On Robust Exterior Orientation

Robert M. Haralick
Kenneth B. Thornton

Abstract

This paper discusses (1) the method by which we have robustified the exterior orientation algorithm to eliminate outliers and (2) our method for analyzing the performance of the algorithm. Several experimental results are presented to illustrate the efficacy of the proposed techniques.

1 Introduction

In the problem of exterior orientation, we seek to recover the position and orientation of the camera, given a set of corresponding fixed 3D model points and noisy observed 2D perspective projection image points. Algorithms for solving this problem differ in their parameterization of the rotation matrix. Older algorithms used Euler angles, while more recent algorithms use quaternions. We use an algorithm based on the quaternion representation. It has been fully described elsewhere (Hinsken, 1988; Haralick and Shapiro, 1991).

If some of the corresponding model and image point pairs are not correct, the standard least-squares exterior orientation algorithm will fail. In the next section, we describe the method by which we make the standard least-squares algorithm robust. In section 3, we show the results of some experiments that verify the efficacy of the proposed technique.

We then consider the following scenario. We are given a number of corresponding sets of model and image points. For each set of points, the measurement noise on the image points is considered to be i.i.d. from a $N(0,\sigma^2)$ distribution. However, σ^2 is not guaranteed to be the same for each set. Under these conditions, and given a limit on the allowable rotation error, we show how the false alarm and misdetection rate of a classification task associated with this performance requirement varies as a function of the setting of a threshold on a particular test statistic. Section 4 sets up this problem and Section 5 shows the results of our experiments. Finally, in Section 6 we give a few concluding comments.

2 Robustifying Exterior Orientation

First, let us establish some notation. On the variables which follow, we use no superscript to denote known or fixed quantities, the tilde superscript to denote observed quantities, and the hat superscript to denote estimated or fitted quantities.

So, let (x_n, y_n, z_n), (p_n, q_n, s_n), and (u_n, v_n) be the true coordinates of the n^{th} point in model, camera, and perspective projection image coordinates, respectively. The relationship between (x_n, y_n, z_n) and (u_n, v_n) is given in two parts. The first part consists of a transformation to camera coordinates

$$\begin{pmatrix} p_n \\ q_n \\ s_n \end{pmatrix} = \mathbf{R}(a, b, c, d) \begin{pmatrix} x_n - x_0 \\ y_n - y_0 \\ z_n - z_0 \end{pmatrix}, \tag{1}$$

where $\mathbf{R}(a, b, c, d)$ is a rotation matrix parameterized by a quaternion with components (a, b, c, d) and (x_0, y_0, z_0) is the camera position in model coordinates. The second part consists of a perspective projection onto the image

$$\begin{pmatrix} u_n \\ v_n \end{pmatrix} = \frac{f}{s_n} \begin{pmatrix} p_n \\ q_n \end{pmatrix}, \tag{2}$$

where f is the focal length. The observed image position is related to the ideal image position by

$$\begin{pmatrix} \tilde{u}_n \\ \tilde{v}_n \end{pmatrix} = \begin{pmatrix} u_n \\ v_n \end{pmatrix} + \begin{pmatrix} \Delta u_n \\ \Delta v_n \end{pmatrix}, \tag{3}$$

where Δu_n and Δv_n are considered to be independent zero-mean Gaussian random variables with variance σ^2 if the n^{th} point is not an outlier.

The exterior orientation problem may be stated as follows. Given

$$\begin{pmatrix} \tilde{u}_n \\ \tilde{v}_n \end{pmatrix}, \begin{pmatrix} x_n \\ y_n \\ z_n \end{pmatrix}, \qquad n = 1, \ldots, N \tag{4}$$

determine the vector $\hat{\Theta}' = (\hat{x}_0, \hat{y}_0, \hat{z}_0, \hat{a}, \hat{b}, \hat{c}, \hat{d})$ which minimizes

$$\epsilon^2 = \sum_{n=1}^{N} w_n((\tilde{u}_n - \hat{u}_n)^2 + (\tilde{v}_n - \hat{v}_n)^2) \tag{5}$$

where w_n are weights such that $w_n \geq 0$ and $\sum_{n=1}^{N} w_n = N$, and $(\hat{u}_n, \hat{v}_n)$ are the projections of the model points.

Usually, all the weights are taken as unity and the nonlinear regression problem is solved by iteratively linearizing the model about the current estimated vector of parameters, solving for corrections to the parameters, forming the new parameter vector, relinearizing, etc.. In the exterior orientation problem, an initial guess for the vector of parameters can be obtained from 3 matching point pairs using a 3-point perspective projection algorithm (Haralick et. al., 1991). Estimation of the corrections and determination of the final covariance matrix of the parameters follows the standard Gauss-Markov model (see, for example, Koch, 1988).

