Estimating 3-D Orientation from 1-D Projections

K. J. Dillon*, B. S. Denney, and R. J. P. de Figueiredo
Neural Computing Systems, Irvine, California, USA

ABSTRACT

We focus on the specific problem of inversion of a collection of one-dimensional projections to reconstruct a three-dimensional image of a rigid body and estimate its orientation. We assume correspondence of the points in the projections is given and derive several useful results leading to an iterative solution algorithm and a fundamental understanding of its possible scope. Finally, the algorithm was simulated on synthetic HRR data which verified its function and convergence.

Keywords: Inverse imaging, Pose Estimation, High Range Resolution (HRR) Radar

1. INTRODUCTION

This work was motivated by the problem of improving pose estimates of vehicles using radar to improve tracking and recognition algorithms. As ground vehicle movement is restricted to a two-dimensional surface, we were mainly concerned with the component of orientation within this surface, referred to as yaw in the aviation field[1]. When we apply the theory to the application of pose estimation, therefore, we specifically are concerned with the component of orientation locally parallel to the ground plane. The remaining components of orientation, namely the pitch and roll, will be contained collectively in a "tilt" variable for our purposes. However, we will approach the problem of estimating three-dimensional orientation generally in the next section.

In our analysis, we make the following assumptions:

1. The object under observation is a rigid body. This is an obvious assumption for application to vehicles generally. It is weaker perhaps when working in fields such as medical tomographic imaging or face recognition.

2. The plane wave approximation is valid. This is a very safe assumption in dealing with vehicles observed from afar as the wavefront curvature on the scale of the vehicle should be negligible. A similar approach to ours would probably be successful without using this approximation, but for the applications we are interested in the additional complications would be unnecessary.

3. The collection of projections span three dimensions. Additionally, as has been discussed elsewhere, a minimum of five looks containing nine points each is needed for a unique solution[2]. What is meant by uniqueness of the solution for our purposes will be discussed later.

4. Correspondence of points between the projections is given. This is the most sweeping assumption made in this analysis. Confidently achieving correspondence at times can be very difficult, and the success of doing so can be very data and target dependent. We will not address this problem here.

This approach when applied to radar is based on the modeling of the object as a collection of scattering centers, and the HRR signature as a one-dimensional projection of those scatterers[3]. Each individual signature is assumed to be an observation of the object, from which we extract range measurements to each of the scatterers, from a separate viewpoint, or "look". This is shown in the following figure.

