Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics
San Antonio, TX, USA - October 2009

On the Absolute Orientation Problem in Computer Vision

Rodney G. Roberts
Dept. of Electrical and Computer Eng.
Florida State University
Tallahassee, FL, USA
rroberts@eng.fsu.edu

Daniel W. Repperger
Air Force Research Laboratory
711 HPW AFRL/RHCV
Wright-Patterson AFB, OH 45433
Daniel.Repperger@wpafb.af.mil

***Abstract*—An important problem in computer vision is to determine the orientation of a rigid body in an image. This can be accomplished by matching points or line segments that naturally appear on the object. Several elegant and computationally fast algorithms based on the singular value decomposition and quaternions have been introduced to solve this problem. In this article, the authors first examine the important special case of identifying the attitude of 2D objects and introduce a particularly elegant solution based on the mathematical structure of the complex plane. Motivated by this simple solution to the 2D case, a new derivation of the 3D case based on the polar decomposition is presented. This derivation is in many ways more natural than previous derivations, particularly when the model and data contain no noise.**

***Index Terms*—Absolute orientation, least squares, polar decomposition.**

I. INTRODUCTION

A fundamental problem in computer vision is the determination of the orientation of a rigid object. An effective approach to this problem is to match a set of points on the object with the corresponding points on a model. In particular, the following mathematical problem appears in a number of references [1]-[6].

Two point sets $\{\mathbf{a}_i\}$ and $\{\mathbf{b}_i\}$ of N vectors in the plane or in 3-space are related by

$$\mathbf{b}_i = R\mathbf{a}_i + \mathbf{t} + \mathbf{n}_i \tag{1}$$

where R is a rotation matrix, $\mathbf{t}$ is a translation vector, and $\mathbf{n}_i$ is a noise term. The set $\{\mathbf{a}_i\}$ corresponds to the location of several specified points on a model of the object while $\{\mathbf{b}_i\}$ represents the corresponding points on the object in an image. The goal is to determine R and $\mathbf{t}$ to minimize

$$F(\mathbf{t}, R) = \sum_{i=1}^{N} \|\mathbf{b}_i - (R\mathbf{a}_i + \mathbf{t})\|^2 \tag{2}$$

where $\|\cdot\|$ is the standard 2-norm. The vector $\mathbf{t}$ and rotation matrix R represent the location and orientation of the object in the image. The problem of matching line segments, although more complicated, results in essentially the same type of optimization problem [2].

Several approaches to this problem have been described in the literature including matrix-based solutions such as the singular value decomposition (SVD) [1]-[3] and quaternion-based solutions [4],[5]. In the next section, we introduce a new proof for the 2D case based on simple properties of complex numbers. In particular, it is shown that the solution for the orientation of the object is given by the polar form of a particular complex number. In Section III, this solution motivates a natural solution to the general 3D problem based on the polar decomposition of a particular matrix. In fact, the polar decomposition is the obvious solution when no noise is present in the problem. After solving for the case when there is noise in the data, we provide a new proof that the same solution holds when there is noise in both the model and the data. Lastly, conclusions appear in Section IV.

II. SOLVING THE PLANAR CASE USING THE COMPLEX PLANE

A popular solution to the orientation problem is based on the singular value decomposition (SVD). However, as one would expect, the much simpler 2D case does not require the sophistication of an SVD, not only because of the smaller dimension size, but more importantly, because of the commutativity of the rotation operation. In this case, it is convenient to formulate the problem in terms of complex numbers. Suppose that $\mathbf{a} = \begin{bmatrix} a_x & a_y \end{bmatrix}^T$ and $\mathbf{b} = \begin{bmatrix} b_x & b_y \end{bmatrix}^T$ are vectors in the plane. If we write these vectors in complex number notation as $a = a_x + ja_y$ and $b = b_x + jb_y$, then the inner product $\mathbf{a} \cdot \mathbf{b}$ of the two vectors is given in complex number notation as $\mathrm{Re}(a^*b)$ where a^* denotes the complex conjugate of the complex number a and where $\mathrm{Re}(z)$ denotes the real part of z. Furthermore, the norm squared $\|\mathbf{a}\|^2$ of the vector $\mathbf{a}$ is given by $|a|^2 = a^*a$, and the rotation matrix corresponding to a counterclockwise rotation of θ radians is given by the complex number $e^{j\theta}$. Based on this formulation, the goal is to minimize the objective function