If some of the corresponding point pairs are incorrect, the standard equally weighted least squares (EWLS) technique is known to produce poor results (Haralick et. al., 1989). In our robust technique, we reweight each point after each iteration of the procedure. After iteration k, we have the fitted values $(\hat{u}_n^k, \hat{v}_n^k)$ and the residual

$$\hat{r}_n^k = \begin{pmatrix} \tilde{u}_n \\ \tilde{v}_n \end{pmatrix} - \begin{pmatrix} \hat{u}_n^k \\ \hat{v}_n^k \end{pmatrix}.$$

In order to determine the new weight for point n, we first determine its robust Mahalanobis distance. Given a set $\{x_1, \ldots, x_N\}$ of points with sample average $\bar{x}$ and sample covariance S (assumed to be positive-definite), the Mahalanobis distance of any point n from the average $\bar{x}$ is defined as (Seber, 1984)

$$m_n = ((x_n - \bar{x})' S^{-1} (x_n - \bar{x}))^{\frac{1}{2}}.$$

We have used the projection algorithm described in Rousseeuw and van Zomeren (1990) to obtain approximate robust distances. It is also possible, of course, to estimate these distances directly if robust estimates of multivariate mean and covariance are available.

Let $M = (m_1^2, \ldots, m_N^2)$ be the set of all univariate distances. This set should be distributed as χ_2^2. If a random variable l is distributed as χ_2^2, then we know (Abramowitz and Stegun, 1972) that

$$\Pr\{l \leq 9.21\} = 0.99.$$

From each m_n^2, we calculate the new weight as follows

$$w_n = \begin{cases} \exp\left(-\frac{m_n^2}{2}\right), & \text{if } m_n^2 \leq 9.21; \\ 0, & \text{otherwise.} \end{cases}$$

This reweighting function was proposed by Krarup (see Förstner, 1989).

After determining the weights, they are normalized to sum to N. Notice that this procedure rejects some residuals as outliers. After the optimization procedure converges, all points having zero-weight are assume to be outliers and the least-squares analysis described above is performed by subtracting from N the appropriate number of outlier points.

In summary, the complete iteratively-reweighted least-squares procedure is as follows

(1) Get initial guess Θ^t. For all n, set w_n 1.

(2) Form residuals r_n and determine the squared error ϵ^2.

(3) Determine univariate squared Mahalanobis distances m_n^2.

(4) Determine weights w_n.

(5) Using weights, determine $\Delta\Theta$, the correction to the current parameter vector.

(6) Determine new residuals r_n and new ϵ^2.

(7) If a fixed number of iterations is reached or if the new ϵ^2 is greater than the old ϵ^2, stop. Otherwise, go to (3).

3 Experimental Protocol and Results

In our experiments, we were concerned with the accuracy of the rotation estimate. The true quaternion $Q = (a, b, c, d)$ and estimated quaternion $\hat{Q} = (\hat{a}, \hat{b}, \hat{c}, \hat{d})$ are both unit vectors. The rotation error measure we used to describe their difference was $e = \log(1 -$

Average Log Rotation Error vs. SNR								
SNR	Non-Robust Without Outliers				Robust With Outliers			
	NG=25	NG=22	NG=18	NG=14	NG=25	NG=22	NG=18	NG=14
80	-9.15	-9.05	-8.91	-8.70	-9.06	-8.99	-8.87	-8.68
70	-8.15	-8.05	-7.91	-7.70	-8.06	-7.98	-7.87	-7.68
60	-7.15	-7.05	-6.91	-6.70	-7.06	-6.98	-6.87	-6.68
50	-6.15	-6.05	-5.91	-5.70	-6.06	-5.98	-5.87	-5.68

Table 1: The average log-rotation error vs. SNR for various amounts of outliers using the equally-weighted least-squares (EWLS) technique with outliers removed and the robust least-squares (RLS) technique with outliers included. NG is the number of good data points. For the robust technique, the total number of data points was 25.

$Q \cdot \hat{Q}$). Since a rotation may be represented as a point on the unit sphere S^3 in 4D space, we could have used arc distance on the sphere as the error criterion. The distance we chose is easier to compute.

