# 3D Photography: Point Based Rigid Registration

## 1 Three Point Registration

Registration is the task of finding a transformation from one coordinate system to another so that all features that appear in both data sets are aligned with each other. We limit ourselves to rigid transformations with input given as 3D points. This means we need to find 3 translational parameters and 3 rotational parameters.

**Problem Definition**

Given the coordinates of points measured in two cartesian coordinate systems, find the rigid transformation between the two systems. Once we have this transformation, we can transform points in on coordinate system to the other.

Given the coordinates of three points in two coordinate systems, let us call them "left" and "right". The points are $p_{r1}, p_{r2}, p_{r3}$. We find the transformation between coordinate systems with the following construction:

1. Choose one of the points to be the origin, let us say $p_1$.

2. Construct the x axis:

$$x = \frac{p_2 - p_1}{|p2 - p1|}$$

3. Construct the y axis:

$$\begin{aligned} y &= (p_3 - p_1) - [(p_3 - p_1) \cdot x]x \\ y &= \frac{y}{|y|} \end{aligned}$$

4. Construct the z axis:

$$z = x \times y$$

5. Build the rotation matrices for both point sets:

$$R_l = [x_l, y_l, z_l] \quad R_r = [x_r, y_r, z_r]$$

6. We can find the rotation matrix $R$ that takes the "right" point set into the "left" point set:

$$R\,R_r \;=\; R_l$$

7. Solving for R, the rotation between the "right" coordinate system to the "left" is given by:

$$R = R_l R_r^T$$

8. The translation between the "right" coordinate system to the "left" is given by:

$$t = p_{l1} - Rp_{r1}$$

   This assumes we have a "match" of the 3 points in each data set with the other. Further, it is an exact method - there can be no imprecision with respect to noisy data.

# 2 Example

Figure 1 shows a statue being scanned from 2 different positions. Each scan position has scanned the same 3 points in its internal scan coordinates. We need to find the transformation to register the two point sets. In this example we wil find the registration that registers the "right" points with the "left" points. The points in their respective left and right scanner coordinates are below:

$$P_{l1} = (0, 2, 2);\;\; P_{l2} = (0, 4, 2);\;\; P_{l3} = (0, 2, 4)$$

$$P_{r1} = (0, 5, 0);\;\; P_{r2} = (2, 5, 0);\;\; P_{r3} = (0, 5, 2)$$

Using the method in section 1, we can form the local rotation matrix for each set of 3 corresponding points:

$$R_l = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \;\; R_r = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix};$$

The rotation matrix R which takes the "right" coordinate system into the "left" is defined as:

$$R\,R_r \;=\; R_l; \quad R \;=\; R_l\,{R_r}^T$$

and substituting for $R_l$ and ${R_r}^T$ we get:

$$R \;=\; \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix} \;=\; \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix};$$

