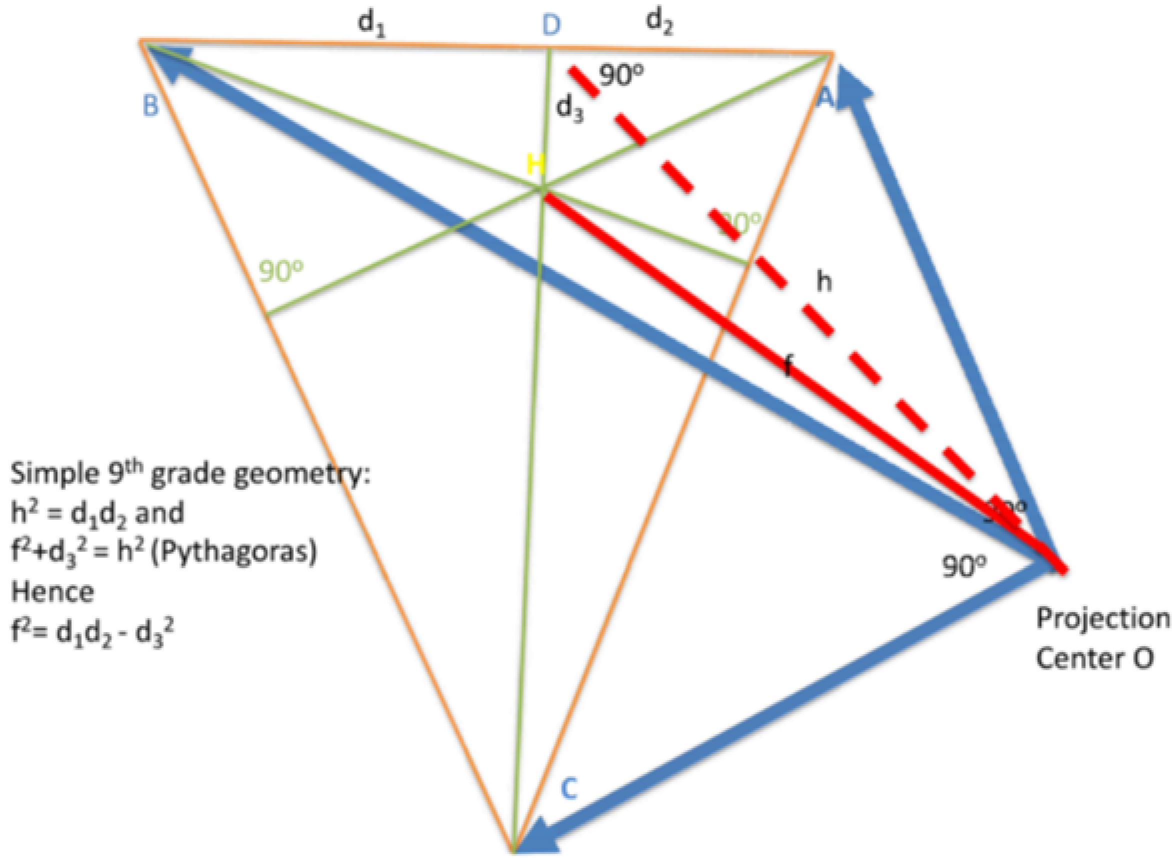
Focal length calculations:



$$f = \sqrt{d_1 d_2 - d_3^2} \times \frac{\text{pixel width}}{\text{measured width}}$$

Alternative method:

Denote two vanishing points coordinates (u_0, v_0, w_0) and (u_1, v_1, w_1) with the orthocenter as origin. Then $\frac{(u_1 u_2 + v_1 v_2)}{f^2} + w_1 w_2 = 0, f = \sqrt{\frac{-w_1 w_2}{u_1 u_2 + v_1 v_2}}$

Localization:

- **How to check if a transformation is a projective transformation:**

H is a transformation from $\mathbb{P}^2$ to $\mathbb{P}^2$:

$$H \sim K \begin{pmatrix} r_1 & r_2 & T \end{pmatrix}$$

Is it a projective transformation? Let us inspect its determinant:

$$\det \begin{pmatrix} r_1 & r_2 & T \end{pmatrix} = T^T (r_1 \times r_2)$$

which vanishes only if the camera lies in the ground plane $Z = 0$. In this case all points would project on a line.

