The Analytic 3-D Transform for the Least-Squared Fit of Three Pairs of Corresponding Points

David M. Weinstein
Email: dweinste@cs.utah.edu

UUCS-98-005

Department of Computer Science
University of Utah
Salt Lake City, UT 84112 USA

March 1, 1998

Abstract

We derive the analytic transformation for minimizing the summed-squared-distance between three movable points in one three-space pose to three corresponding fixed points in another three-space pose. This change of basis is a general rigid-body transformation (translation and rotation), with the addition of a uniform scale. We also derive and present the root-mean-squared distance between the final transformed points and the fixed points.

1 Introduction

The coregistration problem—taking spatial data from one coordinate frame and transforming it to correspond with spatial data from another coordinate frame—is fundamental in constructing computational models from multi-modal data. In particular, when a researcher wishes to combine data which has been acquired using two different systems, or even the same system but using different parameters or at different times, at least one data set must be "transformed" to achieve spatial correlation.

In general, data has an intrinsic coordinate frame, and therefore data retrieved from such a system on different occasions is already in the same space by default. However, if the object being scanned has moved *within* that space, a coregistration process is required to move the poses into alignment.

In the vision literature, this problem is referred to as the "pose recovery" problem [1]. The notion being that information about an object is obtained in a 2-D (image) or 3-D (laser range data) scan. If known points can be identified from that data (*i.e.,* a correspondence can be determined) then the pose recovery problem seeks to determine the pose of the scanned object from those points.

There are general purpose, iterative, algorithms for solving the least-squares coregistration problem for an arbitrary number of points [2, 3]. Here, we focus on an analytic solution to the specific problem of three corresponding points in two poses, and an analytic measurement of the root-mean-squared (RMS) distance/error between the points of the resultant poses.

2 Methods

We will begin with the intuition behind our method, with a high-level discussion of why the method works. Then we will derive the mathematics of the solution and prove that our intuition is correct. We conclude by computing the final RMS distance between the vertices of the transformed and fixed triangles.

2.1 Intuitive Derivation

Since we have three points in each pose, from this point forward we will refer to the points as triangle vertices. Specifically the points from the pose to be transformed will be triangle $a : (a1, a2, a3)$, and the points from the fixed pose will be triangle

$p : (p1, p2, p3)$. Each triangle contains the three nodes from the pose, with the nodes ordered such that correspondence is maintained (*i.e.,* the coregistration transformation will take $a1 \to p1, a2 \to p2$, and $a3 \to p3$). The triangle a has centroid C_a and normal $\vec{N}_a$, and the triangle p has centroid C_p and normal $\vec{N}_p$.

We begin the transformation by translating a so its centroid, C_a, aligns with the centroid of p, C_p. Next, we find the rotation matrix which aligns the normal $\vec{N}_a$ to $\vec{N}_p$. After these transformations have been performed, the triangles are located in the same plane, and have the same centroid. Since pC and aC are now the same point, we will simply refer to them as C; similarly the normal of both triangles will be referred to as $\vec{N}$.

Next, we will rotate the new triangle a about $\vec{N}$ (with fixed point C), in order to minimize the summed-squared-distance between the vertices. From here on out, we will refer to this summed-squared-distance between the corresponding vertices, (that is, the term this algorithm is devised to minimize,) as the *distance between the triangles.*

Finally, we determine a scale factor, s, for the triangle a, and scale the distance between its vertices $(a1, a2, a3)$ and the centroid C in order to once again minimize the distance between the triangles. We refer to the resultant transformed triangle as $\ddot{a}$.

Compositing all of these transformations, we obtain the final transformation matrix for coregistering the two poses. We are also able to obtain an analytic expression for the remaining distance between the poses. A rigorous mathematical derivation is presented below.

2.2 Mathematical Derivation

Mathematically, we have seven degrees of freedom we are solving for in this transformation. Two degrees describe the normal to the transformed triangle $\ddot{a}$, three describe the location of the centroid $C_{\ddot{a}}$, one describes the rotation about the normal $\vec{N}_{\ddot{a}}$ which brings the poses into the best alignment, and a final scale factor optimizes over the remaining space of similar triangles.

