Robot Perception: Pose from 3D Point Correspondences or the Procrustes Problem

Kostas Daniilidis

Procrustes Problem

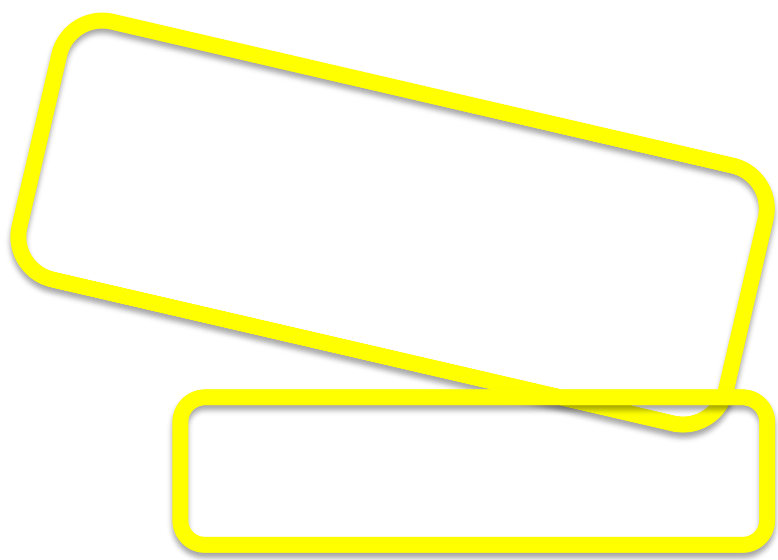
Given two shapes find the scaling, rotation, and translation that fits one into the other.

3D-3D Pose or Procrustes Problem

Given correspondences of points $A_i \in \mathbb{R}^3$ and $B_i \in \mathbb{R}^3$ find the scaling, rotation, and translation transformation, called *similitude* transformation, that satisfies

$$A_i = sRB_i + T$$

for $R \in SO(3)$, $T \in \mathbb{R}$, and $s \in \mathbb{R}^+$.