Since $\det(K) = f^2$, H is invertible iff

$$T^T (r_1 \times r_2) \neq 0$$

- **Explain how you can find the pose R, T of a camera given the projection of four coplanar points whose coordinates are known in the world.**

1. Find H up to a scale factor from the point coorrespondences

2. Compute $H' = K^{-1}H$. Let H''s columns be $\begin{pmatrix} a & b & c \end{pmatrix}$

3. Minimize

$$\|\begin{pmatrix} a & b & c \end{pmatrix} - \lambda \begin{pmatrix} r_1 & r_2 & T \end{pmatrix}\|_F$$

w.r.t. $\lambda \in \mathbb{R}, r_1, r_2, T \in \mathbb{R}^3$
s.t. $r_1^T r_2 = 0$ and $\|r_1\| = \|r_2\| = 1$

Let

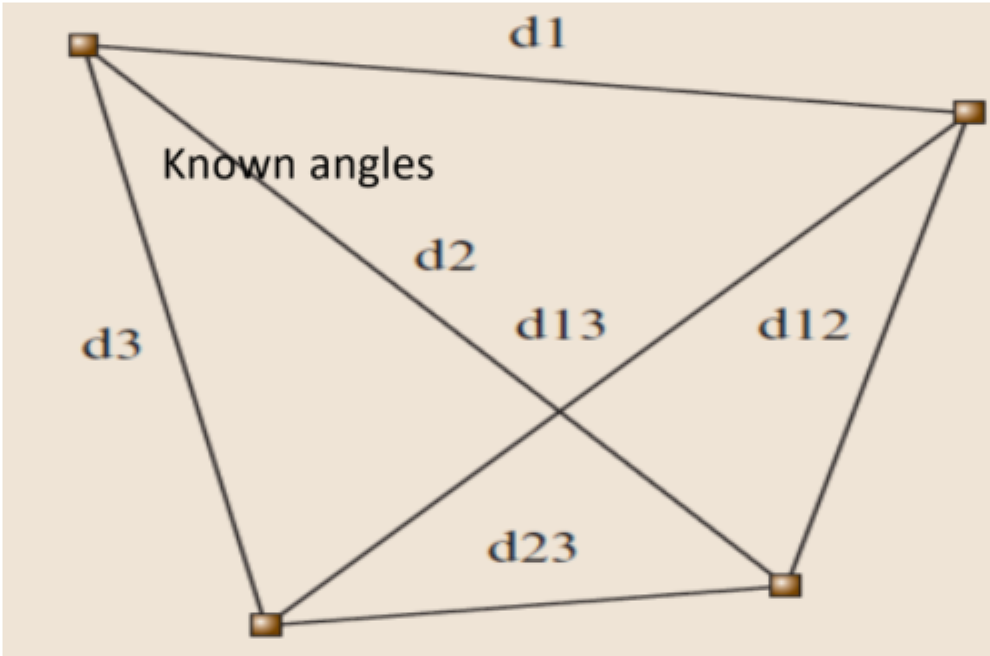
$$\begin{pmatrix} a & b & c \end{pmatrix} = U_{3x2} \begin{pmatrix} s_1 & 0 \\ 0 & s_2 \end{pmatrix} V_{2x2}^T.$$

Then

$$\begin{pmatrix} r_1 & r_2 \end{pmatrix} = U_{3x2} V_{2x2}^T \quad \text{and} \quad \lambda = \frac{s_1 + s_2}{2}$$

4. $T = c/\lambda$ and $R = \begin{pmatrix} r_1 & r_2 & r_1 \times r_2 \end{pmatrix}$. Make R to have determinant.