The complete transformation is a function of the degrees of freedom expressed above. Those constraints can be expressed as matrix operations in homogeneous coordinates, where applying each operation brings the triangles closer to alignment. The product of all the matrix operations is the complete transform. The optimal choice for the various operations (translation, change of basis, rotation and scale) are independent.

The mathematical justification for breaking up the transformation as done here, is based on the general formula for the transformation as a function of the seven degrees of freedom. Given two variables that correspond to degrees of freedom in two different stages of the transformation, the second mixed partial with respect to these variables is always zero. However, because matrix multiplications do not always commute, the order of the operations is important. For example, before choosing an optimal rotation, it greatly simplifies the problem if both triangles have been centered about the origin, and are located in the same plane.

The matrix operations concatenate as follows:

$$\ddot{a}_i = \mathbf{T_{C_p}}\,\mathbf{B_P}^t\,\mathbf{S}\,\mathbf{\Theta}\,\mathbf{B_A}\,\mathbf{T_{C_{-a}}}\,a_i$$

Reading from right to left, these matrices operate on a to:

1. $\mathbf{T_{C_{-a}}}$: Translate the vertices to be centered about the origin;
2. $\mathbf{B_A}$: Rotate the vertices into the xy-plane;
3. $\mathbf{\Theta}$: Rotate the vertices within the xy-plane;
4. $\mathbf{S}$: Scale the distance from the vertices to the origin;
5. $\mathbf{B_P}^t$: Rotate the vertices out of the xy-plane, into the coordinate frame of of p;
6. $\mathbf{T_{C_p}}$: Translate the vertices to have the same centroid as p.

Each of the above operations is described below.

2.2.1 Translation

First, we translate our coordinate frames to line up the centroids of the triangles. In Appendix A, we prove that the centroids must always be aligned for two triangles to have a minimal distance.

$$TC = C_p - C_a$$

This translation vector can be stored as a matrix $\mathbf{T_C}$,

$$\mathbf{T_C} = \begin{bmatrix} 1 & 0 & 0 & TC_x \\ 0 & 1 & 0 & TC_y \\ 0 & 0 & 1 & TC_z \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

or as the partial matrices $\mathbf{T_{C_p}}$, $\mathbf{T_{C_{-p}}}$, and $\mathbf{T_{C_{-a}}}$,

$$\mathbf{T_{C_p}} = \begin{bmatrix} 1 & 0 & 0 & C_{p_x} \\ 0 & 1 & 0 & C_{p_y} \\ 0 & 0 & 1 & C_{p_z} \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad (1)$$

$$\mathbf{T_{C_{-p}}} = \begin{bmatrix} 1 & 0 & 0 & -C_{p_x} \\ 0 & 1 & 0 & -C_{p_y} \\ 0 & 0 & 1 & -C_{p_z} \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

$$\mathbf{T_{C_{-a}}} = \begin{bmatrix} 1 & 0 & 0 & -C_{a_x} \\ 0 & 1 & 0 & -C_{a_y} \\ 0 & 0 & 1 & -C_{a_z} \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad (2)$$

where $\mathbf{T_C} = \mathbf{T_{C_p}} + \mathbf{T_{C_{-a}}}$.

We apply $\mathbf{T_{C_{-a}}}$ to the the triangle a, thus locating its centroid at the origin, resulting in the new triangle $\dot{a}$. Since we want the centroids to be aligned (for optimizing future transforms), we apply $\mathbf{T_{C_{-p}}}$ to p and generate a new fixed triangle, $\dot{p}$ which has its centroid at the origin, as well. In actuality, p will never be translated to the origin; rather, a will eventually be translated *from the origin to* p. However, for the sake of derivation and conceptual clarity, we describe p as being temporarily translated.

2.2.2 Change of Basis

Next, we find the normal of $\dot{a}$. As shown in Appendix B, the two triangles must have the same normal to minimize the distance between $\dot{a}$ and $\dot{p}$. So, the optimal transform will take $\vec{N_{\dot{a}}}$ to $\vec{N_{\dot{p}}}$. To derive this rotation, we first determine the coordinate frames of the two poses. For each of these we find three orthonormal vectors which span the space. Without loss of generality, we describe $\dot{p}$:

$$\begin{aligned} \vec{N_{\dot{p}}} &= \overline{(\dot{p1} - \dot{p3})} \times \overline{(\dot{p1} - \dot{p2})} \\ \vec{U_{\dot{p}}} &= \overline{(\dot{p3} - \dot{p1})} \\ \vec{V_{\dot{p}}} &= \vec{N_{\dot{p}}} \times \vec{U_{\dot{p}}} \end{aligned}$$

The bases $(\vec{U_{\dot{p}}}, \vec{V_{\dot{p}}}, \vec{N_{\dot{p}}})$, and $(\vec{U_{\dot{a}}}, \vec{V_{\dot{a}}}, \vec{N_{\dot{a}}})$ both span $\mathbb{R}^3$. We call the spaces spanned by these vectors $\mathcal{P}$ and $\mathcal{A}$, and define $\mathbf{B_P}$ and $\mathbf{B_A}$ to be the matrices with rows composed of these basis vectors. $\mathbf{B_A}$ defines the change of basis from $\mathcal{A} \to \mathbb{R}^3$, and $\mathbf{B_P}$ defines the change of basis from $\mathcal{P} \to \mathbb{R}^3$. Similarly, to go from $\mathbb{R}^3 \to \mathcal{P}$, we need $\mathbf{B_P}^{-1}$;

however, since $\mathbf{B_P}$ is a rotation matrix, the inverse is simply the transpose, that is $\mathbf{B_P}^{-1} = \mathbf{B_P}^t$.

$$\mathbf{B_A} = \begin{bmatrix} U_{\dot{a}_x} & U_{\dot{a}_y} & U_{\dot{a}_z} & 0 \\ V_{\dot{a}_x} & V_{\dot{a}_y} & V_{\dot{a}_z} & 0 \\ N_{\dot{a}_x} & N_{\dot{a}_y} & N_{\dot{a}_z} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad (3)$$

$$\mathbf{B_P} = \begin{bmatrix} U_{\dot{p}_x} & U_{\dot{p}_y} & U_{\dot{p}_z} & 0 \\ V_{\dot{p}_x} & V_{\dot{p}_y} & V_{\dot{p}_z} & 0 \\ N_{\dot{p}_x} & N_{\dot{p}_y} & N_{\dot{p}_z} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

$$\mathbf{B_P}^t = \begin{bmatrix} U_{\dot{p}_x} & V_{\dot{p}_x} & N_{\dot{p}_x} & 0 \\ U_{\dot{p}_y} & V_{\dot{p}_y} & N_{\dot{p}_y} & 0 \\ U_{\dot{p}_z} & V_{\dot{p}_z} & N_{\dot{p}_z} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad (4)$$

$$(5)$$

Therefore, if we want to transform a point from $\mathcal{A} \rightarrow \mathcal{P}$, we simply transform it through the matrix: $\mathbf{B_{PA}} = \mathbf{B_P}^t \mathbf{B_A}$.

Transformations $\mathbf{B_A}$ and $\mathbf{B_P}$ transform $\dot{a}$ and $\dot{p}$ into triangles in the xy-plane (still with centroids at the origin). We refer to these new triangles as $\hat{a}$ and $\hat{p}$, respectively.

2.2.3 Rotation

At this point, we have determined five of our seven degrees of freedom. The remaining two degrees represent the optimal rotation of the vertices of $\hat{a}$ about the z-axis, and their scaled distances from the origin. Both parameters will be optimized to minimize the distance between $\hat{a}$ and $\hat{p}$.

The distance between the triangles can be written as a function of a rotation θ about the z-axis, and can be computed as the sum of the squared distances between the corresponding points. These individual squared distances are also functions of θ. First, we define two distances:

$$\widehat{ra_i} = \|\vec{\hat{a}_i}\|^2,$$
$$\widehat{rp_i} = \|\vec{\hat{p}_i}\|^2,$$

Now we can define θ_i as the angular distance from $\vec{\hat{a}_i}$ to $\vec{\hat{p}_i}$ about the z-axis. The angular from normalized vector $\vec{U1}$ to normalized vector $\vec{U2}$, rotating about their cross-product, $\vec{N} = (\vec{U1} \times \vec{U2})$, is given by:

$$\arcsin(sign(\vec{U1} \times (\vec{U2} \times \vec{N})) \times \sqrt{(1 - (\vec{U1} \cdot \vec{U2}))^2})$$

We define f to be the distance between the triangles as a function of θ:

$$f(\theta) = \sum_{i=1}^{3} (\widehat{ra_i}^2 + \widehat{rp_i}^2 - 2\widehat{ra_i}\widehat{rp_i}\cos(\theta + \theta_i))$$

Finding the minimum of $f(\theta)$ by setting its derivative equal to zero and solving for θ, we find:

$$\theta = -\theta_0 + \arctan\left(\frac{\widehat{ra_1}\widehat{rp_1}\sin(\theta_0 - \theta_1) + \widehat{ra_2}\widehat{rp_2}\sin(\theta_0 - \theta_2)}{\widehat{ra_0}\widehat{rp_0} + \widehat{ra_1}\widehat{rp_1}\cos(\theta_0 - \theta_1) + \widehat{ra_2}\widehat{rp_2}\cos(\theta_0 - \theta_2)}\right)$$

We construct the rotation matrix $\mathbf{\Theta}$ to rotate about the z-axis (anchored at the origin) by θ, and compute the new vertices $\tilde{a}_i$, and new vectors $\vec{\tilde{a}}_i$.

$$\mathbf{\Theta} = \begin{bmatrix} cos(\theta) & -sin(\theta) & 0 & 0 \\ sin(\theta) & cos(\theta) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{6}$$

2.2.4 Scale

Finally, we want to find the ideal scale factor. Using the new vectors $\vec{\tilde{a}}_i$ and the $\vec{p_i}$ vectors computed above, distance as a function of scale is given by:

$$d(s) = \sum_{i=1}^{3} \|\vec{p_i} - s\vec{\tilde{a}}_i\|^2$$

Once again, we solve for the s which minimizes this distance by setting the derivate equal to zero. Doing so, we find:

$$s = \frac{\sum_{i=1}^{3} \vec{\tilde{a}}_i \cdot \vec{\tilde{p}}_i}{\sum_{i=1}^{3} \|\vec{\tilde{a}}_i\|^2}$$

The scale factor matrix, $\mathbf{S}$ can be expressed as:

$$\mathbf{S} = \begin{bmatrix} s & 0 & 0 & 0 \\ 0 & s & 0 & 0 \\ 0 & 0 & s & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{7}$$

2.2.5 Full Transform

Having derived all of our component matrices (1), (2), (3), (4), (6), (7), we can now express our full transform as a product of these components. We note that the rotate and scale matrices required our points to be centered about the origin, with normal along the z-axis. Additionally, our change of bases required that our triangle also be centered about the origin. We accomplish this by rearranging our transforms (this is equivalent to how we have been discussing the vertices all along):

$$\mathbf{TT} = \mathbf{T_{C_p}}\,\mathbf{B_P}^{\mathbf{t}}\,\mathbf{S}\,\mathbf{\Theta}\,\mathbf{B_A}\,\mathbf{T_{C_{-a}}}$$

If we construct matrices **A** and **P** from our original vertices,

$$\mathbf{A} = \begin{bmatrix} a1_x & a2_x & a3_x & 0 \\ a1_y & a2_y & a3_y & 0 \\ a1_z & a2_z & a3_z & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix},$$

$$\mathbf{P} = \begin{bmatrix} p1_x & p2_x & p3_x & 0 \\ p1_y & p2_y & p3_y & 0 \\ p1_z & p2_z & p3_z & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix},$$

Then the new vertices $\ddot{a}_i$ form the columns of the resultant matrix $\ddot{\mathbf{A}}$:

$$\ddot{\mathbf{A}} = \mathbf{T_{C_p}}\,\mathbf{B_P}^{\mathbf{t}}\,\mathbf{S}\,\mathbf{\Theta}\,\mathbf{B_A}\,\mathbf{T_{C_{-a}}}\,\mathbf{A}$$

and the RMS distance between the vertices of the new triangle and the old triangle is:

$$\text{RMS distance} = \frac{\sqrt{\sum_{i=1}^{3}\sum_{j=0}^{2}(\ddot{\mathbf{A}}_{i,j} - \mathbf{P}_{i,j})^2}}{3}$$

3 Acknowledgments

This work was supported in part by the National Science Foundation. I would like to thank Gordon Kindlmann, Dean Brederson, and Peter-Pike Sloan for their valuable comments and suggestions.

4 Appendices

4.1 Appendix A

The summed-squared-distance between the vertices $(a1, a2, a3)$ of triangle a and the vertices $(p1, p2, p3)$ of triangle p, is:

$$\sum_{i=1}^{3}(pi_x - ai_x)^2 + (pi_y - ai_y)^2 + (pi_z - ai_z)^2$$

Similarly, if we allow the a vertices to be translated by an arbitrary vector t: (t_x, t_y, t_z), we find the distance to now be:

$$\sum_{i=1}^{3}(pi_x - (ai_x + t_x)^2) + (pi_y - (ai_y + t_y))^2 + (pi_z - (ai_z + t_z))^2$$

Expressing this distance as a function of t and setting the partial derivatives equal to zero, we find optimal values for tx, ty, and tz:

$$tx = \frac{-(p1_x + p2_x + p3_x) + (a1_x + a2_x + a3_x)}{3},$$
$$ty = \frac{-(p1_y + p2_y + p3_y) + (a1_y + a2_y + a3_y)}{3},$$
$$tz = \frac{-(p1_z + p2_z + p3_z) + (a1_z + a2_z + a3_z)}{3}.$$

That is, the ideal translation is exactly that which brings the centroids into alignment.

4.2 Appendix B

Here we prove that two triangles with centers at the origin must have the same normal to minimize the summed-squared-distance between their vertices. Without loss of generality, we consider the case where the first triangle, p, is located in the xz-plane, and the second triangle, a was originally also located in the xz-place, but is now allowed to rotate out of the xz-plane, about the z-axis by an amount θ. We show that the optimal value of θ is zero; thus the summed-squared-distance between corresponding points is minimized when both triangles have the same normal.

Exploiting the facts that the triangle is centered about the origin (thus, $p3 = -p1 - p2$), the rotation is about the z-axis (thus there's no change in z-values), and the y-values were all originally zero (because the points were all originally in the xz-plane), we derive the summed-squared-distance to be:

$$\begin{aligned} d(\theta) \;=\; & (p1_x - a1_x * \cos(\theta))^2 + (a1_y * \sin(\theta))^2 + (p1_z - a1_z)^2 + \\ & (p2_x - a2_x * \cos(\theta))^2 + (a2_y * \sin(\theta))^2 + (p2_z - a2_z)^2 + \\ & ((-p1_x - p2_x) - (-a1_x - a2_x) * \cos(\theta))^2 + ((-a1_x - a2_x) * \sin(\theta))^2 + \\ & ((-p1_z - p2_z) - (-a1_z - a2_z))^2. \end{aligned}$$

Taking the derivative and setting it equal to zero, we find four critical values (*i.e.,* local minima and maxima). Taking the derivative a second time, and plugging in those critical values, we find that only $\theta = 0$ corresponds to a local minimum; that is, the triangles are closest when they have exactly the same normal.

References

[1] O.D. Faugeras and M. Hébert. The representation, recognition, and locating of 3D shapes from range data. *International Journal of Robotics Research*, 5(3):27–52, 1986.

[2] J.H. Challis. A procedure for determining rigid body transformation parameters. *J. Biomechanics*, 28(6):733–737, 1995.

[3] G.H. Golub and C.F. Van Loan. *Matrix Computations*. Johns Hopkins, Baltimore, 1989.