$$F(t, e^{j\theta}) = \frac{1}{N} \sum_{i=1}^{N} |b_i - (e^{j\theta}a_i + t)|^2 \tag{3}$$

where the complex number t and the real number $\theta \in [0, 2\pi)$ represent the position and orientation of the object, respectively. Since determining the optimal $t_0(\theta)$ for a given θ is a routine least squares calculation, we merely state the result that

$$t_0(\theta) = \bar{b} - e^{j\theta}\bar{a} \tag{4}$$

where $\bar{b} = \frac{1}{N}\sum_{i=1}^{N} b_i$ and $\bar{a} = \frac{1}{N}\sum_{i=1}^{N} a_i$. We define $\tilde{a}_i = a_i - \bar{a}$ and $\tilde{b}_i = b_i - \bar{b}$ and say that $\{\tilde{a}_i\}$ and $\{\tilde{b}_i\}$ are the unbiased versions of $\{a_i\}$ and $\{b_i\}$, respectively. It then follows that

$$\begin{aligned} F(t, e^{j\theta}) &= \tfrac{1}{N}\textstyle\sum_{i=1}^{N} |\tilde{b}_i - e^{j\theta}\tilde{a}_i|^2 \\ &= \tfrac{1}{N}\textstyle\sum_{i=1}^{N} \left[|\tilde{b}_i|^2 + |\tilde{a}_i|^2 - 2\mathrm{Re}((\tilde{b}_i)^* e^{j\theta}\tilde{a}_i) \right]. \end{aligned} \tag{5}$$

Since $\frac{1}{N}\sum_{i=1}^{N}\left[|\tilde{b}_i|^2 + |\tilde{a}_i|^2\right]$ is fixed by the data, we want to maximize $\mathrm{Re}\left[\left(\frac{1}{N}\sum_{i=1}^{N}(\tilde{b}_i)^*\tilde{a}_i\right)e^{j\theta}\right]$ with respect to θ, which is clearly achieved by choosing

$$\theta = -\arg\left(\sum_{i=1}^{N}(\tilde{b}_i)^*\tilde{a}_i\right) = \arg\left(\sum_{i=1}^{N}\tilde{a}_i^*\tilde{b}_i\right) \tag{6}$$

where $\arg(z)$ denotes the argument of the complex number z and has range $[0, 2\pi)$. The location and orientation of the object is then given by (4) and (6), respectively.

Another approach, that will serve as a guide in the next section to solve the 3D case, is to write

$$\begin{aligned} F(t, e^{j\theta}) &= \tfrac{1}{N}\textstyle\sum_{i=1}^{N}\left[|\tilde{b}_i|^2 + |\tilde{a}_i|^2\right] - 2\mathrm{Re}(c^* e^{j\theta}) \\ &= \tfrac{1}{N}\textstyle\sum_{i=1}^{N}\left[|\tilde{b}_i|^2 + |\tilde{a}_i|^2\right] - |c|^2 - 1 + |c - e^{j\theta}|^2 \end{aligned} \tag{7}$$

where $c = \frac{1}{N}\sum_{i=1}^{N}\tilde{a}_i^*\tilde{b}_i$. We thus want to minimize $|c - e^{j\theta}|^2$, which is clearly achieved by $\theta = \arg(c)$ or, equivalently, $\theta = \arg\left(\sum_{i=1}^{N}\tilde{a}_i^*\tilde{b}_i\right)$.

III. SOLVING THE GENERAL CASE USING THE POLAR DECOMPOSITION

A popular approach to solving the 3D case is based on quaternions [4], [5]. While quaternions are a generalization of complex numbers, the complex number approach of the previous section more naturally leads to a matrix solution based on the polar decomposition. We begin by reformulating the problem in matrix notation by letting $A = \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_N \end{bmatrix}$, $B = \begin{bmatrix} \mathbf{b}_1 & \cdots & \mathbf{b}_N \end{bmatrix}$, and $N = \begin{bmatrix} \mathbf{n}_1 & \cdots & \mathbf{n}_N \end{bmatrix}$. Equation (1) then becomes

$$B = RA + \mathbf{t}\mathbf{e}^T + N \tag{8}$$

where $\mathbf{e} = \begin{bmatrix} 1 & \cdots & 1 \end{bmatrix}^T$, and the optimization problem becomes the minimization of

$$F(\mathbf{t}, R) = \|B - (RA + \mathbf{t}\mathbf{e}^T)\|_F^2 \tag{9}$$

subject to R being a rotation matrix where $\|\cdot\|_F$ denotes the Frobenius norm, which is given by the square root of the sum of the squares of the matrix elements.

