Important Copyright Notice:

The provision of this paper in an electronic form in this site is only for scholarly study purposes and any other use of this material is prohibited. What appears here is a near-publication draft of the final paper as appeared in the journal or conference proceedings. This is subject to the copyrights of the publishers. Please observe their copyrights.
Fast Estimation of Epipolar Geometry Using High Breakdown M-estimators

Reza Hoseinnezhad
Swinburne University of Technology
Faculty of Engineering and Industrial Sciences
John St., Hawthorn, Victoria 3122, Australia
rhoseinnezhad@swin.edu.au

Alireza Bab-Hadiashar
Swinburne University of Technology
Faculty of Engineering and Industrial Sciences
John St., Hawthorn, Victoria 3122, Australia
abab-hadiashar@swin.edu.au

Abstract
The high breakdown M-estimator (HBM) is introduced in this paper as an outstanding choice compared to modern high breakdown estimators for epipolar geometry estimation and motion segmentation problems. It is mathematically demonstrated that since HBM automatically guides its cost optimization using an iterative reweighted least square regression (instead of using random sampling like high breakdown RANSAC-based techniques), its computational cost is substantially cheaper than that of modern high breakdown estimators. In a number of experiments involving both synthetic and real image pairs, the performance of HBM, MSSE and pbM-estimator to solve fundamental matrix estimation problems are compared. The results verify that in terms of computational cost, HBM significantly outperforms the modern RANSAC-based high breakdown estimator and the pbM-estimator, while the estimation accuracies of HBM and its tolerances to high fractions of gross or pseudo outliers are similar to the other estimators. For real-time epipolar geometry estimation and multiple motion segmentation problems that involve multi-structured data segmentation (large fractions of outliers), application of the HBM-estimator for fast computation is highly recommended.

1 Introduction
The fundamental matrix provides a constraint between corresponding points in two views. More precisely, a sequence of images can be analyzed by looking at the relationship between corresponding points of pairs of images and the fundamental matrix provides a constraint which establishes such a relationship. Estimation of fundamental matrix is important for several problems such as matching, recovering of structure, and motion segmentation [12]. Consider the movement of a set of point image projections from an object which undergoes a rotation and non-zero translation between views. After the motion, the set of homogeneous image points \( \{x_i\}, i = 1, \ldots, n \), as viewed in the first image is transformed to the set \( \{x'_i\} \) in the second image, positions related by:

\[
x'_i \top F x_i = 0
\]

where \( x_i = [x_i, y_i, 1] \top \) is a homogeneous image coordinate and

\[
F = \begin{bmatrix}
f_1 & f_2 & f_3 \\
f_4 & f_5 & f_6 \\
f_7 & f_8 & f_9
\end{bmatrix}
\]

is the fundamental matrix and has seven degrees of freedom, as only the ratio of parameters is significant and also the matrix has zero determinant. Equation (1) can be written in the following linear form:

\[
f_1 x'_i x_i + f_2 x'_i y_i + f_3 x'_i + f_4 y'_i x_i + f_5 y'_i y_i + f_6 y'_i + f_7 x_i + f_8 y_i + f_9 = 0
\]

which seems to be an eight-dimensional archetype for solution by robust regression techniques.

In the computation of fundamental matrix, however, outliers typically arise from gross errors such as correspondence mismatches or the inclusion of movements inconsistent with the majority. The latter might be caused by features being on occluding contours, shadows or independently moving objects. Therefore, robust estimators are commonly used to estimate the epipolar geometry from image pairs in computer vision applications. Specially in cases involving multiple independently moving objects, each movement can include a small portion of total image pairs and only high breakdown robust estimators can segment the motions.

Since the introduction of RANSAC [2], a quarter of century ago, several high breakdown robust estimators have been specially designed to solve computer vision problems and used to estimate epipolar geometry (such as RESC [14], MSSE [1], ASSC [13], MSAC [12], and Projection based...
M-estimators (also called pbM) [10, 11]. All such estimators include an optimization scheme to search the parameter space and find the parameter values which optimize the objective function of the estimator. Almost all high breakdown robust estimators (except pbM) use random sampling to optimize their objective functions.