---

* For further information email dillon@ncs.com; www.neuralsystems.com

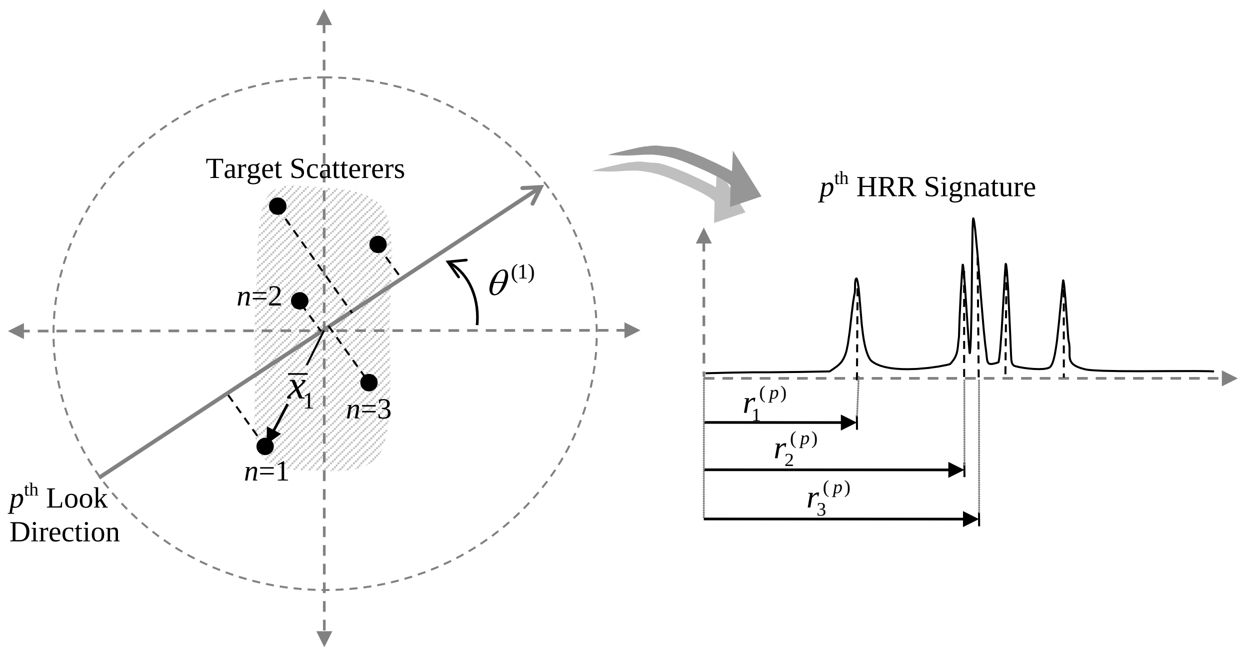

Figure 1: Two-dimensional example depicting notation, and showing the application to a HRR signature.

2. THEORY

Our first objective is to derive analytic solutions for the optimal estimate of the object's orientation and three-dimensional point locations, which we refer to as "model". We use the mean-squared error as our criterion for optimality. The return data model $r_n^{(p)}$ which is the range for a particular look and scatterer follows:

$$\begin{gathered} r_n^{(p)} = \left(\bar{x}_n - \bar{x}_0\right)^T R^{(p)}\, \bar{w} \\ \bar{w}, \bar{x}_n \in \Re^3, \\ R \in \mathrm{SO}(3) \end{gathered} \tag{1}$$

where $\bar{w}$ is the look vector, $\bar{x}_n$ is the location of the n^{th} individual scatterer and $R^{(p)}$ is the three-dimensional rotation[4] of the target model about $\bar{x}_0$, some fixed point on the object, for the p^{th} look. Note that we are only concerned with relative locations of the points on the object; the absolute location of the object is of no concern. The mean-squared error in our estimate is therefore

$$\bar{E}(R) = \sum_{p=1}^{P}\sum_{n=1}^{N}\left[\left(\bar{x}_n - \bar{x}_0\right)^T R^{(p)}\, \bar{w} - r_n^{(p)}\right]^2 \tag{2}$$

where there are P looks and N scatterers in the data.

1.1 Finding the Orientation Given the Model

First we will only be concerned with averaging over n (the scatterer index) as we are concentrating on a particular p (the look index). Our approach is to find the optimal orientation by taking the derivative of the mean-squared error with respect to rotations as follows.

$$\overline{E}(R)=\sum_{n=1}^{N}\left[\left(\overline{x}_n-\overline{x}_0\right)^T R^{(p)}\,\overline{w}-r_n^{(p)}\right]^2 \tag{3}$$

For some function *J*(*R*) of a rotation *R*, the directional derivative in unit rotation direction *D* is

$$J'(R:D)=\lim_{\alpha\to\infty}\frac{J(D^{\alpha}R)-J(R)}{\alpha} \tag{4}$$

where D^{α} is a rotation in the amount α about an axis *l* [5]. The derivative is taken and set to zero for all *l* leading to the necessary condition

$$R^{(p)}\,\overline{w}=\left\{\sum_{n=1}^{N}\left[\left(\overline{x}_n-\overline{x}_0\right)\cdot\left(\overline{x}_n-\overline{x}_0\right)^T\right]-\lambda I\right\}^{-1}\sum_{n=1}^{N}\left[\left(\overline{x}_n-\overline{x}_0\right)r_n^{(p)}\right] \tag{5}$$

where *I* is the 3 by 3 identity matrix and λ is some scalar. This result is not yet complete as we note that not all values of λ will yield a solution, though there should be more than one solution. The case of λ =0 fits with the straightforward derivation of optimal $R^{(p)}\,\overline{w}$ performed (as opposed to optimal $R^{(p)}$) by setting the gradient of equation (3) with respect to $R^{(p)}\,\overline{w}$ equal to zero yields

$$R^{(p)}\,\overline{w}=\left\{\sum_{n=1}^{N}\left[\left(\overline{x}_n-\overline{x}_0\right)\cdot\left(\overline{x}_n-\overline{x}_0\right)^T\right]\right\}^{-1}\sum_{n=1}^{N}\left[\left(\overline{x}_n-\overline{x}_0\right)r_n^{(p)}\right] \tag{6}$$

