Application of Robust Estimation in Close-Range Photogrammetry

S. A. Veress
Department of Civil Engineering, University of Washington, Seattle, WA 98195
Huang Youcai
Department of Engineering Surveying, Wuhan Technical University of Mapping and Surveying, Wuhan, China

ABSTRACT: The least-squares estimator has been widely used in photogrammetric data processing. When gross errors are made in observations, the reliability of least-squares estimates is greatly affected. Techniques for detecting and locating gross errors have been rapidly developed in various fields. This paper uses a bisquare estimator as a robust estimator, which has been modified for detecting and removing gross errors.

INTRODUCTION

To increase the reliability of least-squares estimates, traditional data rejection in close-range photogrammetry is based on the evaluation of the magnitude of residuals. If the residuals are three times larger than the standard deviation, they are considered gross errors and are to be deleted from data processing. The least-squares estimator is very sensitive to gross errors, which may lead to some problems: (1) The standard deviation may not be truely representative of the accuracy of the majority of observations, and (2) gross errors may not necessarily be in the observations with large residuals. Therefore, if gross errors are not detected and located correctly, least-squares estimates may become falsified. Two methods may be suggested to correct these problems: (1) find an efficient method to detect and locate gross errors, or (2) design estimators which can delete gross errors automatically from the data processing. For the latter, robust estimators have been introduced in photogrammetry by several authors (Kubik *et al.*, 1986; Krarup, 1980; Stefanovic, 1980; Willke, 1969) as an effective way to eliminate gross errors.

According to Huber (1972), Hampel (1974), and Mosteller and Tukey (1977), a robust estimator must have two properties to be efficient: resistance and robustness. Resistance means that estimated parameters should change only slightly when a small part of the data is replaced by new nubmers (i.e., the estimates should limit the effect of any small change in the data). In this sense, small changes include gross errors in a small fraction of the data, and small errors in all the data. An estimator has robustness of efficiency over a range of distributions if its variance (its mean-squared error) is close to the minimum for each distribution. it may guarantee that the estimator is good when samples are from a distribution that is not known precisely.

The bisquare estimator, as a robust estimator, is used in this paper. Modifications have been made to achieve the best efficiency in the rejection of gross errors.

THEORETICAL DERIVATIONS OF THE BISQUARE ESTIMATOR AND ITS MODIFICATIONS

The advantages of the bisquare estimator, as compared to the least-squares estimator, have been summarized by Hoaglin *et al.* (1982). Gross (1972) gave several numerical examples for the evaluation of the bisquare estimator. For the least-squares estimator, the objective function is expressed as follows:

$$\rho(y;t) = (y-t)^2, \tag{1}$$

where y is an observation and t is the expected value of y. The estimated value of t can be obtained by minimizing the objective function Equation 1 with respect to t; then,

$$\sum_{i=1}^{n} \psi\,(y_i;t) = \sum_{i=1}^{n} (y_i-t) = 0, \tag{2}$$

where Equation 2 is a so-called ψ function. Finally, the estimated value of t can be obtained from

$$t_n = \sum_{i=1}^{n} y_i\,/n. \tag{3}$$

The sensitivity of t_n to gross errors may be illustrated by the curve (a straight line in this case) of the ψ function in Figure 1. It is apparent that the ψ function of the least-squares estimator has no bounds. That is the reason gross errors can completely upset the values of the least-squares estimates.

The bisquare estimator is the solution t_n of the ψ function; i.e.,

$$\sum_{i=1}^{n} \psi(\mathrm{u_i}) = 0, \tag{4}$$

$$\begin{aligned} \psi(\mu) &= \mu(1-\mu^2)^2, \qquad |\mu| < 1 \\ \psi(\mu) &= 0, \qquad |\mu| > 1 \end{aligned} \tag{5}$$

where

$$u_i = \frac{y_i - t_n}{K.S_n} = \frac{r_i}{K.S_n}, \tag{6}$$

$$S_n = \text{median}|y_i - t_n|, \text{ and} \tag{7}$$

where Equation 4 is called the bisquare estimator. S_n is the median absolute residual, which is less sensitive to gross errors than the standard deviation. r_i is the residual of observation i.