In our experiments, the number of image points was fixed at 25 and spread evenly over a 5×5 grid covering the unit square $[-1,1] \times [-1,1]$. The configuration was fixed throughout the experiments. The focal length was set to give a field of view of $\pi/2$ radians. The signal-to-noise ratio (SNR) was defined as $-20\log_{10}(\frac{\sigma}{S})$, where $S = 2$ is the image sidelength. Thus, the SNR determines the noise variance σ^2. Outliers were generated by replacing an image point chosen at random by another point sampled from a uniform $[-1,1] \times [-1,1]$ distribution. For each trial, we determine the corresponding set of points in camera coordinates by backprojecting each image point by an amount uniformly distributed over the range $[10,30]$. The location of each coordinate of the origin of the world coordinate system given in camera coordinates is also generated uniformly in the range $[10,30]$. A random rotation is generated by rotating each axis by a random amount generated uniformly over the range $[0,\pi]$. Given the rotation matrix, it is then possible to determine the location of the camera in the coordinates of the object system. The quaternion representing the rotation was obtained from the rotation matrix.

To test the ability of the algorithm to reject outliers, we tested the performance of the equally-weighted least-squares algorithm with outliers removed against the performance of the robust algorithm without the outliers removed. We set the percentage of outliers (PO) to be one of $\{0,15,30,45\}$ and the signal-to-noise ratio SNR to be one of $\{50,60,70,80\}$. For each setting of PO and SNR, we ran 1000 trials. We then compared the average rotation error e obtained using both the equally-weighted least-squares algorithm (with outliers removed) and the robust least-squares algorithm with outliers. We also compared the average estimated SNR to the true SNR using both algorithms. Tables 1 and 2 show these results, respectively. The similarity of performance of the 2 algorithms confirms that we were able to reject the outliers using the proposed technique.

Average Estimated SNR in dB vs. Number of Good Points								
NG	Non-Robust Without Outliers				Robust With Outliers			
	80dB	70dB	60dB	50dB	80dB	70dB	60dB	50dB
25	80.099	70.099	60.099	50.099	79.906	69.900	59.886	49.899
22	80.099	70.099	60.099	50.099	79.909	69.915	59.915	49.909
18	80.128	70.128	60.128	50.128	80.044	70.000	59.995	49.941
14	80.174	70.174	60.174	50.174	80.137	70.135	60.143	50.154

Table 2: The average estimated SNR in dB vs. the number of good data points using both the equally-weighted least-squares (EWLS) technique with outliers removed and the robust least-squares (RLS) technique with outliers included. NG is the number of good data points. For the robust technique, the total number of data points was 25.

4 Performance Criterion

We envision a scenario where there is a performance criterion on the estimated quaternion $\hat{Q} = (\hat{a}, \hat{b}, \hat{c}, \hat{d})$. For example, a manufacturer may desire that the error between the true quaternion and the estimated quaternion not exceed some threshold value. In this environment, we assume that the noise which perturbs the measured image point positions is i.i.d. within each image, but may vary between images.

The algorithm does not know Q, the true quaternion, but must decide whether the error $e = \log(1 - Q \cdot \hat{Q})$ between the true and estimated quaternions exceeds some threshold. We use the statistic $t = \log(\text{Var}(\hat{a}) + \text{Var}(\hat{b}) + \text{Var}(\hat{c}))$ to make this determination. (Since the quaternion has only 3 degrees of freedom, we only estimate corrections to $\hat{a}^k$, $\hat{b}^k$,$\hat{c}^k$ during each iteration k and thus only use the estimated variance of these components.)

5 Experimental Protocol and Results

In this set of experiments, the corresponding model and image points were generated as before. In this case, no outliers were added to the data and the signal-to-noise ratio SNR was generated uniformly over the range [50, 90]. Over 1000 trials, we tabulated t vs. e, where t and e are the test statistic and error measure described above, respectively. We then formed the joint cumulative distribution of e and t by first constructing a table with grid points $\{e_0, \ldots, e_A\}$, and $\{t_0, \ldots, t_B\}$, so that $e_0 < e_1 < \cdots < e_A$ and $t_0 < t_1 < \cdots < t_B$, where $e_0 = e_{\min}$, $e_A = e_{\max}$, $t_0 = t_{\min}$, and $t_B = t_{\max}$. Each bin (e_i, t_j) in the table contained the count $\{\#(e,t) | e < e_i, t < t_j\}$.