The difference here is the optimum is taken over all three dimensions of space rather than simply over rotation yielding a potentially different result in the presence of noise. The forms of the results can be made identical by adding the constraint that the length of $R^{(p)}\,\overline{w}$ must be constant, i.e. rotations do not distort length.

Geometrically, this result describes an infinite number of solutions for $R^{(p)}$ as a single-look projection cannot resolve each scatterer to better than some rotation about the look vector. Our approach to dealing with this ambiguity in our application is to assume the vehicle is flat on the ground.

1.1.1 Finding Solutions of optimally rotated look vector

We will now derive the analytical solutions for λ. Since $\overline{w}$ is a unit vector by definition and $R^{(p)}$ is a rotation

$$\left(R^{(p)}\,\overline{w}\right)^T\left(R^{(p)}\,\overline{w}\right)=1 \tag{7}$$

and therefore

$$\left(\sum_{n=1}^{N}\left[\left(\overline{x}_n-\overline{x}_0\right)r_n^{(p)}\right]\right)^T\left(\left\{\sum_{n=1}^{N}\left[\left(\overline{x}_n-\overline{x}_0\right)\cdot\left(\overline{x}_n-\overline{x}_0\right)^T\right]+\lambda I\right\}^{-1}\right)^2\sum_{n=1}^{N}\left[\left(\overline{x}_n-\overline{x}_0\right)r_n^{(p)}\right]=1 \tag{8}$$

If we let $M=\sum_{n=1}^{N}\left[\left(\overline{x}_n-\overline{x}_0\right)\cdot\left(\overline{x}_n-\overline{x}_0\right)^T\right]$ and $s=\sum_{n=1}^{N}\left[\left(\overline{x}_n-\overline{x}_0\right)r_n^{(p)}\right]$ writing the solution as a polynomial in λ yields

$$
\begin{aligned}
&\left[|M|^2 - s^T \operatorname{adj}(M)^2 s\right] + \\
&\left[2|M|\operatorname{trace}(\operatorname{adj}(M)) - 2\operatorname{trace}(M) s^T \operatorname{adj}(M) s + 2 s^T \operatorname{adj}(M) M s\right]\lambda + \\
&\begin{bmatrix} \operatorname{trace}(\operatorname{adj}(M))^2 + 2|M|\operatorname{trace}(M) - \operatorname{trace}(M)^2 s^T s + \\ 2\operatorname{trace}(M) s^T M s - s^T M^2 s - 2 s^T \operatorname{adj}(M) s \end{bmatrix}\lambda^2 + \\
&\left[2|M| + 2\operatorname{trace}(\operatorname{adj}(M))\operatorname{trace}(M) - 2\operatorname{trace}(M) s^T s + 2 s^T M s\right]\lambda^3 + \\
&\left[\operatorname{trace}(M)^2 + 2\operatorname{trace}(\operatorname{adj}(M)) - s^T s\right]\lambda^4 + \\
&\left[2\operatorname{trace}(M)\right]\lambda^5 + \\
&\lambda^6 = 0
\end{aligned}
\tag{9}
$$

It is straightforward to calculate the roots of this polynomial computationally and get the solutions, λ, and it is a simple matter to determine which root yields the best mean squared fit. It should be possible to use a geometrical insight to select the solution, however. The meaning of all these solutions remains an open question.