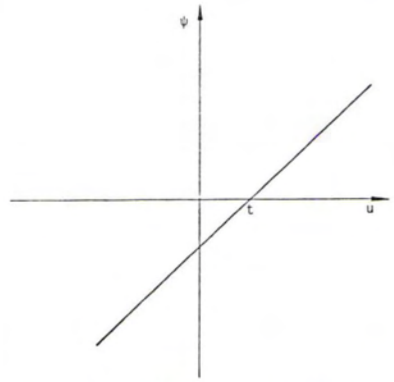

FIG. 1. ψ function of least-squares estimator.

0099-1112/87/5302–171$02.25/0

μ_i is a studentized residual. When the standard deviation is replaced by $k.S_n$, K is called the tuning constant, which depends on the types of data sets. Theoretically, if there is no gross error, μ_i should be smaller than 1. The resistance of efficiency of the bisquare estimator can be illustrated by the ψ function in Figure 2. The curve of the ψ function is bound by $\psi(\mu_i)=0$, $|\mu| > 1$.

Setting $w_i=(1-\mu_i^2)^2$, w_i is the weight of observation i, and the ψ function of the bisquare estimator can be rewritten as

$$\sum_{i=1}^{n} w_i \mu_i = 0, \tag{8}$$

or, in a matrix form,

$$\begin{aligned} \mathbf{A}^T\mathbf{WAX} - \mathbf{A}^T\mathbf{WE} &= 0, \\ \mathbf{X} &= (\mathbf{A}^T\mathbf{WA})^{-1}\mathbf{A}^T\mathbf{WE}, \end{aligned} \tag{9}$$

where $w_i=0$ when $|\mu_i| > 1$, $\mathbf{E}$ is a constant vector, $\mathbf{A}$ is a coefficient matrix, and $\mathbf{X}$ is a matrix of estimated parameters.

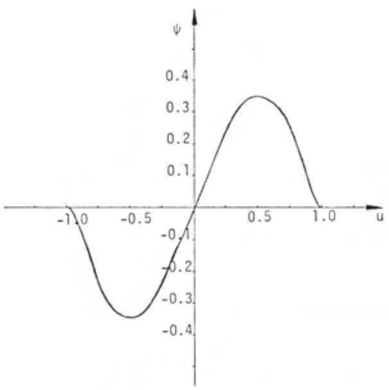

FIG. 2. ψ function of bisquare estimator.

Comparing Equation 8 with Equation 2, the bisquare estimator actually is the weighted least-squares estimator. Therefore, it belongs to the maximum likelihood estmation family and has the robustness of efficiency due to its minimum variance of the weighted observations.

Equation 8 cannot be solved directly because μ_i is unknown. An iteration approach is required to solve this problem. The initial values of μ are computed by using the residuals of the least-squares estimated observations. The iteration will be terminated after some criteria are satisfied. The iteration procedure can be summarized as follows:

(1) Compute initial values of w_i, $i=1,\ldots,n$. n is the number of observations. t_n, the least-squares estimate of n observations, is used for calculating μ_i and w_i.
(2) Use a weighted least-squares adjustment program (Equation 9) to compute a new set of w_i, $i=1,\ldots,n$.
(3) Loop back to (1). Iteration will be terminated after some criteria are satisfied. The criteria will be mentioned later on.