- **PNP problem with 3 points:**



$$d_i^2 + d_j^2 - 2d_i d_j \cos\delta_{ij} = d_{ij}^2$$

- **Procrustes**

Returning to the Procrustes problem (6.4.1), if $Q \in \mathbb{R}^{p\times p}$ is orthogonal, then

$$\begin{aligned} \| A - BQ \|_F^2 &= \sum_{k=1}^{p} \| A(:,k) - B\cdot Q(:,k) \|_2^2 \\ &= \sum_{k=1}^{p} \| A(:,k) \|_2^2 + \| BQ(:,k) \|_2^2 - 2Q(:,k)^T B^T A(:,k) \\ &= \| A \|_F^2 + \| BQ \|_F^2 - 2\sum_{k=1}^{p} \left[Q^T(B^TA)\right]_{kk} \\ &= \| A \|_F^2 + \| B \|_F^2 - 2\mathrm{tr}(Q^T(B^TA)). \end{aligned}$$

Thus, (6.4.1) is equivalent to the problem

$$\max_{Q^TQ=I_p} \mathrm{tr}(Q^TB^TA).$$

If $U^T(B^TA)V = \Sigma = \mathrm{diag}(\sigma_1, \ldots, \sigma_p)$ is the SVD of B^TA and we define the orthogonal matrix Z by $Z = V^TQ^TU$, then by using (6.4.2) we have

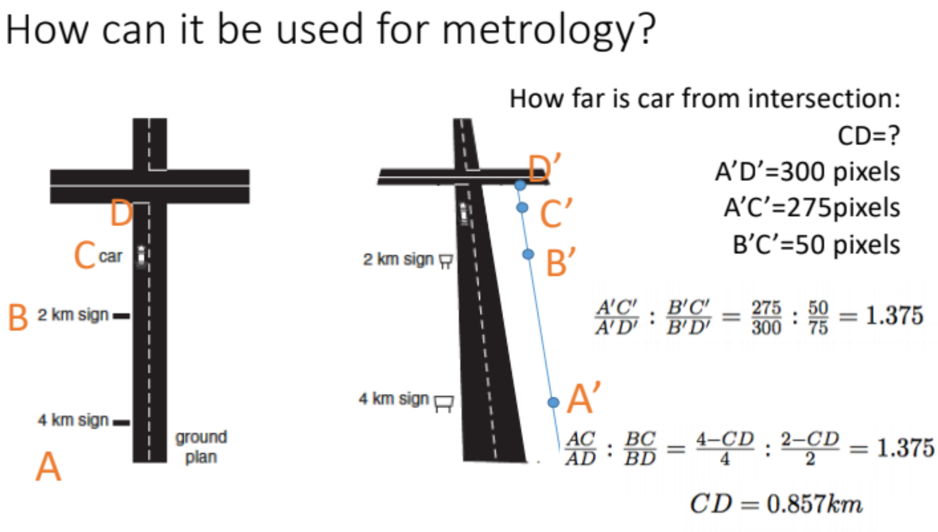
$$\mathrm{tr}(Q^TB^TA) = \mathrm{tr}(Q^TU\Sigma V^T) = \mathrm{tr}(Z\Sigma) = \sum_{i=1}^{p} z_{ii}\sigma_i \le \sum_{i=1}^{p} \sigma_i .$$

The upper bound is clearly attained by setting $Z = I_p$, i.e., $Q = UV^T$.

- **When to use PnP vs Procrustes**

- These two methods are used for localization and both use point-to-point correspondences
- PNP
    - For PNP problem we are given N correspondence (Xi,Yi,Zi,xi,yi); i.e. world coordinates and the corresponding ray direction in the camera system
- Procrustes
    - For procrustes problems the n-dimension correspondences are already given, for example, in 3d, the correspondences are (X1i,Y1i,Z1i,X2i,Y2i,Z2i); i.e. world coordinates and corresponding camera coordinates
    - For the ray direction, you don't know the depth, so there would be an unknown λ.

Distance Transfer and Cross Ratios



[Goal Directed Video Metrology]