1.1.2 General Form of Solution

Once we have selected λ we can calculate the product $R^{(p)}\,\overline{w}$. We now have the equation

$$\overline{a} = R^{(p)}\,\overline{w} \tag{10}$$

where $\overline{a}$ is the known vector or the product we solved for, and $\overline{w}$ is also known. In general we can see that there are an infinite number of possible matrices which are solutions to this problem. We can however find the form of the general solution for $R^{(p)}$. Assuming we know of a particular rotation $R_0^{(p)}$ which is a solution,

$$\overline{a} = R_0^{(p)}\,\overline{w} \tag{11}$$

For example we could use the rotation which rotates $\overline{w}$ to $\overline{a}$ within the plane determined by the two vectors. This can be computed by simple geometry using the dot and cross products. Now, we know in general we can write any rotation as the product of multiple rotations, so our general solution can be described as the product of $R_0^{(p)}$ with another rotation. Therefore we describe the general solution to our problem as

$$\overline{a} = R_0^{(p)} R_1^{(p)}\,\overline{w} \tag{12}$$

where $R_1^{(p)}$ is unknown so far. But we notice that since all rotation matrices are invertible we can manipulate this as

$$R_0^{-1(p)}\overline{a} = R_1^{(p)}\,\overline{w} \tag{13}$$

but since $R_0^{(p)}$ is a solution by definition, the left hand side of the above equation simply equals $\overline{w}$ leaving $R_1^{(p)}$ be a solution to

$$\overline{w} = R_1^{(p)}\,\overline{w} \tag{14}$$

This means $\overline{w}$ remains unchanged when the rotation $R_1^{(p)}$ is applied to it. Since $\overline{w}$ is a vector, $R_1^{(p)}$ must in general be a rotation about $\overline{w}$ in some angle, ψ. We write this as $R(\psi, \overline{w})$. Therefore we can write the general solution to our initial problem as

$$R^{(p)} = R\left(\cos^{-1}(\overline{a}\cdot\overline{w}), \overline{a}\times\overline{w}\right)R(\psi, \overline{w}) \tag{15}$$

where we used the dot and cross products between $\overline{a}$ and $\overline{w}$ to get the angle and direction of one particular solution. ψ could be any angle since the projection for any given look p will be ambiguous to a rotation about the look vector. To resolve this ambiguity, we make use of assumptions about the vehicle orientation with respect to the ground to approximate the complete pose information. Basically, we assume the vehicle sits flat in the ground plane with a small unknown pitch and roll (which we combine into a general "tilt"). Multiple looks over a sufficient range of angles allow this to be done without ever performing recognition.

1.2 Finding Model Given Orientation

Now we consider the case of known $R^{(p)}$ and unknown $\overline{x}_n$ in the signal model.

$$r_n^{(p)} = \left(\overline{x}_n - \overline{x}_0\right)^T R^{(p)}\,\overline{w} \tag{16}$$

The squared error we now wish to consider is

$$\overline{E}^2 = \sum_{p=1}^{P}\left[\left(\overline{x}_n - \overline{x}_0\right)^T R^{(p)}\,\overline{w} - r_n^{(p)}\right]^2 \tag{17}$$

since we must average over multiple looks to find the optimal location of each scatterer $\overline{x}_n$. Taking the gradient with respect to $\overline{x}_n$ and setting equal to zero yields

$$\left(\overline{x}_n - \overline{x}_0\right) = \left\{\sum_{p=1}^{P}\left[\left(R^{(p)}\,\overline{w}\right)\left(R^{(p)}\,\overline{w}\right)^T\right]\right\}^{-1}\sum_{p=1}^{P}\left[r_n^{(p)}R^{(p)}\,\overline{w}\right] \tag{18}$$

from which the locations of the scatterers can be computed directly

1.2.1 Simultaneously Solving for Model and Orientation

To analyze the potential ambiguities which would limit an algorithm which solves for model and orientation simultaneously, we write the equation (2) error in matrix form as

$$\overline{E}(R) = \sum_{rows}\sum_{cols}\left[X^T A - D\right]^T\left[X^T A - D\right] \tag{19}$$