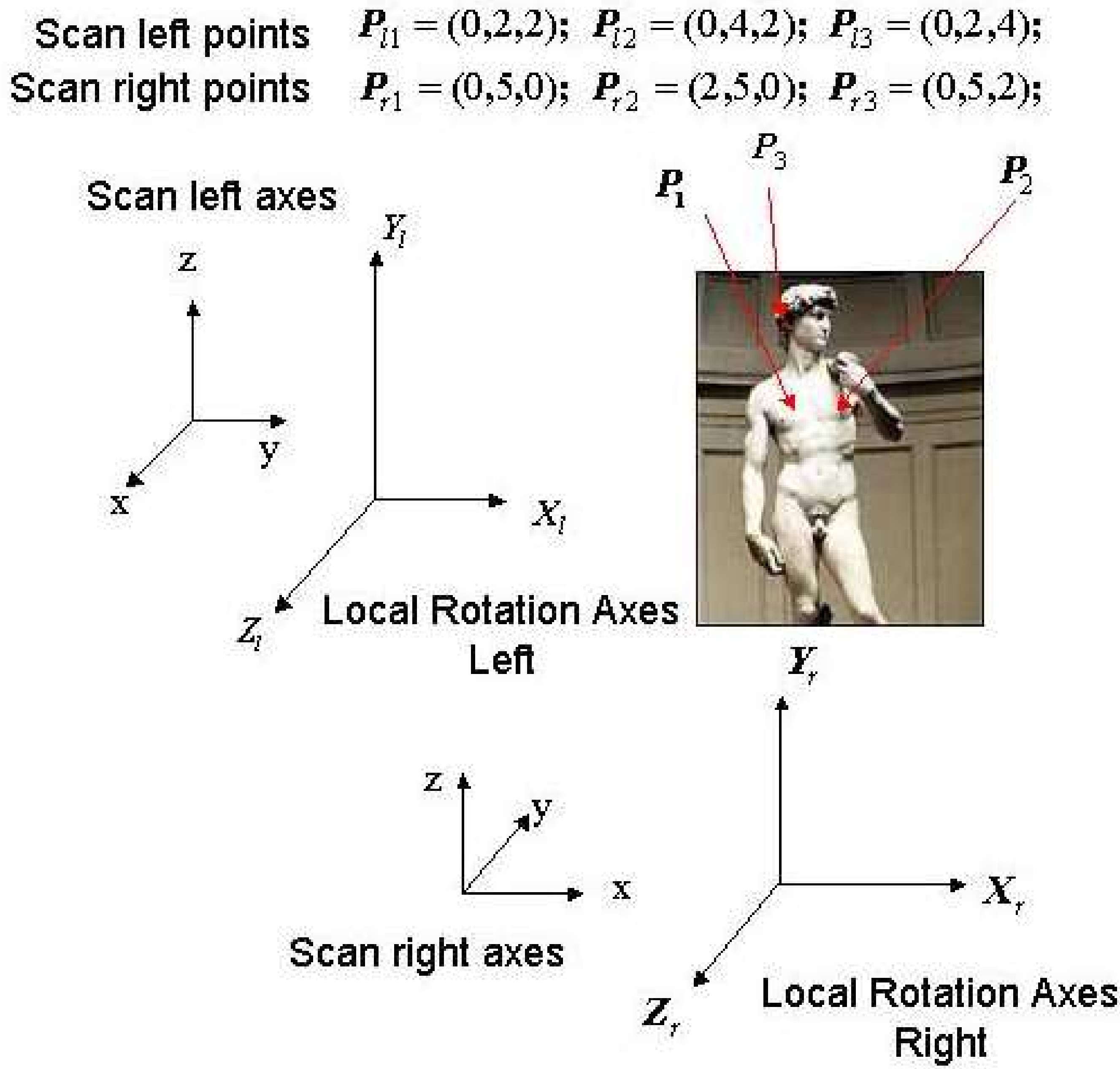


Figure 1: Scan registration example. The statue is scanned from 2 positions (left and right). Each scan has its own scan coordinates, which are then used to create local coordinate systems for each scan. The key is that the local systems are formed using the same 3 points, scanned from different positions. CAVEAT: the 3 points cannot be co-linear.

To get the translation vector, we can compute any of the below:

$$t \;=\; p_{l1} - R\,p_{r1}$$

$$t \;=\; p_{l2} - R\,p_{r2}$$

$$t \;=\; p_{l3} - R\,p_{r3}$$

which will all compute the same translation vector, which in this case is:

$$t \;=\; [5, 2, 2]$$

We can verify this is correct by transforming $P_{r1}$ into $P_{l1}$:

$$P_{l1} = R\,p_{r1} + t \;=\; \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} [0, 5, 0]^T + [5, 2, 2]^T \;=\; [0, 2, 2]^T$$

Each of the points in the "right" data set will transform the same way.

## 3 Direct Method to Compute R and T

Method due to K. Arun, et. al., IEEE PAMI, Vol 9, no 5, pp 698-700, Sept 1987

If we have a noisy point set, we can still calculate the transform using a direct method. Given 2 point sets we can find the best rotation and translation (in a Least Squares sense) that transforms the 2 point sets. This works if the 2 point sets aren't an exact match (which is the usual case). For this method, you need to have at least 6 points. The method starts by decoupling the rotation from the translation, and solving for the rotation.

The error we want to minimize is the Euclidean distance between each point and its transformed counterpart. If we have point correspondences in the 2 data sets $(p_i, p_i')$ then the error is:

$$E = \sum_{i=1}^{n} \|p_i' - (Rp_i + T\|^2 \tag{1}$$

To decouple the rotation, we simply take the centroid of each data set, and subtract it out from each data point. Let $p$ be the centroid of data set $p_i$ and $p'$ be the centroid of data set $1p_i'$. Then the new translated coordinates of each of the 2 data sets is:

$$q_i = p_i - p; \;\; q_i' = p_i' - p' \tag{2}$$

Now that we have each data set "centered" at its centroid, we need to find the rotation that registers the points. we can rewrite the error equation (1), leaving out the translation as:

$$E = \sum_{i=1}^{n} \|q_i' - Rq_i\|^2 \tag{3}$$

Once we find this rotation matrix $R$, we can then use the centroids $(p', p)$ to find the translation that minimizes the registration error as:

$$T = p' - Rp \tag{4}$$

To find rotaion matrix $R$, we can rewrite the error (3) as:

$$E = \sum_{i=1}^{n} (q_i' - Rq_i)^T (q_i' - Rq_i) \tag{5}$$

Expanding this we get:

$$E = \sum_{i=1}^{n} (q_i'^T q_i' + q_i^T q_i - 2q_i'^T Rq_i) \tag{6}$$

To minimize this error, we can see that it implies *maximizing* the quantity:

$$E' = \sum_{i=1}^{n} q_i'^T R q_i \tag{7}$$

$$= Trace\,(R \sum_{i=1}^{n} q_i q_i'^T) \tag{8}$$

$$= Trace\,(RH) \tag{9}$$

where $H = \sum_{i=1}^{n} q_i q_i'^T$. H is known as the covariance matrix. We are computing the covariance matrix of the two normalized data sets. This matrix has as its eigenvector the rotation axis which maps the two point sets into each other. By using a Singular Value Decomposition (SVD) of the matrix, you can directly form the transformation matrix. The matrix H can be decomposed using a singular value decomposition (SVD):

$$H = U \Lambda V^T$$

Without going into details or proofs, it can be shown that $Trace\,(RH)$ can be maximized when

$$R = VU^T \tag{10}$$

This solves for the rotation, and the translation can be found using equation(4).
To recap:

1. Find the centroid of each data set, and subtract the centroid from each data point - centers the data set at a new origin, the centroid. We can think of these data sets as the normalized data sets.

2. Once the data sets are both centered at their centroids, we need to find the rotation about these common origins that will bring them into registration.

3. Create the covariance matrix $H$ of the normalized data sets.

4. Use the SVD decomposition of $H$ to find rotation matrix $R$ and you are done. You can find software to take the SVD in many places, including the book "Numerical Recipes" (We will provide you with a C code Numerical Recipes library if you need it).

$$H = \sum_i \begin{bmatrix} q'_{i,x} q_{i,x} & q'_{i,x} q_{i,y} & q'_{i,x} q_{i,z} \\ q'_{i,y} q_{i,x} & q'_{i,y} q_{i,y} & q'_{i,y} q_{i,z} \\ q'_{i,z} q_{i,x} & q'_{i,z} q_{i,y} & q'_{i,z} q_{i,z} \end{bmatrix}$$

Note: Verify $Det(R) = 1$. If not, then algorithm may fail.