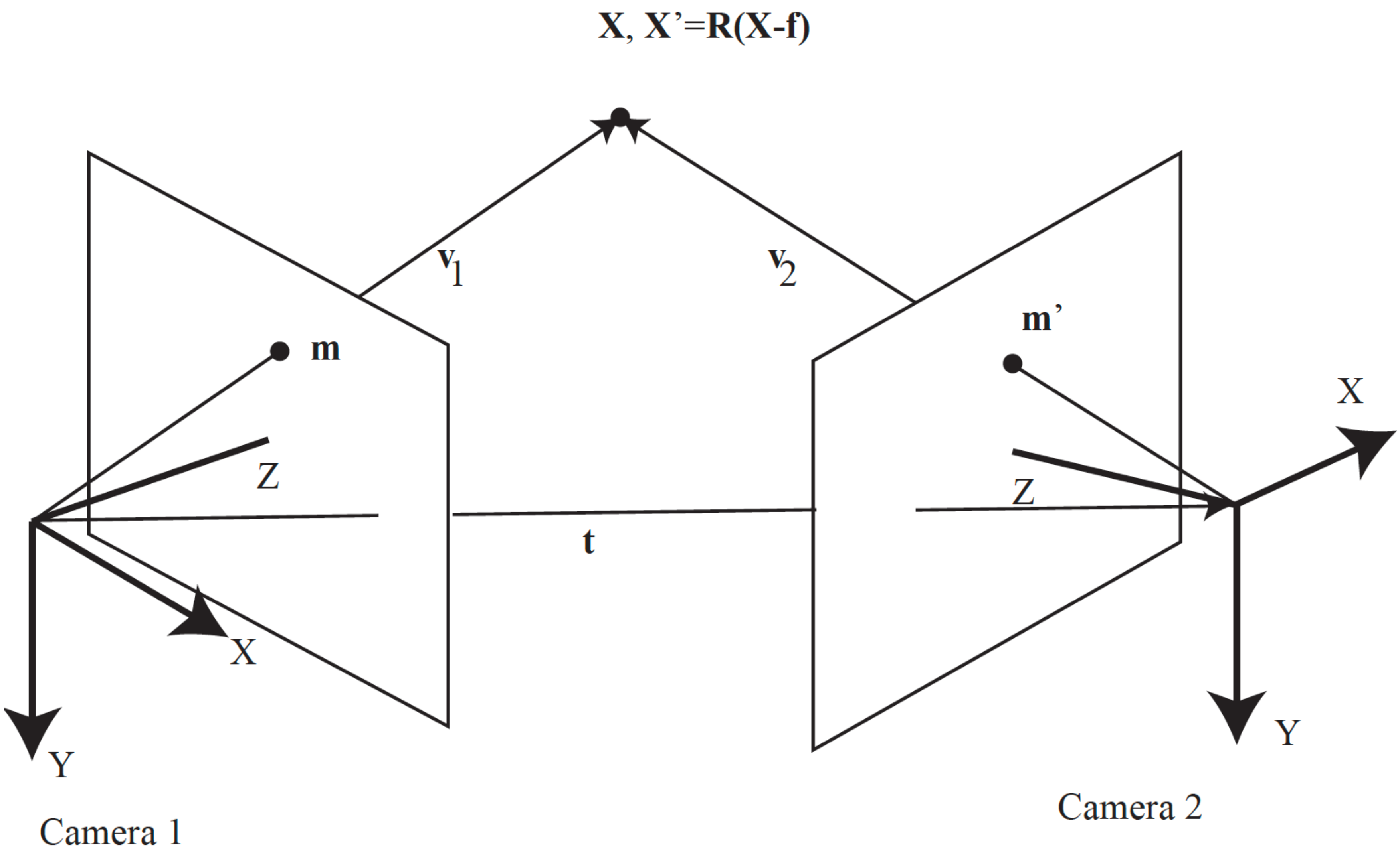
Coordinate geometry: Triangulation

D.A. Forsyth,

University of Illinois at Urbana Champaign

Reconstruct point from two views

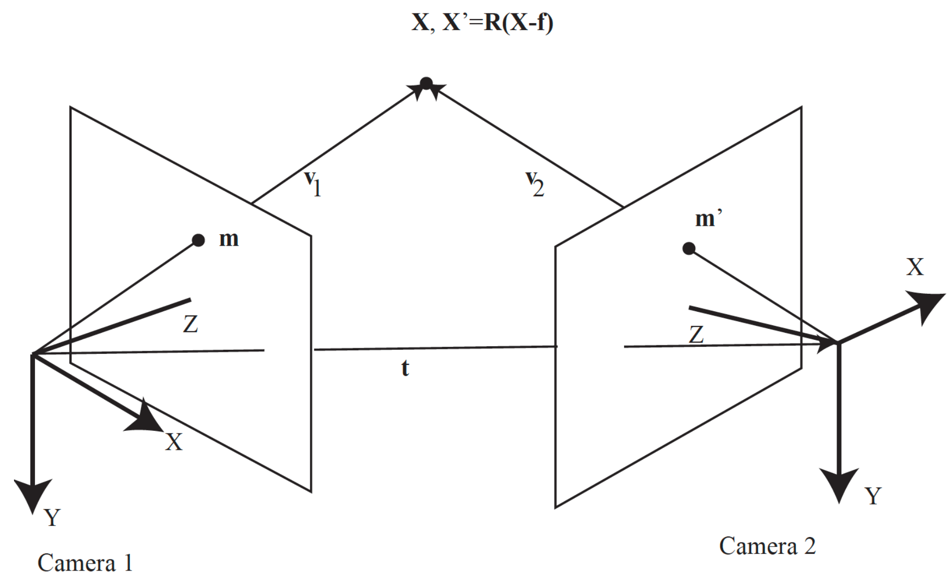


Notice

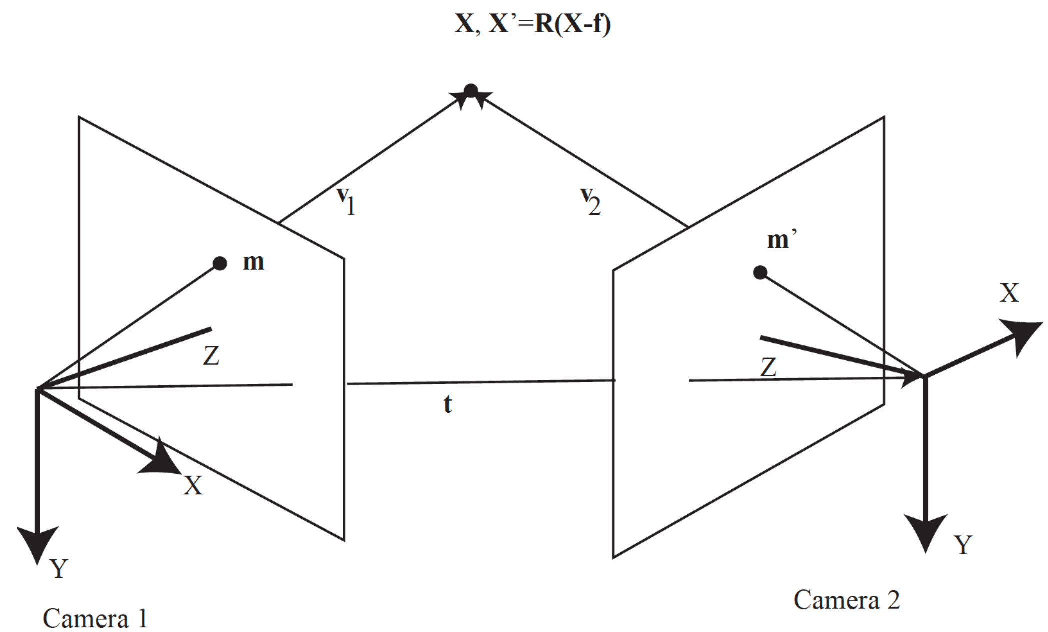
- m and m’ are in world coordinates

- But
  - you have calibrated cameras
  - so if x, x’ are in camera coords

$$\mathbf{m} = \mathcal{K}_1^{-1}\mathbf{x}$$
$$\mathbf{m}' = \mathcal{K}_2^{-1}\mathbf{x}'$$



Now choose a coordinate system so that the first camera has focal point at the origin, looks down the z-axis, and has image plane at $z = 1$, as in Section 30.2. To get the second camera, rotate the first camera by $\mathcal{R}^T$, then translate it by $\mathbf{t}$, so that $\mathbf{f}_2 = \mathbf{t}$. Notice that this means that a point at $\mathbf{X} = [X_1, X_2, X_3]^T$ in the first camera's coordinate system appears at $\mathbf{X}' = \mathcal{R}(\mathbf{X} - \mathbf{t})$ (if the camera rotates left, then all the points in the image frame move right). Figure 32.3 shows this setup.