where X is a 3 by N matrix whose columns consist of the vectors $\overline{x}_n - \overline{x}_0$. Similarly, A is a 3 by P matrix with columns $\overline{a}^{(p)}$. D is a N by P matrix in which the $(n, p)^{\text{th}}$ element is $r_n^{(p)}$. In this form it is easy to see the form of potential ambiguities in our minimum error result. Solutions for X and A which result in the same value for the product $X^T A$ will yield the same error minimum. In general, if we get different solutions X' and A', we can write them as UX and VA , with unknown 3 by 3 matrices U and V. For this to result in the same error minimum, we require

$$X^T A = (X')^T A' = (UX)^T (VA) = X^T (U^T V) A \tag{20}$$

And it follows that for $X^T U^T VA$ to equal $X^T A$, $U^T V$ must equal the identity matrix. Additionally, since the columns of both A and A' must be unit length (they are rotated unit vectors $R^{(p)}\hat{w}$), the matrix V must preserve length, i.e.

$$\|V y\| = \|y\| \tag{21}$$

for any vector y. This leads to the following observation,

$$\begin{aligned} (V\ y)^T (V\ y) &= y^T y \\ y (V^T V) y &= y^T y \end{aligned} \tag{22}$$

which implies $V^T V$ equals the identity. This along with our previous result means U must equal V, and is by definition an orthogonal matrix. Visually, this means the transformation consists of any possible combination of a rotation and a reflection.

As a result, we can see the minimum error result can be ambiguous within a possible orthogonal transformation (rotation and/or reflection), U, of the pose vectors, A, estimated (also shown considering the analogous tomography situation in [2]). The target model, X, will be transformed by U^T which is the opposite rotation/reflection.

The strategy for solving for the model and orientation simultaneously would be to start with a rough estimate of either model or orientation and solve for the other, and iterate this process. Due to the ambiguity in orthogonal transform between the model and orientation when solved simultaneously, we can expect our algorithm's resulting accuracy to be somewhat limited by the value of our initial estimate.

3. RESULTS

We considered how the algorithm was capable of converging assuming accurate range information and various degrees of inaccuracy in the initial estimates of pose and tilt of the vehicle. Here we use pose to mean the yaw as mentioned earlier. The tilt, i.e. the combination of roll and pitch, is assumed to be always unknown and randomly varying somewhat between looks as the vehicle bounces over uneven terrain for example. The following figure illustrates a single simulation where HRR data is generated synthetically for several look angles. In this simulation the truth is known but errors are added.

The random numbers vary between each look and between each experiment. The poses of the algorithm output are compared to the known poses of the model within the synthetic data and the heading estimate errors are plotted in Figure 3 for a range of tests. In this case, the algorithm was run using new synthetic data twenty times for a wide range of input error standard deviations and for four different tilt variances.

The error in the algorithm's final result is plotted versus the error in the initial estimates used by the algorithm. For example, we found that if we are given HRR data describing a target and the pose of the target was already known within ten degrees, we could improve the accuracy of the pose estimate to within about one or two degrees. As would be expected, the algorithm performance degrades as the initial estimates become more inaccurate, but there is still a significant improvement.