Random sampling is a random search scheme in the sample space for the best elemental subset (or $p$-tuples) that results in the parameter estimates with the optimal objective function value. An elemental subset is a subset of $p$ data samples ($p$ is the dimension of the parameter space) that defines a full rank system of equations from which a model candidate can be computed. If $N$ elemental subsets are randomly selected, then with a probability of:

$$P_{\text{success}} = 1 - [1 - \epsilon^p]^N$$

at least one of them is a good elemental subset (i.e. all its samples belong to the inlier structure), where $\epsilon$ is the fraction of inliers samples. Thus, for a given success probability $P_{\text{success}}$, at least:

$$N = \frac{\log(1 - P_{\text{success}})}{\log(1 - \epsilon^p)}$$

elemental subsets should be randomly examined.

For cases involving multi-structural data, the above number of random $p$-tuples is substantial and the computational cost of segmentation is commonly too high for real-time (or near real-time) applications. This is because the fraction of inliers is usually unknown in advance and the ratio $\epsilon$ has to be taken as the smallest possible fraction of inliers in the application which would be very small if the application involves segmentation of multiple data structures. The plots shown in Figure 1 demonstrate that substantially large numbers of random $p$-tuples are required for small $\epsilon$ ratios, hence the segmentation is computationally expensive.

The computational cost of random sampling is specially expensive for fundamental matrix estimation, because the dimension of parameter space ($p = 8$ in case of using the eight-point algorithm [4]) is large and as Figure 1 shows, the number of random $p$-tuples in this case is substantially higher compared to two or three dimensional visual data segmentation problems.

Subbarao and Meer [11] have devised an alternative approach to random sampling (proposed as an optimization strategy for pbM). PbM employs a gradient-based optimization technique that searches over a Grassman manifold in the $p$-dimensional space of parameter vectors. Since the objective function of pbM estimator is not differentiable, simplifying assumptions are made and therefore, an approximate estimate of the conjugate gradient is calculated in each iteration of the search process. Due to the locality of the search process and the error of gradient calculation, usually the optimization procedure does not converge and requires to be repeated for numerous times (initialized with random elemental subsets). The number of initializations is specially substantial with the 8-dimensional data involved in fundamental matrix estimation problems.

We have devised a high breakdown estimator with a differentiable objective function that can be optimized through an iterative reweighted least square regression scheme. Since the redescending M-estimators employ similar iterative updating formulas for their search scheme, we have called the new technique: High Breakdown M-estimator or HBM for short. But the performance of HBM to segment multi-structured visual data is superior to M-estimators as the maximum breakdown point of the classical redescending M-estimators is 50% but HBM can tolerate a substantially larger fraction of gross or pseudo outliers. HBM was initially developed to solve range data segmentation problems and the results have been discussed in a paper accepted to be presented in ICCV’2007 [6]. Following the introduction of HBM and its application to 3D range data segmentation problems, we have applied this estimator to motion segmentation problems. HBM performs exclusively well in such problems as due to the high dimension of data ($p = 8$), substantial numbers of random $p$-tuples are required by RANSAC-based robust estimators. Specially in problems involving segmentation of multiple motions, the fraction of inliers can be small and cause an even higher cost of computation with RANSAC-based techniques.

This paper introduces the HBM as an outstanding choice compared to modern high breakdown estimators for epipolar geometry estimation and motion segmentation problems by presenting the results of using HBM to solve fundamental matrix estimation problems. The results of comparative experiments reported in this paper demonstrate that HBM

![Figure 1. Required number of elemental subsets plotted versus $\epsilon$ for different data dimensions.](image-url)
significantly outperforms the modern RANSAC-based high breakdown estimators and the pbM-estimator in terms of computational cost. This advantage is gained mainly because HBM does not use random sampling for its cost optimization and automatically guides its search using iterative reweighted least square regression.

2 A Review of High Breakdown M-estimator

Consider a vision problem involving segmentation of several data structures. From each structure, $n_i$ measurement samples denoted by $\{y_i; i = 1, \ldots, n_i\}$ are available and each sample $y_i \in \mathbb{R}^p$ is corrupted with independent and identically distributed (i.i.d.) noise:

$$ y_i = y_{io} + \delta y_i; \quad \delta y_i \sim GI(0, \sigma^2 I_p) \quad (6) $$

where $y_{io}$ is the true value of $y_i$. $GI(.)$ stands for a general symmetric distribution of independent measurement noise samples and $\sigma$ is the unknown scale of noise. Characterizing the distribution by its first two central moments in equation (6) implies that noise distribution is assumed to be Gaussian.

Each data structure is modeled by the following linear errors-in-variables (EIV) regression model:

$$ y_{io}^T \theta^* - \alpha = 0; \quad i = 1, \ldots, n_i \quad (7) $$

where $\theta^* \in \mathbb{R}^p$ and $\alpha$ are the model parameters yet to be estimated for each structure and the following constraints are imposed to eliminate the ambiguity of the model parameters being defined up to a multiplicative constant:

$$ ||\theta|| = 1; \quad \alpha \geq 0. \quad (8) $$

In addition, in a fundamental matrix estimation problem, the constraint of a zero-determinant fundamental matrix $F$ should also be imposed.

For a given parameter estimate $\hat{\Theta} = [\hat{\theta}^T, \hat{\alpha}]^T$, each data sample $y_i$ corresponds to an algebraic distance $r_i = y_i^T \hat{\theta} - \hat{\alpha}$. With traditional regression models, these distances are called residuals and we also use this popular term in this paper. In the least $k$-th order statistics (LkOS) estimator, the objective function is the $k$-th order statistics of the squared residuals:

$$ J_{\text{LkOS}}(\hat{\Theta}) = r_{k:n}^2 \quad (9) $$

where $n$ is the total number of available data samples. The order $k$ is given by $k = \lceil \epsilon n \rceil$ where $\epsilon$ is a presumed lowest possible fraction of inliers in the application. Provided there are moderate number of samples in the target structure, the breakdown point of LkOS estimator is $(1 - \epsilon) \times 100\%$. The objective function (9) is a non-differentiable function of $\hat{\Theta}$ and is usually minimized using random sampling.

In HBM, the objective function is chosen as the functional form of the $k$-th order statistics of the squared residuals. For a given parameter estimate $\hat{\Theta}$, the squared residuals $\{z_i = r_i^2; \quad i = 1, \ldots, n\}$ have a statistical distribution that can be estimated by the following kernel density estimator:

$$ f_{\hat{\Theta}}(z) = \frac{1}{nh} \sum_{i=1}^{n} K \left( \frac{z - z_i}{h} \right) \quad (10) $$

where $K(.)$ is a kernel function with the following properties:

$$ \int_{-\infty}^{+\infty} K(u) du = 1 \quad (11) $$

$$ K(u) = K(-u) \geq 0 \quad (12) $$

$$ K(u_1) \geq K(u_2) \quad \text{for} \quad |u_1| \leq |u_2| \quad (13) $$

and $h$ is the kernel bandwidth. The value of the bandwidth has a weak influence on the result of the M-estimation [10] and we use the following formulae, as suggested in [10,11], to calculate the kernel bandwidth using the median of absolute differences (MAD) estimate:

$$ h = n^{-\frac{1}{5}} \text{med}_{i=1}^{n} \{|z_i - \text{med}_{j=1}^{n} z_j|\} \quad (14) $$

The objective function of HBM is given by:

$$ J_{\text{HBM}}(\hat{\Theta}) = z_e = F_{\hat{\Theta}}^{-1}(\epsilon) \quad (15) $$

where $F_{\hat{\Theta}}^{-1}(\cdot)$ is the inverse cumulative distribution function (inverse CDF) of the squared residuals. The CDF of the squared residuals is the following differentiable function of $\hat{\Theta}$:

$$ F_{\hat{\Theta}}(z) = \frac{1}{nh} \sum_{i=1}^{n} \int_{-\infty}^{z} K \left( \frac{\xi - z_i}{h} \right) d\xi. \quad (16) $$

Therefore, the inverse CDF is also differentiable and minimized by solving:

$$ \left. \frac{\partial F_{\hat{\Theta}}^{-1}(\epsilon)}{\partial \hat{\Theta}} \right|_{\hat{\Theta} = \Theta^*} = \left. \frac{\partial z_e}{\partial \hat{\Theta}} \right|_{\hat{\Theta} = \Theta^*} = 0. \quad (17) $$

For any parameter estimate we have:

$$ \epsilon = F_{\hat{\Theta}}(J_{\text{LkOS}}) = F_{\hat{\Theta}}(z_e) \quad (18) $$

By differentiating both sides of (18), we derive:

$$ 0 = \frac{\partial F_{\hat{\Theta}}(z_e)}{\partial \hat{\Theta}} = \frac{1}{nh} \sum_{i=1}^{n} \left( \frac{\partial}{\partial \hat{\Theta}} \int_{-\infty}^{z_e} K \left( \frac{\xi - z_i}{h} \right) d\xi \right) = \frac{1}{nh} \sum_{i=1}^{n} \left( \frac{\partial z_i}{\partial \hat{\Theta}} K \left( \frac{z_i}{h} \right) + \int_{-\infty}^{z_e} \frac{\partial}{\partial \hat{\Theta}} K \left( \frac{\xi - z_i}{h} \right) d\xi \right) \quad (19) $$
To minimize the objective function, the condition (17) should be satisfied. Thus, in the above equation the term \( \frac{\partial z_i}{\partial \Theta} \) is zero:

\[
0 = \frac{1}{nh^2} \sum_{i=1}^{n} \int_{-\infty}^{\infty} \frac{\partial}{\partial \Theta} K \left( \frac{z_i - z}{h} \right) \, d\xi
= \frac{1}{nh^2} \sum_{i=1}^{n} \int_{-\infty}^{\infty} \frac{\partial}{\partial z} \frac{\partial}{\partial \Theta} K \left( \frac{z_i - z}{h} \right) \, d\xi
\]

and therefore:

\[
\sum_{i=1}^{n} \frac{1}{h^2} K \left( \frac{z_i - z}{h} \right) - K(-\infty) \frac{\partial z_i}{\partial \Theta} = 0. \tag{20}
\]

The dependence of the bandwidth on the parameter estimate has been ignored in the above derivations as the bandwidth given by equation (14) does not substantially vary with \( \Theta \) (and \( \frac{\partial h}{\partial \Theta} \) is small) and the size of the bandwidth (and therefore its variations) do not substantially affect the performance of the estimator [10,11]. From the kernel properties (11)-(13) we have \( K(-\infty) = 0 \).

Since \( z_i = r_i^2 \) the term \( \frac{\partial z_i}{\partial \Theta} \) can be replaced with \( 2r_i \frac{\partial r_i}{\partial \Theta} \) and the following equation is derived:

\[
\sum_{i=1}^{n} \frac{1}{h^2} K \left( \frac{z_i - z}{h} \right) r_i \frac{\partial r_i}{\partial \Theta} = 0. \tag{22}
\]

Like a similar equation for redescending M-estimators [8], equation (22) can be iteratively solved by updating the parameters through iterative reweighted least squares regression on the data with the following weights:

\[
w_i = \frac{1}{h^2} K \left( \frac{z_i - r_i^2}{h} \right). \tag{23}
\]

To facilitate the computation of the weights in equation (23), the functional form of the \( k \)-th order statistics, \( z_k \), can be approximated with its sample value (provided there are adequate number of data samples):

\[
w_i = \frac{1}{h^2} K \left( \frac{r_{k,m}^2 - r_i^2}{h} \right) = \frac{1}{h^2} K \left( \frac{r_i^2 - r_{k,n}^2}{h} \right). \tag{24}
\]
However, as it was shown in the previous paragraphs, the dependence of $\rho(.)$ on $r_{k,n}^2$, is the direct result of iterative minimization of the functional form of $r_{k,n}^2$, which in turn allows HBM to have a high breakdown point.

In practice, since the $\rho(.)$ function assigns relatively large weights to residuals larger than $\sqrt{r_{k,n}^2}$, these weights have to be cared in order to stop the estimator returning a bridging fit where there are close structures in data. In our implementation, such weights are set to zero to remove their influence and minimize the computational cost.

Fig. 3 shows the optimization algorithm of HBM objective function. Like the redescending M-estimators, the proposed iterative reweighted least squares procedure in steps 4-10 minimizes the term $\sum_i w_i r_i^2$ using orthogonal least squares regression in each iteration. However, this minimization can be trapped in a local minimum of the objective function and should be repeated with different random initializations. However, the required number of repetitions, $N_{\text{init.}}$, can be far less than the number of random elemental subsets given by equation (5), particularly in applications involving numerous data structures (implying a small fraction of inliers and requiring large number of random samples, as discussed before).

A similar process (involving multiple random initializations) is also implemented in the pbM estimator with conjugate gradient search [11]. However, as it was mentioned before, the objective function of pbM is not differentiable and in formulation of the first order conjugate gradient method (searching over the Grassman manifold: $\theta^\top \nabla \theta = 1$), the dependence of $\alpha$ on $\theta$ is ignored. Because of this negligence, in many iterations the optimization procedure either diverges or converges to a local minimum and therefore, it needs to be repeated for a large number of times. Our experiments verify this argument, showing that the minimum number of initializations required by pbM is several times larger than HBM (see Section 3).

It is important to note that the algorithm shown in Figure 3 is only the optimization part of HBM and should be followed by a segmentation algorithm. In our experiments, we have used the MSSE segmentation algorithm [1] because of its low computational cost, high level of consistency and small bias in applications involving close data structures [5, 7].

### 3 Experimental Results

This section presents the results of fundamental matrix estimations by HBM and two modern high breakdown estimators in a number of experiments involving both synthetic and real image pairs. In each experiment, the set of homogenous image points in the first and second images are denoted by $\{x_i = [x_i, y_i, 1]^\top\}$ and $\{x'_i = [x'_i, y'_i, 1]^\top\}$ and related by equation (1) and its linear form (3).

We employ the Hartley’s 8-point algorithm [4] to solve for candidate fits of the $f_i$ parameters in the linear equation (3) to the data. The Sampson/Weng distances [12]:

$$d_i = \frac{x_i^\top F x_i}{\sqrt{\left( \frac{\partial f}{\partial x_i} \right)^2 + \left( \frac{\partial f}{\partial y_i} \right)^2}}$$

are used as residuals ($r_i$’s in the formulations presented in Section 2 and in HBM algorithm presented in Figure 3) for different robust estimation algorithms examined in this paper.

Our focus is on three important features of the robust estimators: estimation error, breakdown point and computation time. Among RANSAC-based techniques, we have chosen MSSE [1] which employs the LkOS estimator for its optimization part and has its specific method for segmentation. Our previous studies have shown that besides a high breakdown point (because of similarity to LkOS), MSSE outperforms other modern high breakdown robust estimators in terms of asymptotic and finite sample bias [5, 7]. Furthermore, we recall from Section 2 that the optimization algorithm implemented in HBM is inspired by LkOS. Therefore, in comparison of HBM with MSSE, the main focus will be on convergence speed of the two optimization algorithms: one based on random sampling, and the other using a guided search for the minimum of objective functions in parameter space, with few random initializations. We have also compared the performance of HBM with the pbM-estimator which uses conjugate gradient search for optimization [11].

In our first experiment, we have replicated the Torr and Murray’s experiment involving estimation of fundamental matrix using synthetic data [12]. This experiment was developed to compare the performance of the art techniques (in 1997) and we follow the same path to compare the pbM (as the most recent robust estimator in the current literature) and MSSE (demonstrated to have a high breakdown point, higher consistency and smaller bias [5, 7]) with HBM in terms of their breakdown point, accuracy of outlier detection and computation time.

The data points are randomly generated in the region of $\mathbb{R}^3$ visible to two positions of a synthetic camera having intrinsic parameters

$$C = \begin{bmatrix} 1.00 & 0.00 & 0.36 \\ 0.00 & 1.50 & 0.36 \\ 0.00 & 0.00 & 0.0014 \end{bmatrix}$$

equivalent to an aspect ratio of 1.5, an optic center at the image center $[256 \ 256]^\top$, and a focal length of $f = 703$ (pixels), providing a field of view of $40^\circ$. The projection of a point $X$ in the first position is $x = C[I|0]X$ and in the second is $x' = C[R|t]X$ where the camera makes a
rotation $[R]$ and translation $t$. The motion is random and different in each test.

The experiment has been repeated for a variety of fractions of outliers ranging from 5% to 70%. For each fraction of outliers, 100 synthetic datasets (with random motions) were examined each including 2000 corresponding points ($\lceil 2000\epsilon \rceil$ true matches and $\lceil 2000(1 - \epsilon) \rceil$ outliers where $1 - \epsilon$ is the fraction of outliers). Each image point was perturbed by Gaussian noise with a standard deviation of 1.0 (pixels). Figure 4 shows a typical set of point correspondences arising from some arbitrary random motion. The solid and dashed connecting lines show the inlier and outlier correspondences, respectively - the true dichotomy is shown in (a) and the estimated one in (b). The point symbols ◦ and □ show the points $(x, y)$ and $(x', y')$, respectively. For the purpose of clarity, only 40 inlier matches and 10 outlier matches are displayed.

In each of the 100 trials (for each fraction of outliers), the fundamental matrix has been estimated using MSSE, pbM and HBM and with each estimate, the standard deviation of the Sampson/Weng distances (of actual noise free projections of the synthetic world points to the estimated epipolar lines) were calculated, and the average of all 100 standard deviations were recorded. Besides, the average of computation times (over 100 trials for each fraction of outliers) were calculated and recorded. The standard deviation records are presented in Figure 5. As it is expected from the examined high breakdown estimators, the standard deviations are small for all fractions of outliers (up to 70%) and HBM has a slightly larger estimation error. However, in the worst case, the standard deviation is 2.55 pixels which is well below the “four pixels” criterion used by Torr and Murray [12] to rule out the performance of inaccurate robust estimators in the test.

In Figure 6(a), the average computation times are plotted versus fractions of outliers for each estimator. For MSSE, the “random sampling”-based optimization technique consumes a substantially long computation time for high fractions of outliers. This is expected, as the minimum number of required random $p$-tuples, given by equation (5), significantly increases for small $\epsilon$ (large fractions of outliers) with $p = 8$. On the other hand, a close-up of the plots, presented in Figure 6(b), shows that for small fractions of outliers (up to 45%), computation time of MSSE is shorter than pbM and slightly shorter than HBM, and HBM is always faster than pbM. Therefore, in the applications with more than 45% outliers (which commonly occur in practice), HBM is advantageous in terms of computation.

The three robust estimators have been also applied on real image sequences. The first case shown in Figure 7, involves estimation of fundamental matrix from two images of a table which can be obtained from Peter Kovesi’s website [9]. Using a Harris corner detector [3], 231 points in image one were found as shown in Fig 7(a) and their corre-
sponding points in image two, shown in Fig 7(b), were also found by normalized correlation matching.

The second experiment shown in Figure 8, involves estimation of fundamental matrix from two images of a person’s head which can be obtained from Zhang’s website [15]. Fig 8(a) shows the 146 point found by Harris corner detector [3] and Figure 8(b) shows their corresponding points in image two, found by normalized correlation matching. In both Figure 7(b) and Figure 8(b), the dashed connecting lines represent outliers and solid lines show the inlier matches (yet to be identified by the robust estimators).

In each experiment, MSSE, pbM and HBM have been applied to estimate the epipolar geometry and the number of identified inlier matches, estimation error (standard deviation of the Sampson distances for the inliers), and the computation times have been recorded as listed in Table 1. The results again show that HBM is substantially faster than the other two techniques while the estimation error and segmentation performance of the three estimators are comparable.

4 Conclusions

A computationally efficient high breakdown robust estimator, called HBM, was introduced for solving epipolar geometry estimation problems. Since HBM has a differentiable objective function, a closed form updating formula (solving a weighted least squares problem) can be mathematically derived to optimize its objective function.

The closed mathematical form of HBM and its stability (theoretically supported by stability properties of re-descending M-estimators) combined with its high breakdown point and its fast convergence speed make this estimator an outstanding choice for solving epipolar geometry estimation problems.

A number of experiments involving fundamental matrix estimation, using both synthetic and real image pairs have been conducted to benchmark the performance of the proposed estimator. The computation time of HBM is compared with other state of the art high breakdown estimators using the same computing platforms (CPU, memory, software, etc.) and the results show that HBM outperforms the other techniques in terms of computation, while having a similarly high breakdown point and estimation accuracy. For real-time epipolar geometry estimation and multiple motion segmentation problems that involve multi-structured data segmentation (large fractions of outliers), practitioners are recommended to apply the HBM-estimator for fast com-


Figure 8. Head image sequence: (a) the points in the first image and (b) their matching points.

Acknowledgment

This research was supported by the Australian Research Council and Pacifica Group Technologies (PGT) through the ARC Linkage Project grant LP0561923.

References