As mentioned above, we assume that the user desires that e is below some threshold e_{thr}. If the true e is above this threshold, the performance criterion is not met. If it is below this threshold, the performance criterion is met. In practice, the program does not know e. It only knows t. Thus, the program decides whether the performance criterion is met by setting a threshold t_{thr} on t. Thus, for any combination of thresholds e_{thr} and t_{thr} we can estimate the false alarm and misdetection rates of the algorithm. The false alarm (FA) rate is the probability that when the performance is met, the program says it is not. The misdetection (MD) rate is the probability that when the performance is not met, the program says it is. Letting, $N_1 = \{\#(e,t)|e < e_{thr}, t < t_{thr}\}$, $N_2 = \{\#(e,t)|e < e_{thr}, t > t_{thr}\}$, $N_3 = \{\#(e,t)|e > e_{thr}, t < t_{thr}\}$, and $N_4 = \{\#(e,t)|e > e_{thr}, t > t_{thr}\}$. we defined FA by $\frac{N_2}{N_1+N_2}$ and MD by $\frac{N_3}{N_3+N_4}$. In addition, the error rate is given by $\frac{N_2+N_3}{1000}$.

In this set of experiments, we placed a prior on the Gaussian noise that was uniform on a log scale. This reflects the assumption that although the noise standard deviation is not known, higher noise standard deviations are less likely than low noise standard deviations.

Figure 1 illustrates the FA rate versus the MD rate as the performance criterion e_{thr} is varied from 0.30 to 0.80 of its maximum value obtained over the 1000 trials. When the SNR is 50 dB, this corresponds to a noise standard deviation of $3\sigma \approx S/100$, where S is the image sidelength. For a 500×500 pixel image, this corresponds essentially to bounding the noise perturbation to plus or minus 5 pixels. If the quaternion were decomposed into the axis and angle representation of a rotation, and we assumed that the rotation axis was estimated perfectly, the error criterion we used is related to the error in estimating the angle of rotation about this axis. Then, assuming the rotation axis to be known, as the error in degrees between the estimated angle of rotation about the rotation axis and the true angle of rotation about the rotation axis goes down by a factor of 2, the log rotation error goes down by approximately 0.6. Figure 1 shows that as long as the performance requirement is 40% or greater than the maximum obtained over 1000 trials, the misdetection and false alarm rates can both be kept below 10%.

Each setting of the threshold t_{thr} gives a point on the graph. Figure 2 illustrates the error rate versus t_{thr}. On this graph it is quite clear that by choosing the threshold t_{thr} properly, the error rate of the algorithm can be minimized.

6 Conclusion

In this paper we have shown how the rotation error of the exterior orientation problem varies as a function of the signal-to-noise ratio. By robustifying the standard equally-weighted least-squares algorithm, we were able to achieve nearly the same results using the robust least-squares algorithm with outliers as with the equally-weighted least-squares algorithm with outliers removed. We also saw that if there is a uniform prior placed on the SNR, it is possible to judiciously choose a criterion threshold for minimizing the error rate of the algorithm for any performance criterion that a user might set.

References

[1] Abramowitz, M. and Stegun, I.A. (1972), "Handbook of Mathematical Functions," National Bureau of Standards Applied Mathematics Series 55, Tenth Printing.

[2] Förstner, W. (1989), "Robust Statistical Methods for Computer Vision," Tutorial notes.

[3] Haralick, R.M., Joo, H., Lee, C.N., Zhuang, X., Vaidya, V., Kim, M.B. (1989), "Pose Estimation from Corresponding Point Data," IEEE Transactions on Systems, Man, and Cybernetics, Vol. 19, No. 6, pp. 1426-1446.