It is obvious that the criterion for rejecting gross errors in the bisquare estimation is based on the evaluation of the magnitude of residuals. As mentioned previously, gross errors may not show up in the large residuals. Huber (1981) gave a detailed discussion of this subject. After least-squares adjustment, the residual r can be expressed as

$$\begin{aligned} r_i &= y_i - \hat{y}_i \\ &= (1 - h_i)y_i - \sum h_{ik}y_k, \ (k \neq i) \end{aligned} \tag{10}$$

where h_i is a diagonal element of matrix $\mathbf{H}=\mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T$, which is called the hat matrix by Huber (1972); in matrix form

$$\begin{aligned} \mathbf{R} &= \mathbf{Y} - \hat{\mathbf{Y}} \\ &= \mathbf{Y} - \mathbf{AX} \\ &= \mathbf{Y} - \mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{Y} \\ &= \mathbf{Y}(\mathbf{I} - \mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T), \end{aligned} \tag{11}$$

where y_i or the $\mathbf{Y}$ matrix is an observation, and $\hat{y}_i$ or the $\hat{\mathbf{Y}}$ matrix is its least-squares estimate. $\mathbf{A}$ is a coefficient matrix of observation equations.

It is obvious that, if h_i is close to 1, a gross error in y_i will not necessarily show up in r_i. But it might show up elsewhere, say in y_k. h_k might be smaller than h_i. The criterion of rejecting gross errors based on the magnitude of residuals thus is not desirable. It may lead to some good observations being rejected and some bad observations being included in the data set. A point with relatively large h is called a leverage point.

APPLICATION OF THE BISQUARE ESTIMATOR IN SPACE RESECTION

A space resection is used to illustrate the efficiency of the bisquare estimator to detect and reject gross errors. For this illustration only, a large number of control points (21 in this experiment) was utilized even though such a sample is not practical in routine operation.

The collinearity equations were used in the space resection. They can be written as

$$\begin{aligned} x - x_0 &= -f\frac{a_1(X - X_c) + b_1(Y - Y_c) + c_1(Z - Z_c)}{a_3(X - X_c) + b_3(Y - Y_c) + c_3(Z - Z_c)} \\ y - y_0 &= -f\frac{a_2(X - X_c) + b_2(Y - Y_c) + c_2(Z - Z_c)}{a_3(X - X_c) + b_3(Y - Y_c) + c_3(Z - Z_c)}, \end{aligned} \tag{12}$$

where X, Y, and Z are the coordinates of control points; X_c, Y_c, and Z_c are the coordinates of the camera station; a_i, b_i, and c_i (i = 1, 2, 3) are the elements of the rotation matrix; f is the focal length of the camera; and x_0 and y_0 are the coordinates of the principal point.

The orientation martrix $\mathbf{M}$ is formed from the orientation angles ϕ, ω, and κ.

$$\begin{aligned} \mathbf{M} &= \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix} \\ a_1 &= \cos\phi\cos\kappa \\ a_2 &= -\cos\phi\sin\kappa \\ a_3 &= -\sin\phi \\ b_1 &= \cos\omega\sin\kappa - \sin\omega\sin\phi\cos\kappa \\ b_2 &= -\cos\phi\sin\kappa \\ b_3 &= -\sin\omega\cos\phi \\ c_1 &= \sin\omega\sin\kappa + \cos\omega\sin\phi\cos\kappa \\ c_2 &= \sin\omega\cos\kappa - \cos\omega\sin\phi\sin\kappa \\ c_3 &= \cos\omega\cos\phi \end{aligned} \tag{13}$$

The observation equations, in the general form, can be expressed as

$$\begin{aligned} v_{xi} &= \frac{\partial F_1}{\partial X_c}\Delta X_c + \frac{\partial F_1}{\partial Y_c}\Delta Y_c + \frac{\partial F_1}{\partial Z_c}\Delta Z_c + \frac{\partial F_1}{\partial \phi}\Delta\phi + \frac{\partial F_1}{\partial \omega}\Delta\omega \\ &\quad + \frac{\partial F_1}{\partial \kappa}\Delta\kappa + F_1(X_{c0}, Y_{c0}, Z_{c0}, \phi_0, \omega_0, \kappa_0) - x_i, \\ v_{yi} &= \frac{\partial F_2}{\partial X_c}\Delta X_c + \frac{\partial F_2}{\partial Y_c}\Delta Y_c + \frac{\partial F_2}{\partial Z_c}\Delta Z_c + \frac{\partial F_2}{\partial \phi}\Delta\phi + \frac{\partial F_2}{\partial \omega}\Delta\omega \\ &\quad + \frac{\partial F_2}{\partial \kappa}\Delta\kappa + F_2(X_{c0}, Y_{c0}, Z_{c0}, \phi_0, \omega_0, \kappa_0) - y_i, \end{aligned} \tag{14}$$