A. The Noise-free Case

We first examine the simplest possible case, i.e., when no noise is present. In this ideal case, we have an exact equality, which can be written in matrix form as

$$RA + \mathbf{t}\mathbf{e}^T = B \tag{10}$$

where $\mathbf{e} = \begin{bmatrix} 1 & \cdots & 1 \end{bmatrix}^T$. The translation term $\mathbf{t}$ can be found by post-multiplying (10) by $\frac{1}{N}\mathbf{e}$ to obtain

$$\mathbf{t} = \bar{\mathbf{b}} - R\bar{\mathbf{a}} \tag{11}$$

where $\bar{\mathbf{a}} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{a}_i$ and $\bar{\mathbf{b}} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{b}_i$. Writing $\tilde{A} = A - \bar{\mathbf{a}}\mathbf{e}^T$ and $\tilde{B} = B - \bar{\mathbf{b}}\mathbf{e}^T$, we have $R\tilde{A} = \tilde{B}$ or, equivalently, $\tilde{A}^T R^T = \tilde{B}^T$. Lastly, pre-multiplying both sides by $\tilde{A}$ gives

$$\tilde{A}\tilde{A}^T R^T = \tilde{A}\tilde{B}^T. \tag{12}$$

We assume that the 3×3 matrix $\tilde{A}\tilde{B}^T$ has full rank. Then the unique polar decomposition of $\tilde{A}\tilde{B}^T = PQ$ is given by the left hand side of (12) where $P = \tilde{A}\tilde{A}^T$ is positive definite and $Q = R^T$ is an orthogonal matrix. Since $P = \tilde{A}\tilde{A}^T$ is positive definite, it follows that the determinants of $\tilde{A}\tilde{B}^T$ and Q have the same sign so that $R = Q^T$ is a rotation matrix if and only if $\det(\tilde{A}\tilde{B}^T) > 0$. If $\det(\tilde{A}\tilde{B}^T) < 0$, then R is a reflection matrix and the orientation problem is ill-defined.

The polar decomposition is a natural solution to the problem when no noise is present. Before continuing to the more general case, we observe that the solution is particularly simple when $\tilde{A}$ has the property that $\tilde{A}\tilde{A}^T = kI$. In that case, we merely scale $\tilde{A}\tilde{B}^T$ to obtain an orthogonal matrix. An example of this occurs when the columns of A correspond to the vertices of a platonic solid. Unfortunately, choosing a data matrix so that $\tilde{A}$ has this property may result in a numerically unstable solution when noise is present.

B. Noise in Only the Data

We now return to the problem formulated in (9), which can be rewritten as

$$F(\mathbf{t}, R) = \|\mathbf{t}\mathbf{e}^T - (B - RA)\|_F^2. \tag{13}$$

The optimal solution for $\mathbf{t}$ for a fixed rotation matrix R is given by the pseudoinverse solution

$$\mathbf{t} = (B - RA)(\mathbf{e}^T)^+ = \frac{1}{N}(B - RA)\mathbf{e} \tag{14}$$

so that it follows that $\mathbf{t}$ is once again given by (11). Substituting this expression into (13) gives

$$\begin{aligned} \|R\tilde{A} - \tilde{B}\|_F^2 &= \|\tilde{A}\|_F^2 - 2\mathrm{tr}(\tilde{B}^T R\tilde{A}) + \|\tilde{B}\|_F^2 \\ &= \|\tilde{A}\|_F^2 + \|\tilde{B}\|_F^2 - 2\mathrm{tr}(\tilde{A}\tilde{B}^T R). \end{aligned} \tag{15}$$

Applying the slip-in/slip-out method, we can write our objective function as

$$\begin{aligned} \|R\tilde{A} - \tilde{B}\|_F^2 &= \|\tilde{A}\|_F^2 + \|\tilde{B}\|_F^2 - \|\tilde{A}\tilde{B}^T\|_F^2 - \|R\|_F^2 \\ &\quad + \|\tilde{A}\tilde{B}^T\|_F^2 - 2\mathrm{tr}(\tilde{A}B^T R) + \|R\|_F^2 \\ &= \|\tilde{A}\|_F^2 + \|\tilde{B}\|_F^2 - \|\tilde{A}\tilde{B}^T\|_F^2 - n \\ &\quad + \|\tilde{A}\tilde{B}^T - R^T\|_F^2. \end{aligned} \tag{16}$$

Note the similarity of (16) with (7). The first four terms in the last equality of (16) are independent of R, so our problem becomes the optimization of $\|\tilde{A}\tilde{B}^T - R^T\|_F^2$.