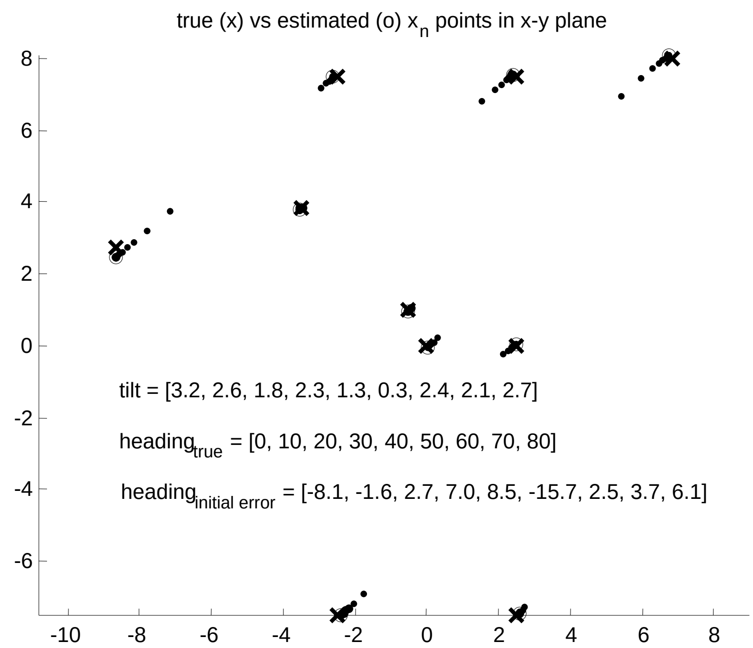

Figure 2: Example simulated model truth vs. converged estimate. The o's are the estimates of the model's scatterer locations and the x's are their true locations. The dots are intermediate locations computed prior to convergence. The tilts and angle errors are synthetic random numbers chosen to produce a specific variance.

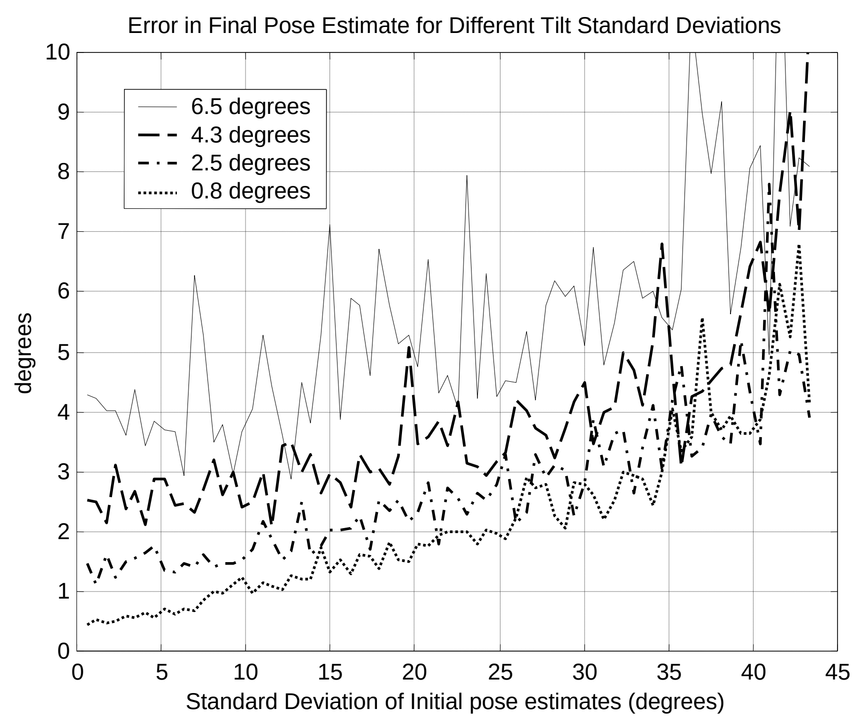

Figure 3: Algorithm Resulting Accuracy

4. CONCLUSION

A number of useful theoretical results were presented in this work concerning the problem of estimating orientation of an object from a collection of one-dimensional projections, which led to the development of a successful pose estimation algorithm. Solutions for calculating the orientation given the model and vise versa were found, and the general forms of the solutions were shown. While it may be possible to solve analytically for the model and orientation simultaneously, this was not necessary for our purposes as we developed a successful iterative algorithm which was shown in simulations to converge to the correct solution.

5. ACKNOWLEDGEMENTS

Portions of this work were funded by the U.S. Air Force, contract number F33615-02-M-1216.

6. REFERENCES

1. B. C. P. Horn, *Robot Vision*, McGraw-Hill, New York, 1986.
2. Basu, S. and Y. Bresler, "*Uniqueness of Tomography with Unknown View Angles*", IEEE Transactions on Image Processing, volume 9, number 6, pp. 1094-1106, 2000.
3. B. S. Denney, R. J. de Figueiredo, *"Scattering-based tomography for HRR and SAR prediction"*, Multidimensional Systems and Signal Processing, 14, pp. 207-222, 2003.
4. S. Altmann, *Rotations, Quarternions and Double Groups*, Oxford University Press, New York, 1986.
5. B. S. Denney, Rotation *Clustering for Robot Vision*, PhD. Thesis, University of California, Irvine, 1996.