You see the point in the first camera at

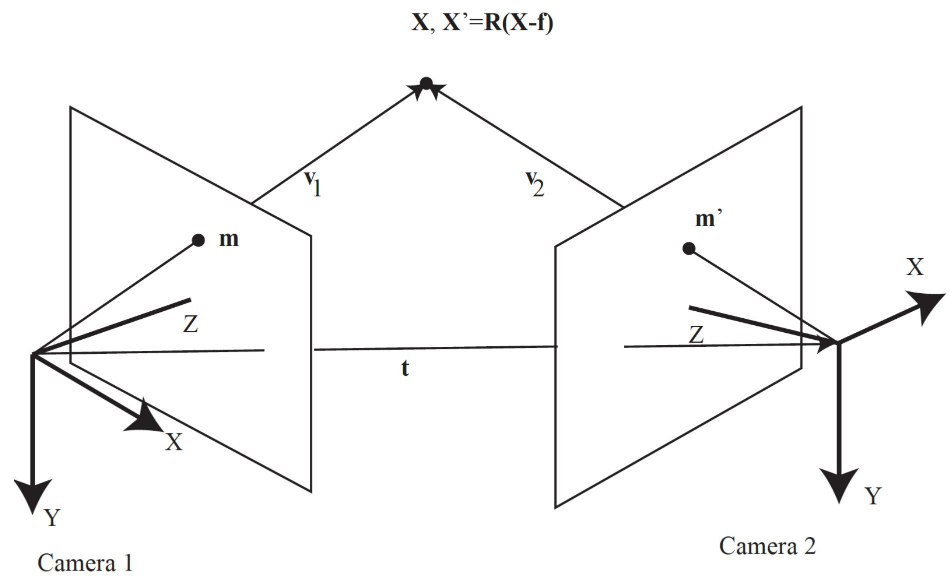
$$\mathbf{m} = \begin{bmatrix} m_1 \\ m_2 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{X_1}{X_3} \\ \frac{X_2}{X_3} \\ 1 \end{bmatrix}$$

and in the second camera at

$$\mathbf{m}' = \begin{bmatrix} m_1' \\ m_2' \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{\mathbf{r}_1^T(\mathbf{X}-\mathbf{t})}{\mathbf{r}_3^T(\mathbf{X}-\mathbf{t})} \\ \frac{\mathbf{r}_2^T(\mathbf{X}-\mathbf{t})}{\mathbf{r}_3^T(\mathbf{X}-\mathbf{t})} \\ 1 \end{bmatrix}.$$

For the moment, assume that $\mathbf{m}$, $\mathbf{m}'$, the rotation and the translation are all known exactly. Then

$$X_3\mathbf{m} = \mathbf{X}$$



and the only unknown is X_3. You can write two linear equations in this unknown, which are

$$\begin{aligned} m_1'(\mathbf{r}_3^T(X_3\mathbf{m}-\mathbf{t})) - (\mathbf{r}_1^T(X_3\mathbf{m}-\mathbf{t})) &= 0 \\ m_2'(\mathbf{r}_3^T(X_3\mathbf{m}-\mathbf{t})) - (\mathbf{r}_2^T(X_3\mathbf{m}-\mathbf{t})) &= 0. \end{aligned}$$

These equations have to be consistent, which means that there is a relationship between $\mathbf{m}$ and $\mathbf{m}'$ that depends on $\mathcal{R}$ and $\mathbf{t}$. The relationship expresses the mapping from points to lines of the previous section. It could be obtained by some aggressive linear algebra, but is better constructed directly, which I do in the next section.