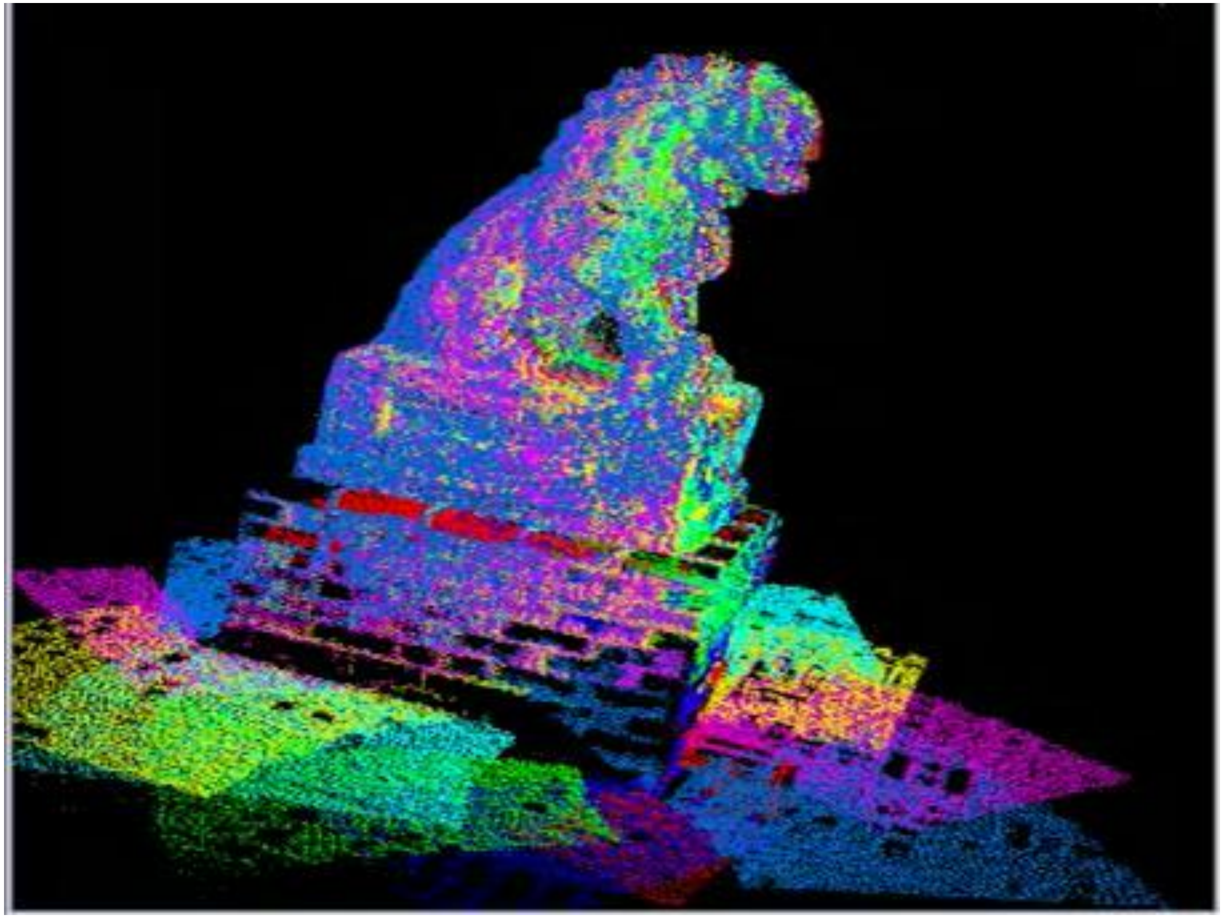
3D-3D Pose or Procrustes Problem

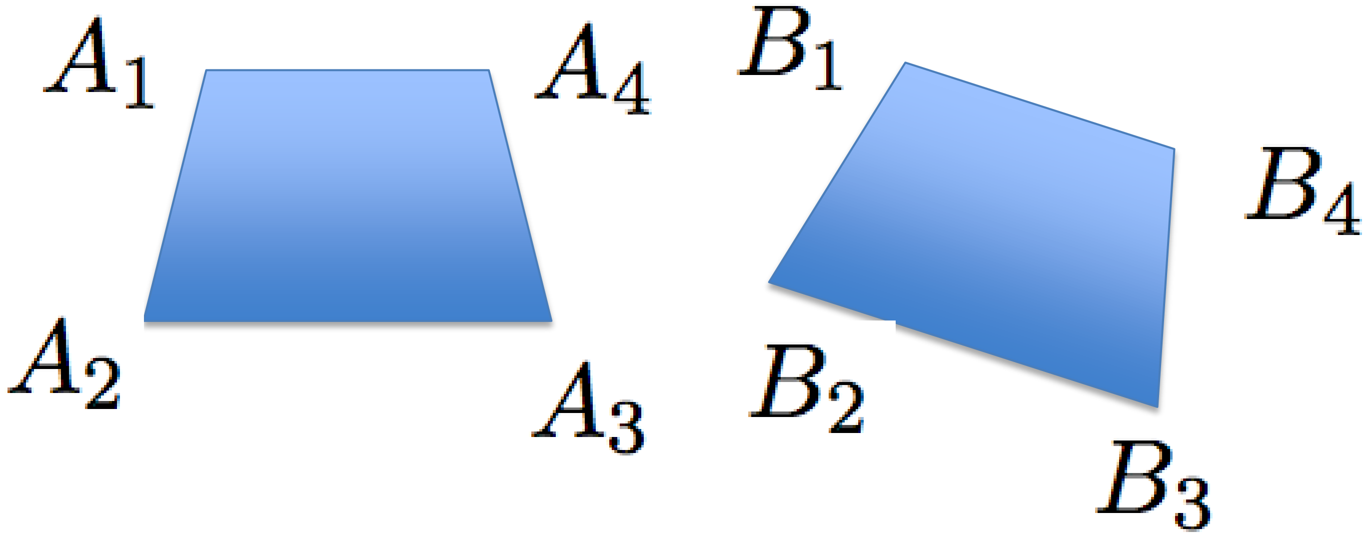
In the camera rigid pose problem scale $s = 1$ is known:

$$Z_i p_i^{cam} = R P_i^{obj} + T$$

This is the last step of the P3P problem or the entire problem of finding rigid pose when we know the depth at every point (e.g., in am RGB-D sensor).