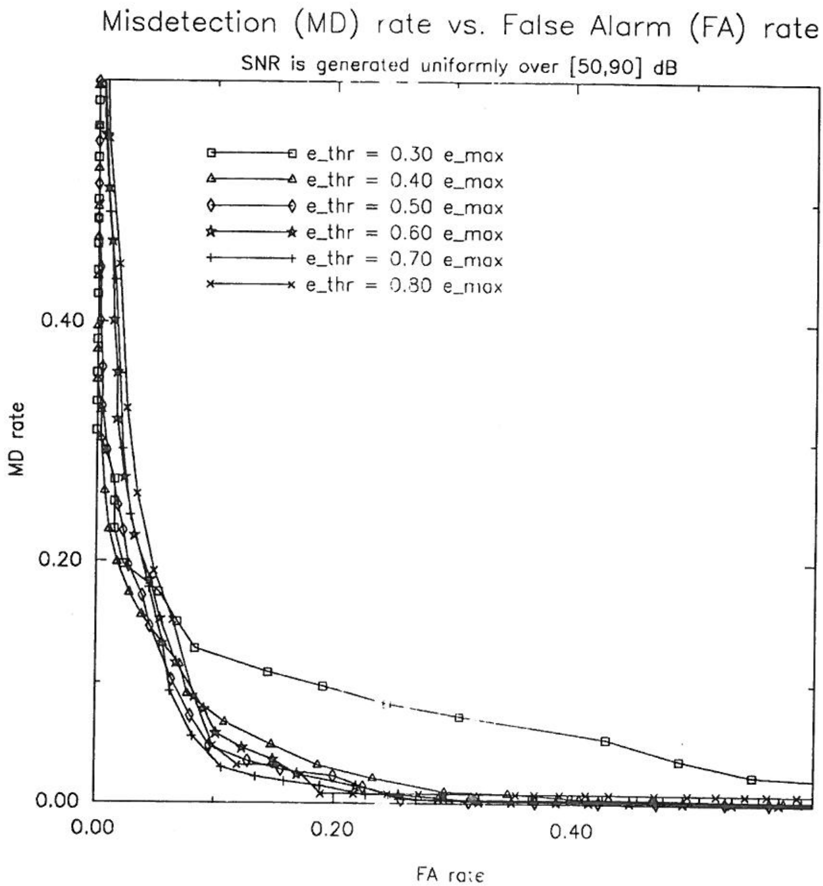

Figure 1: The misdetection (MD) rate versus the false alarm (FA) rate for different $e_{\mathbf{thr}}$.

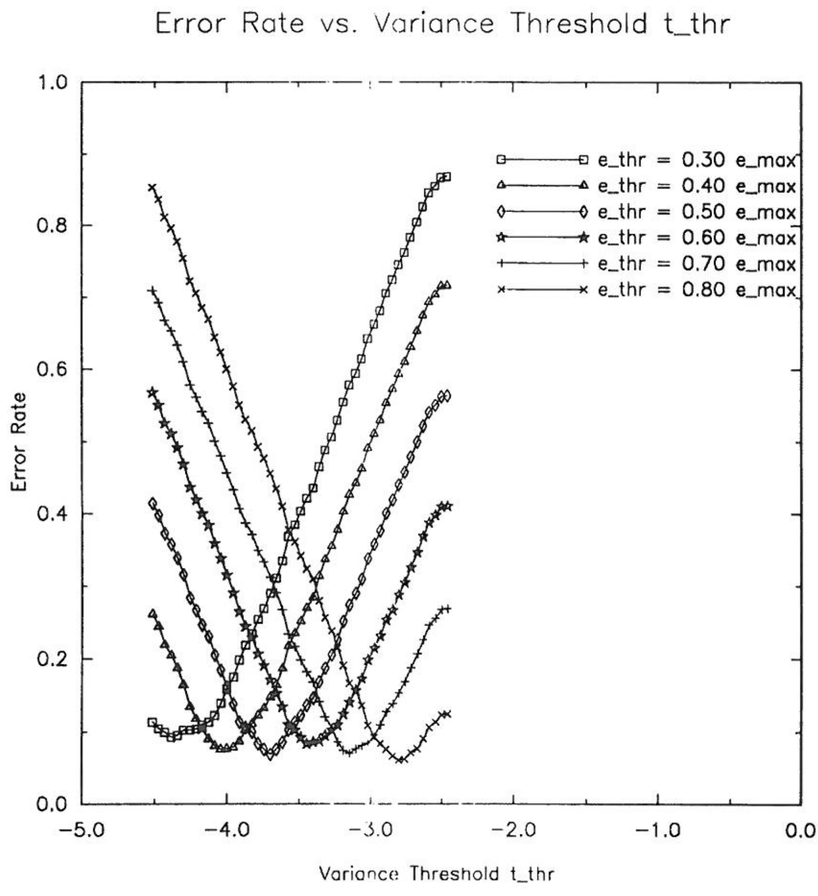

Figure 2: The error rate versus the variance threshold t_{thr} for different e_{thr}.

[4] Haralick, R.M., Lee, C.N., Ottenberg, K., Nöile, M. (1991), "Analysis and Solutions of the Three Point Perspective Pose Estimation Problem," IEEE Conference on Computer Vision and Pattern Recognition, pp. 592-598.

[5] Haralick, R.M. and Shapiro, L.G. (1991), Machine Vision, Addison-Wesley.

[6] Hinsken, L. (1988), "A Singularity Free Algorithm for Spatial Orientation of Bundles," International Archives of Photogrammetry and Remote Sensing, Vol. 27, Comm. B5, Kyoto.

[7] Koch, K.R. (1988), Parameter Estimation and Hypothesis Testing in Linear Models, Springer-Verlag.

[8] Rousseeuw, P.J. and van Zomeren, B.C. (1990), "Unmasking Multivariate Outliers and Leverage Points," Journal of the American Statistical Association, Vol. 85, No. 411.

[9] Seber, G.A.F. (1984), Multivariate Observations, Wiley, N.Y.