Note the special structure of this final version of the optimization problem. We seek an orthogonal matrix that is closest to the 3×3 matrix $\tilde{A}\tilde{B}^T$. More generally, we consider the problem of determining

$$U_0 = \arg \min_{U \in O(n)} \|M - U\|_F \tag{17}$$

where M is an $n \times n$ full rank matrix. We solve this problem by first examining two special cases. First, suppose M is a diagonal matrix: $M = \text{diag}(d_1, \ldots, d_n)$. Then

$$\begin{aligned} \|M - U\|_F^2 &= \|D - U\|_F^2 \\ &= \|D\|_F^2 + \|U\|_F^2 - 2\text{tr}(DU) \\ &= \sum_{i=1}^n d_i^2 + n - 2\sum_{i=1}^n d_i u_{ii}. \end{aligned} \tag{18}$$

Since $\sum_{i=1}^n d_i^2$ and n are fixed, we need to maximize $\sum_{i=1}^n d_i u_{ii}$. Since $-1 \leq u_{ii} \leq 1$, we have that $\sum_{i=1}^n d_i u_{ii} \leq \sum_{i=1}^n |d_i|$ and it follows that $U_0 = \text{diag}(\text{sgn}(d_1), \ldots, \text{sgn}(d_n)) \in O(n)$. Furthermore, if D is positive definite, i.e., if $d_i > 0$ for $i = 1, \ldots, n$, then $U_0 = I$.

For our second case, suppose that M is symmetric. Then we can write M as $M = VDV^T$ where $V \in O(n)$ and $D = \text{diag}(d_1, \ldots, d_n)$. We then have

$$\|M - U\|_F = \|VDV^T - U\|_F = \|D - V^TUV\|_F \tag{19}$$

where the second equality in (19) follows from the fact that the Frobenius norm is invariant under pre- and post-multiplication by orthogonal matrices. By the first case, it follows that the optimal $U_0 \in O(n)$ is given by $V^TU_0V = \text{diag}(\text{sgn}(d_1), \ldots, \text{sgn}(d_n))$, i.e., the optimal orthogonal matrix is given by $U_0 = V\text{diag}(\text{sgn}(d_1), \ldots, \text{sgn}(d_n))V^T$. For the important case when M is symmetric positive definite, this becomes $U_0 = I$.

We are now ready for the general case when M is an arbitrary full rank $n \times n$ matrix. In this case, we can write M in its polar form: $M = PQ$ where P is symmetric, positive definite and $Q \in O(n)$. Then $\|M - U\|_F = \|PQ - U\|_F = \|P - UQ^T\|_F$, which is minimized over $U \in O(n)$ by the orthogonal matrix U_0 where $U_0Q^T = I$, i.e., $U_0 = Q$.

We thus conclude that the optimal solution for determining the orientation of the object in question is given by the orthogonal matrix $R = Q^T$ where PQ is the unique polar decomposition of $\tilde{A}\tilde{B}^T$. We note that when the polar decomposition has been mentioned in the literature as a solution, it has primarily appeared as an afterthought of the SVD solution. While these solutions are arguably equivalent, the SVD solution does not appear to be as natural as the polar decomposition solution. It is important to note once again that R is a rotation matrix if and only if $\det(Q) = 1$; otherwise, R is a reflection matrix.

C. Noise in Both the Model and Data

We next examine the case when there is noise in both the model and the data. In particular, we consider the following problem described in [6], given here with slightly different notation. Suppose we have two sets of N noisy observations given by two $3 \times N$ matrices A and B. We assume that the correspondence problem has already been solved so that the corresponding points are in the same order in A and B and that the translation of the object has already been determined. Our problem then is to find a rotation matrix R and perturbations δA and δB which satisfy

$$R(A + \delta A) = B + \delta B \tag{20}$$

such that $\|\delta A\|_F^2 + \|\delta B\|_F^2$ is minimized. The perturbations δA and δB correspond to noise in the model and data, respectively. This problem, presented in slightly different notation, was solved by Goryn and Hein in [6], but the solution presented there relies heavily on the introduction of some non-obvious substitutions. We present a more natural and intuitive derivation. We begin by first noting that unlike δB, the term δA appears in (20) as $R\delta A$, suggesting that it may be better to rewrite the cost function as $\|\delta A\|_F^2 + \|\delta B\|_F^2 = \|R\delta A\|_F^2 + \|\delta B\|_F^2$, where equality follows from the fact that R is orthogonal. Next, equation (20) can be written as $R\delta A - \delta B = -(RA - B)$. We thus want to minimize $\|R\delta A\|_F^2 + \|\delta B\|_F^2$ subject to $R\delta A - \delta B = -(RA - B)$.

The scalar version of the preceding problem statement suggests a solution to this constrained optimization problem. In the scalar case, we want to minimize $x^2 + y^2$ over the scalars x, y, z subject to the constraint $ax + by = f(z)$ where $f(z) = -cz + d$ and where a, b, c, d are fixed parameters. Geometrically, for a fixed z, the constraint can be interpreted as the equation of a line and the optimal solution over x, y would then correspond to the point on that line which is closest to the origin. We then want to choose z to place the line as close to the origin as possible, which is achieved by minimizing $|f(z)| = |cz - d|$. Once this is done, the optimal x and y can be determined. This suggests minimizing the norm of $RA - B$ in the non-scalar case.

We now present a formal derivation of the solution for the non-scalar case. The constraint (20) in the general case can be written in matrix notation as

$$\begin{bmatrix} \frac{1}{\sqrt{2}}I & \frac{-1}{\sqrt{2}}I \end{bmatrix} \begin{bmatrix} R\delta A \\ \delta B \end{bmatrix} = \frac{-1}{\sqrt{2}}(RA - B) \tag{21}$$

where the $1/\sqrt{2}$ term is included so that the rows of the matrix on the left are not only mutually orthogonal, but are also normalized. This suggests augmenting the matrix so that it becomes orthogonal:

$$\frac{1}{\sqrt{2}} \begin{bmatrix} I & -I \\ I & I \end{bmatrix} \begin{bmatrix} R\delta A \\ \delta B \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} -(RA - B) \\ R\delta A + \delta B \end{bmatrix}. \tag{22}$$

Since the Frobenius norm is invariant under multiplication by orthogonal matrices, we have that the cost function $\|\delta A\|_F^2 + \|\delta B\|_F^2$ is given by

$$\|R\delta A\|_F^2 + \|\delta B\|_F^2 = \frac{1}{2}\|RA - B\|_F^2 + \frac{1}{2}\|R\delta A + \delta B\|_F^2, \tag{23}$$

which is clearly minimized by setting $\delta B = -R\delta A$ and minimizing $\|RA - B\|_F^2$ over the family of orthogonal matrices R so that R is determined in the same manner as before.

IV. CONCLUSION

In this article, we have presented the polar decomposition as the natural method for solving the absolute orientation problem. Previously, the polar decomposition approach was mentioned in the literature as merely an afterthought of the SVD approach. The polar decomposition method for the 3D case was motivated by a complex number approach to the 2D case, which is interesting in its own right. Lastly, we have provided a simpler and more natural proof that the same solution also holds for a noisy model.

ACKNOWLEDGMENT

The authors gratefully acknowledge the support of the Air Force Research Lab.

REFERENCES

[1] K. S. Arun, T. S. Huang, and S. D. Blostein, "Least-squares fitting of two 3-D point sets," *IEEE Trans. PAMI*, vol. 9, no. 5, pp. 698–700, Sept. 1987.

[2] B. Kamgar-Paris and B. Kamgar-Paris, "Matching sets of 3D line segments with application to polygonal arc matching," *IEEE Trans. PAMI*, vol. 19, no. 10, pp. 1090–1099, Oct. 1997.

[3] B. K. P. Horn, H. M. Hilden, and S. Negahdaripour, "Closed-form solution of absolute orientation using orthonormal matrices," *Journal of the Optical Society of America A*, vol. 5, no. 7, pp. 1127-1638, July 1988.

[4] B. K. P. Horn, "Closed-form solution of absolute orientation using unit quaternions," *Journal of the Optical Society of America A*, 4(4):629-642, April 1987.

[5] M. D. Wheeler and K. Ikeuchi, "Iterative estimation of rotation and translation using the quaternion," Technical Report CMU-CS-95-215, Carnegie-Mellon University, December 1995.

[6] D. Goryn and S. Hein, "On the estimation of rigid body rotation from noisy data, " *IEEE Trans. PAMI*, vol. 17, no. 12, pp. 1219–1220, Dec. 1995.