3D-3D Registration enables the creation of 3D models from multiple point clouds:

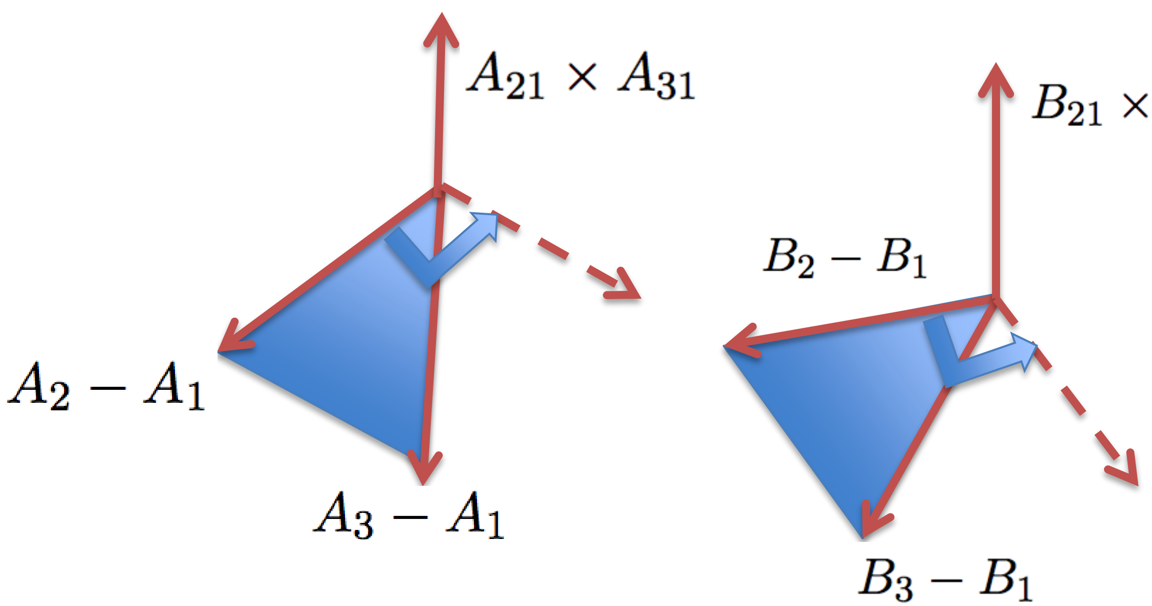


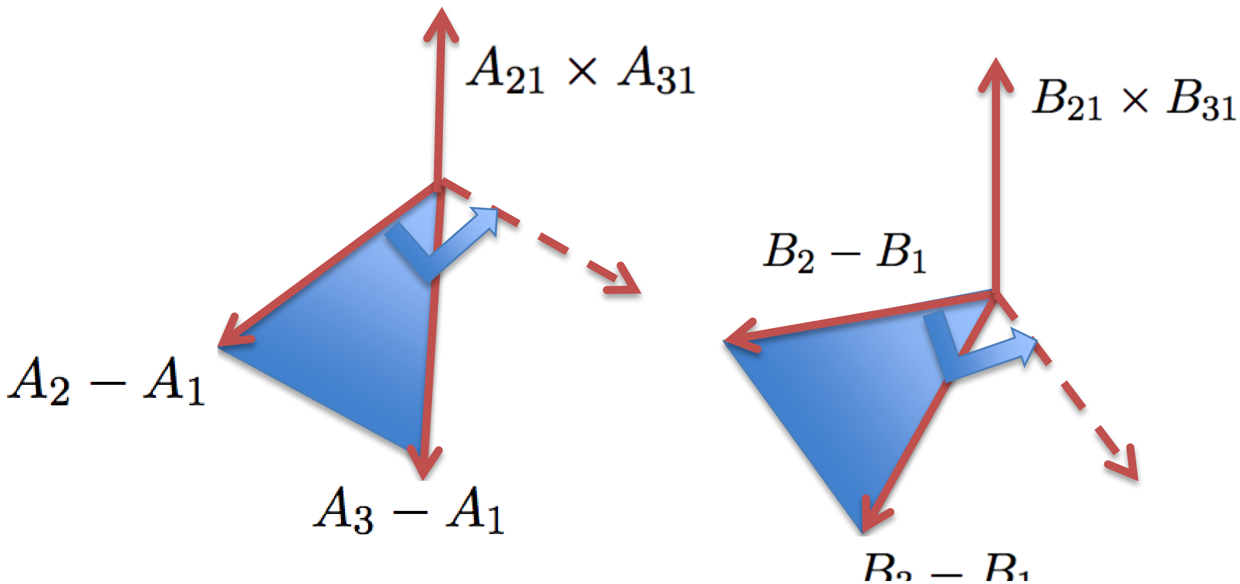


How do we solve for R, T from n point correspondences?

$$A_i = RB_i + T$$

What is the minimal number of points needed?





Three non-collinear points suffice: each triangle $A_{i=1\ldots3}$ and $B_{i=1\ldots3}$ make an orthogonal basis

$$\begin{pmatrix} A_{21} & (A_{21} \times A_{31}) \times A_{21} & A_{21} \times A_{31} \end{pmatrix}$$

and

$$\begin{pmatrix} B_{21} & (B_{21} \times B_{31}) \times B_{21} & B_{21} \times B_{31} \end{pmatrix}$$

Rotation between two orthogonal bases is unique.

We solve a minimization problem for $N > 3$ point correspondences:

$$\min_{R,T} \sum_{i}^{N} \|A_i - RB_i + T\|^2$$

After differentiating with respect to T we observe that the translation is the difference between the centroids:

$$T = \frac{1}{N}\sum_{i}^{N} A_i - R\frac{1}{N}\sum_{i}^{N} B_i = \bar{A} - R\bar{B}$$

We solve a minimization problem for $N > 3$ point correspondences:

$$\min_{R,T} \sum_{i}^{N} \|A_i - RB_i + T\|^2$$

After differentiating with respect to T we observe that the translation is the difference between the centroids:

$$T = \frac{1}{N}\sum_{i}^{N} A_i - R\frac{1}{N}\sum_{i}^{N} B_i = \bar{A} - R\bar{B}$$

We subtract the centroids $\bar{A}$ and $\bar{B}$ and rewrite the objective function as

$$\min_R \|A - RB\|_F^2$$

where