or in matrix form,

$$\mathbf{V}^{T} = |v_{x1}, v_{y1}, \ldots, v_{xn}, v_{yn}|$$

$$\mathbf{X}^{T} = |\partial X_c, \partial Y_c, \partial Z_c, \partial \phi, \partial \omega, \partial \kappa|$$

$$\mathbf{A} = \begin{bmatrix} \frac{\partial F_{12}}{\partial X_c}, & \frac{\partial F_{12}}{\partial Y_c}, & \frac{\partial F_{12}}{\partial Z_c}, & \frac{\partial F_{12}}{\partial \phi}, & \frac{\partial F_{12}}{\partial \omega}, & \frac{\partial F_{12}}{\partial k} \\ \frac{\partial F_{22}}{\partial X_c}, & \frac{\partial F_{22}}{\partial Y_c}, & \frac{\partial F_{22}}{\partial Z_c}, & \frac{\partial F_{22}}{\partial \phi}, & \frac{\partial F_{22}}{\partial \omega}, & \frac{\partial F_{22}}{\partial k} \\ \frac{\partial F_{1n}}{\partial X_c} & \cdot & \cdot & \cdot & \cdot & \cdot \\ \frac{\partial F_{2n}}{\partial X_c} & \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \tag{15}$$

$$\mathbf{E} = \begin{bmatrix} x_1 - F_{11}(X_{c0}, Y_{c0}, Z_{c0}, \phi_0, \omega_0, \kappa_0) \\ y_1 - F_{12}(X_{c0}, Y_{c0}, Z_{c0}, \phi_0, \omega_0, \kappa_0) \\ \cdot \\ \cdot \\ \cdot \\ x_n - F_{1n}(X_{c0}, Y_{c0}, Z_{c0}, \phi_0, \omega_0, \kappa_0) \\ y_n - F_{2n}(X_{c0}, Y_{c0}, Z_{c0}, \phi_0, \omega_0, \kappa_0) \end{bmatrix}$$

where n is the number of control points; X_{c0}, Y_{c0}, and Z_{c0} are the approximate values of coordinates of the camera station; ϕ_0, ω_0, and κ_0 are the approximate values of the rotation angles; and v_i is equal to $-r_i$. The bisquare estimates of parameter $\mathbf{X}$ are expressed in Equation 9.

In the space resection, a leverage point might occur in two possible cases: (1) The distribution of control points is not homogeneous; for example, one control point is located far from the others and the value h corresponding to this point might be very large; and (2) blunders in observations, mistakes in recording and typing, or gross errors of other sources may not be detected and eliminated from the data set. The problem may be resolved by a transformation of the residuals

$$\bar{r}_i = \frac{r_i}{1 - h_i} \tag{16}$$

where $\bar{r}_i$ is a transformed residual of residual r_i. The gross errors can be detected by checking $\bar{r}_i$s because they have been removed from the residuals. To improve the reliability of rejection of gross errors in the bisquare estimation, the r_i should be replaced by $\bar{r}_i$ in Equation 6. The modified bisquare estimator is written as follow:

$$\bar{u}_i = \frac{\bar{r}_i}{KS_n} \tag{17}$$

ANALYSIS OF RESULTS OF THE EXPERIMENTS

The following data are from an actual project. Twenty-one control points were imaged and can be used for space resection. The reliability of those points has been checked by a complete bundle adjustment. Actual situations can be simulated by adding different kinds of errors to the data set, i.e., gross errors of various magnitude in control points, in photocoordinates, and in the original data set. They are listed in Table 1. The values of coordinates of camera station and rotational angles computed by the least-squares estimator and the bisquare estimator are shown in Table 2.

Table 2 shows that points 4, 5, and 12 are the common points rejected by the bisquare estimator. Checking Figure 3, it can be seen that the location of those three points is far from the majority of the other points. It indicates that the geometric positions of the three points may not be desirable.

Figure 4 shows the number of iterations of the modified bisquare estimator in the four cases.

Reviewing Table 2 and Figure 4, the modified bisquare estimator have a very good convergency even though gross errors occur in the coordinates of control points. The required number of iterations is approximately nine. The estimates of median absolute deviation by the bisquare estimator in the four cases are almost the same. That is why the convergency of the bisquare estimator may occur so rapidly.

Table 3 shows further research regarding the residuals and

TABLE 1. OBSERVATIONS AND CONTROL DATA IN DIFFERENT SITUATIONS

	Original data set (case 1)					Gross errors in coord. of control points (case 2)			Gross errors in photocoord. and control coord. remain the same (case 3)		Gross errors in control coord. and photocoord. remain the same (case 4)		
	Photocoordinates		Control coordinates										
No	x (mm)	y (mm)	X (m)	Y (m)	Z (m)	X (m)	Y	Z (m)	x (mm)	y (mm)	X (m)	Y (m)	Z (m)
1	34.512	−1.597	1429.63	1029.34	1647.65								
2	91.433	−3.058	1495.51	1026.66	1660.28								
3	29.425	52.249	1432.49	1096.63	1762.66								
4	17.042	42.451	1415.11	1082.61	1738.48								
5	32.432	47.270	1435.39	1089.18	1751.24								
6	37.300	−28.088	1428.50	1003.06	1598.49								
7	51.048	−26.926	1443.14	1003.87	1601.90								
8	53.113	−31.096	1445.25	999.50	1601.70								
9	64.941	−25.918	1458.09	1004.54	1605.19	(6000.)			(10.)		(50.)		
10	64.967	−27.718	1458.06	1002.67	1604.76	7458.06			54.967		1408.06		
11	−47.073	30.284	1336.72	1066.86	1692.05								
12	53.820	37.800	1328.32	1076.33	1699.99								
13	65.040	−29.513	1458.09	1000.79	1604.49								
14	65.036	−31.235	1458.04	999.00	1604.25								
15	23.153	−29.275	1413.59	1002.19	1595.36								
16	22.986	−32.753	1413.35	998.62	1594.86								
17	37.374	−29.951	1428.52	1001.11	1598.33								
18	37.339	−31.737	1428.44	999.26	1598.04								
19	91.792	−23.735	1487.49	1006.10	1611.75								
20	−18.994	−32.592	1370.18	1000.01	1585.28			(7000.)		(20.)			(90.)
21	37.588	−33.704	1428.67	997.23	1597.74			8597.74		−13.704			−1507.74

Focal length = 614.055 mm

TABLE 2. COMPUTED VALUES OF COORDINATES OF CAMERA STATION AND ROTATION ANGLES

	Least-squares estimator						
	Coord of camera station			Rotational angles (o)			dev.
case	X (m)	Y (m)	Z (m)	Tilt	Swing	Azimuth	S (m)
1	1376.85	1046.98	963.40	1 46.5	316 22.9	136 58.2	0.019
2	897.21	1006.92	1132.10	43 18.5	278 10.4	87 42.8	7.590
3	1379.68	1033.66	967.82	1 03.7	284 50.9	105 07.5	0.729
4	1391.52	1017.64	1004.42	1 14.6	178 01.4	359 30.5	1.003
	Bisquare estimator						
1	1376.06	1047.00	963.35	1 48.8	315 19.7	135 53.3	0.010
2	1376.74	1046.47	963.10	1 44.0	315 35.3	136 10.3	0.012
3	1376.03	1046.89	963.36	1 48.4	315 03.2	135 36.6	0.012
4	1376.74	1046.47	963.10	1 44.0	315 35.3	136 10.3	0.011
	Point number rejected by bisquare estimator						
1				2, 3, 4, 5, 12			
2				4, 5, 10, 11, 12, 21			
3				3, 4, 5, 10, 12, 21			
4				4, 5, 10, 11, 12, 21			

their weights for case 2. One can see a desirable property of the bisquare estimator; smaller residuals have larger weights. This property can strengthen the reliability of estimates. In order to remove gross errors completely, another modification was made in programming, i.e., if an observation x_i or y_i is rejected, both x_i and y_i should be rejected because a large error in the ith control point might not necessarily show up both in residual r_x and r_y. Becaue r_x and r_y have strong correlations with the corresponding control point, the effect of the mistaken control point can be completely removed only by rejecting both r_x and r_y. This modification is only for data processing in photogrammetry.

The criteria used to terminate the iterations in this research were that (1) the differences of computed coordinates of camera station between two iterations should be smaller than 0.001 metre and the differences of rotational angles should be smaller than 0.01 minute; and (2) the number of iterations is smaller than 20.

Three values of k were used in this work ($k=3$, 6, and 9). The smallest value of k (3) is so small that many useful points were rejected. The value of $k=9$ is so large that gross errors were not deleted completely. The value of $k=6$ seems to be the best for this work. The criterion for choosing a value for k is to keep as many observations as possible after deleting gross errors. By evaluating the "hat" matrix, it can be found that the magnitude of h is reduced when the number of observations is increased. It indicates that h_i is a measure of dependence of estimated parameters on point i. In other words, the value of h can show the influence of an observation on the estimated parameters. A desirable situation is that the differences between h's are not significant and the magnitude of h's are much smaller than 1. Increasing the number of observations is a way to reduce the dependence of estimated parameters on individual observation.

CONCLUSION

There are many redundant observations to be chosen for the space resection. It is difficult to detect and locate gross errors correctly by the traditional method. The least-squares estimator will yield unreliable estimates if gross errors exist in the observations. The bisquare estimator may be a powerful tool to use to solve these problems.

The advantages of the bisquare estimator can be summarized as follows:

(1) The reliability of estimates can be guaranteed by completely deleting gross errors in the observations.

(2) The reliability of the estimates can be strengthened by reweighting the observations in data processing.

(3) The stability of geometry may be improved by rejecting a

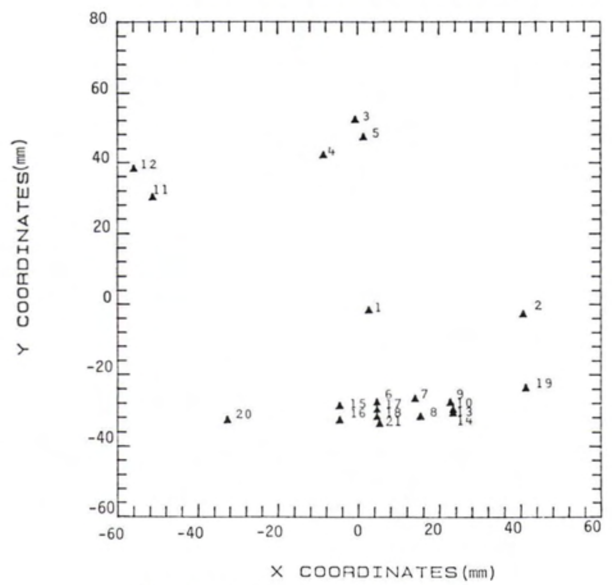

FIG. 3. Distribution of image points.

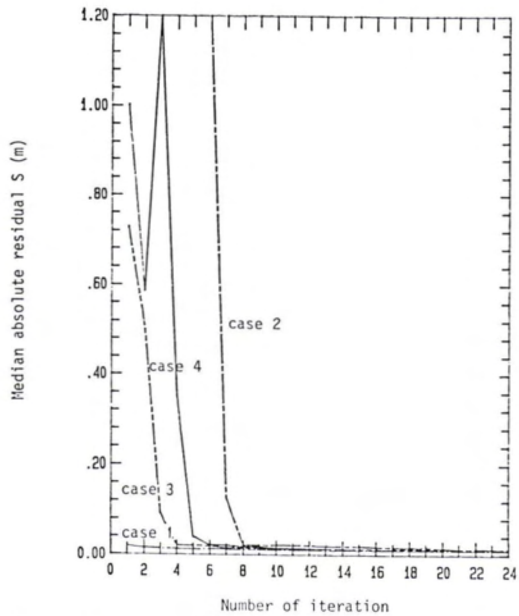

FIG. 4. Convergency of modified bisquare estimator.

TABLE 3. RESIDUALS OF OBSERVATIONS AND THEIR WEIGHTS IN CASE 2

Point	r (m)	w	r (m)	w
1	−0.007	0.978	−0.009	0.962
2	0.021	0.810	0.012	0.941
3	−0.005	0.989	0.000	1.000
4	*	0.000	*	0.000
5	*	0.000	*	0.000
6	−0.024	0.762	−0.024	0.760
7	−0.007	0.976	−0.015	0.906
8	−0.016	0.889	−0.010	0.958
9	0.003	0.995	−0.011	0.994
10	*	0.000	*	0.000
11	*	0.000	*	0.000
12	*	0.000	*	0.000
13	−0.007	0.975	0.006	0.984
14	−0.011	0.949	0.013	0.928
15	0.004	0.993	0.002	0.998
16	0.011	0.948	0.015	0.899
17	0.008	0.971	0.005	0.991
18	0.011	0.950	0.015	0.902
19	0.012	0.937	−0.006	0.986
20	0.006	0.983	0.003	0.996
21	*	0.000	*	0.000

*means rejected gross errors. Iteration number = 20.

few points with large values of h because some large values of h may be due to the undesirable locations of the points, which can lead to overriding influences on the fit. The modified bisquare estimator can reveal the leverage point hiding behind a cloud. The bisquare estimator has some disadvantages. It is time-consuming because an iteration procedure is required. But not many control points are required for the resection in close-range photogrammetry and the speed of convergency is fast, as shown in Figure 1 (only nine iterations are required in the worse situation, case 2). The problem of time consumption can be neglegible.

REFERENCES

Behnken, D. W., and N. R. Draper, 1972. Residuals and Their Variance Patterns, *Technommetrics* 14, pp. 101–111.

Gross, A. M., 1972. Confidence Intervals for Bisquare Regression Estimates, *Journal of the American Statistical Association*, 72, pp. 341–353.

Hoaglin, C., F. Mosteller, and J. W. Tukey, 1983. *Understanding Robust and Exploratory Data Analysis*, John Wiley & Sons, Inc., pp. 348–349.

Huber, P. J., 1981. *Robust Statistics*, New York: John Wiley and Sons, Inc., pp. 153–162.

Krarup, T., 1980. Gotterdammerung Over Least Squares Adjustment, *International Archives of Photogrammetry*, Vol. 23 (B3), pp 369–378.

Kubik, K., D. Merchant, and T. Schenk, 1986. Grosserrors and Robust Estimation, *Technical papers, 1986 ACSM-ASPRS Annual Convention*, Volume 4, Photogrammetry, pp. 250–261.

Mosteller, F., and J. W. Tukey, 1977. *Data Analysis and Regression*, Addison-Wesley Publishing Company.

Stefanovic, P., 1980. Pitfalls in Blunder Detection Techniques, *International Archives of Photogrammetry*, Vol. 23 (B3), pp. 687–700.

Willke, T. A., 1969. Committee on Computational Photogrammetry, 1969 Symposium on Computational Photogrametry, 21–23, Jan., The American Society of Photogrammetry.

(Received 3 July 1986; accepted 3 September 1986; revised 18 September 1986)