A warning: it is not a good idea to *estimate* X_3 using the equations above, because you will never actually know $\mathbf{m}$ and $\mathbf{m}'$ exactly. They are useful only to establish that you can recover X_3.

Triangulation

- Have m, m’, R, t
- Compute reprojection error
  - sum:
    - residual between X projected into 1 and m
    - residual between X projected into 2 and m’
- Choose X that minimizes

Now assume you know $\mathcal{R}$ and $\mathbf{t}$, the intrinsic calibration matrices $\mathcal{K}$ and $\mathcal{K}'$, and have estimated locations $\mathbf{m}$ and $\mathbf{m}'$ for a pair of points that correspond. These estimates may not be exact – for example, they might come from an interest point matcher – but any error is small. You must recover the point in 3D.

The first camera has camera matrix $\mathcal{K}\mathcal{C}_p$. The second camera has camera matrix

$$\mathcal{K}' \left[\mathcal{R} | -\mathcal{R}\mathbf{t}\right]$$

(recall notation from Section 21.3, and check that this camera has focal point at $\mathbf{t}$).

Now write $\mathbf{X} = [X_1, X_2, X_3]$ for a point in 3D in affine coordinates. The residual vector in camera 1 is the vector from the projection of $\mathbf{X}$ to $\mathbf{m}$, so

$$\mathbf{e}_1(\mathbf{X}) = \left[\begin{array}{c} k_{11}\frac{X_1}{X_3} + k_{12}\frac{X_2}{X_3} + k_{13}\frac{1}{X_3} - m_1 \\ k_{22}\frac{X_2}{X_3} + k_{23}\frac{1}{X_3} - m_2 \end{array} \right].$$

The residual vector in camera 2 is the vector from the projection of $\mathbf{X}$ to $\mathbf{m}$, so

$$\mathbf{e}_2(\mathbf{X}) = \left[\begin{array}{c} k'_{11}\frac{\mathbf{r}_1^T(\mathbf{X}-\mathbf{t})}{\mathbf{r}_3^T(\mathbf{X}-\mathbf{t})} + k'_{12}\frac{\mathbf{r}_2^T(\mathbf{X}-\mathbf{t})}{\mathbf{r}_3^T(\mathbf{X}-\mathbf{t})} + k'_{13}\frac{1}{\mathbf{r}_3^T(\mathbf{X}-\mathbf{t})} - m'_1 \\ k'_{22}\frac{\mathbf{r}_2^T(\mathbf{X}-\mathbf{t})}{\mathbf{r}_3^T(\mathbf{X}-\mathbf{t})} + k'_{23}\frac{1}{\mathbf{r}_3^T(\mathbf{X}-\mathbf{t})} - m'_2 \end{array} \right].$$

The *reprojection error* $E_r(\mathbf{X})$ for a point $\mathbf{X}$ in 3D is the sum of distances in each camera from the projections of the point to the measured locations, so

$$E_r(\mathbf{X}) = \mathbf{e}_1^T\mathbf{e}_1 + \mathbf{e}_2^T\mathbf{e}_2.$$

It is natural to obtain $\mathbf{X}$ by simply minimizing the reprojection error.

Start point

- Least squares on

$$
\begin{aligned}
m'_1(\mathbf{r}_3^T(X_3\mathbf{m} - \mathbf{t})) - (\mathbf{r}_1^T(X_3\mathbf{m} - \mathbf{t})) &= 0 \\
m'_2(\mathbf{r}_3^T(X_3\mathbf{m} - \mathbf{t})) - (\mathbf{r}_2^T(X_3\mathbf{m} - \mathbf{t})) &= 0.
\end{aligned}
$$

Triangulation

Procedure: 32.4 *Triangulating by minimizing reprojection error*

Start with a point $\mathbf{c}$ viewed in a calibrated camera and a corresponding point $\mathbf{c}'$ in a second calibrated camera. The rotation $\mathcal{R}$ and translation $\mathbf{t}$ from the first to the second camera are known. The first camera's intrinsic matrix is $\mathcal{K}$ and the second camera's is $\mathcal{K}'$. Compute the reprojection error $E_r(\mathbf{X})$ for a variable 3D point $\mathbf{X}$ and minimize

$$E_r(\mathbf{X})$$

as a function of $\mathbf{X}$. Use a quasi-newton method for minimization.