$$A = \begin{pmatrix} A_1 - \bar{A} & \dots & A_N - \bar{A} \end{pmatrix}$$

and

$$B = \begin{pmatrix} B_1 - \bar{B} & \dots & B_N - \bar{B} \end{pmatrix}$$

We rewrite the Frobenius norm using the trace of the matrix

$$\|A - RB\|_F^2 = tr(AA^T) + tr(BB^T) - tr(RBA^T) - tr(AB^T R^T)$$

and observe that only the two last terms depend on the unknown R yielding a maximization problem.

We rewrite the Frobenius norm using the trace of the matrix

$$\|A - RB\|_F^2 = tr(A^T A) + tr(B^T B) - tr(A^T RB) - tr(B^T R^T A)$$

and observe that only the two last terms depend on the unknown R yielding a maximization problem.

Even without using the properties of the trace we can see that both last terms are equal to

$$\sum_i^N R(B_i - \bar{B})(A_i - \bar{A})^T = tr(RBA^T)$$

The 3D-3D pose problem reduced to

$$\max_R tr(RBA^T)$$

If the SVD of BA^T is USV^T and $Z = V^T RU$

$$tr(RBA^T) = tr(RUSV^T) = tr(ZS) = \sum_1^3 z_{ii}\sigma_i \leq \sum_1^3 \sigma_i$$

and, hence, the upper bound is obtained by setting $Z = V^T RU = I$

$$R = VU^T$$

If the SVD of BA^T is USV^T and $Z = V^T RU$

$$tr(RBA^T) = tr(RUSV^T) = tr(ZS) = \sum_1^3 z_{ii}\sigma_i \leq \sum_1^3 \sigma_i$$

and, hence, the upper bound is obtained by setting $Z = V^T RU = I$

$$R = VU^T$$

To guarantee that it has determinant 1

$$R = V \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \det(UV^T) \end{pmatrix} U^T$$

3D-3D Registration enables the creation of 3D models from multiple point clouds: