

PROMEDS: AN ADAPTIVE ROBUST FUNDAMENTAL MATRIX ESTIMATION APPROACH

Alberto Irurueta Carro

Josep Ramon Morros

Visual Engineering, Barcelona, Spain
airurueta@visual-engin.com

Universitat Politècnica de Catalunya, Barcelona, Spain
ramon.morros@upc.edu

ABSTRACT

Accurate fundamental matrix estimation from computed correspondences is hard to achieve depending on the constraints on computational time and available data (i.e. correspondences and quality scores). Several algorithms exist for this task, like the 8-points, the 7-points algorithm [1] or robust methods such as RANSAC [2], MSAC [3] or LMedS [4]. Robust methods are capable of discriminating correspondence outliers, thus, obtaining better results. Additionally, some variations of the previous methods have been proposed. For instance PROSAC [5] is an improvement of RANSAC which takes into account additional information of the quality of the matches to largely reduce the computational cost of the fundamental matrix estimation process. This work proposes a new robust method for fundamental matrix estimation that combines the benefits of PROSAC and LMedS algorithms, namely improved quality, reduced computational time and less parameters to adjust.

Index Terms — Fundamental Matrix, PROSAC, LMedS

1. INTRODUCTION

Fundamental matrix estimation is based on computing a set of point correspondences between two images and solving the homogeneous set of linear equations obtained from equation (1),

$$\mathbf{m}'^T \cdot \mathbf{F} \cdot \mathbf{m} = 0 \quad (1)$$

where \mathbf{m} is a point on the first image and \mathbf{m}' is the corresponding point the second image, both expressed in homogeneous coordinates. After expansion of (1), and using N corresponding points, the expression becomes (2) which is a homogeneous linear system of equations that can be solved by finding its null-space using the Singular Value Decomposition (SVD) of matrix \mathbf{A} in terms \mathbf{U} , \mathbf{S} and \mathbf{V} , where \mathbf{U} and \mathbf{V} are orthonormal and \mathbf{S} is a diagonal matrix.

$$\mathbf{A} \cdot \mathbf{f} = \mathbf{0}_N \Rightarrow \begin{bmatrix} \tilde{x}_1 \cdot \tilde{x}'_1 & \tilde{y}_1 \cdot \tilde{x}'_1 & \tilde{x}'_1 & \tilde{x}_1 \cdot \tilde{y}'_1 & \tilde{y}_1 \cdot \tilde{y}'_1 & \tilde{y}'_1 & 1 \\ \tilde{x}_N \cdot \tilde{x}'_N & \tilde{y}_N \cdot \tilde{x}'_N & \tilde{x}'_N & \tilde{x}_N \cdot \tilde{y}'_N & \tilde{y}_N \cdot \tilde{y}'_N & \tilde{y}'_N & 1 \end{bmatrix} \cdot \mathbf{f} = \mathbf{0}_N \quad (2)$$

where (2) contains inhomogeneous coordinates $(\tilde{x}_i, \tilde{y}_i)$ of the N corresponding points. Because the Fundamental Matrix (\mathbf{F}) is a 3×3 matrix with rank 2, then (2) can be solved when 8 matches are available by using the last column of \mathbf{V} obtained from SVD and reordering, as in (3). This is the 8-point algorithm [1],

$$\begin{aligned} \mathbf{F} &= \mathbf{U}\mathbf{S}\mathbf{V}^T \Rightarrow \\ \mathbf{f} &= \mathbf{v}_9 = (f_{11}, f_{12}, f_{13}, f_{21}, f_{22}, f_{23}, f_{31}, f_{32}, f_{33})^T \Rightarrow \\ \mathbf{F} &= \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix} \end{aligned} \quad (3)$$

The solution in (3) does not ensure that the resulting fundamental matrix has rank 2, for that reason, using (4), rank 2 is imposed by setting the smallest singular value in \mathbf{S} to zero.

$$\mathbf{F} = \mathbf{U}\mathbf{S}\mathbf{V}^T \Rightarrow \mathbf{F}' = \mathbf{U} \cdot \text{diag}(s_1, s_2, 0) \cdot \mathbf{V}^T \quad (4)$$

When only 7 points are available, (2) has at most rank 7 and its null-space has dimension 2. In such case, the solution to (2) is any linear combination of the last two columns of \mathbf{V} in (3), hence

$$\mathbf{f} = \alpha \cdot \mathbf{v}_8 + (1 - \alpha) \cdot \mathbf{v}_9, \quad 0 \leq \alpha \leq 1 \quad (5)$$

By enforcing rank 2 of the fundamental matrix by setting its determinant to zero, a third degree polynomial on α is obtained. \mathbf{f} can be obtained after finding the real roots of such polynomial. This is the seven points algorithm [1]. If more than 8 correspondences are available in (2), an LMSE solution can be obtained.

2. ROBUST METHODS FOR FUNDAMENTAL MATRIX ESTIMATION

All methods shown up to this point are considered non robust because they take into account all the matches provided. In real situations, some correspondences might contain much larger error than others, which can be considered outliers. Non robust methods largely degrade in such situations. By being able to discriminate these outliers, robust methods can provide much better quality because points found to be outliers are not taken into consideration.

RANSAC [2] is a robust method to estimate parameters of a mathematical model (fundamental matrix in our case) from a set of observed data which contains outliers (i.e. data that do not fit the model). To estimate the fundamental matrix, random subsets of 7 point correspondences are iteratively picked and used to estimate the fundamental matrix using the 7 point algorithm. At each iteration, the confidence on the estimation is obtained by testing the remaining points of one image against the model (using eq. (1)) to predict the corresponding points in the other image and computing the number of hypothetical inliers (correspondences that fit the estimated model at current iteration) and computing the error of the inliers relative to the model. Correspondences are considered to be *inliers* when residuals (difference between predicted and computed correspondences) are smaller than a given threshold t , which is typically 1 pixel. This is, RANSAC attempts to find a solution that minimizes the total residual error (6) by using the discriminant function (7).

$$\min \sum_i f(r_i) \quad (6)$$

$$\text{where } f(r_i) = \begin{cases} 1, & r_i > t \\ 0, & r_i \leq t \end{cases} \quad (7)$$

where r_i is the residual of i -th sample and t the selected inlier threshold. Thus, the ranking of a solution is its cardinality.

The estimated model is stored if it is better (i.e. sufficiently many points have been classified as hypothetical inliers) than models in previous iterations. Iterations continue until a desired level of confidence c on the estimation is found (or until a given number of iterations is achieved). The estimated fundamental matrix from the subset of points that has produced the largest number of inliers is used as the solution of the algorithm.

Although RANSAC produces good results it has some disadvantages. First, the threshold must be set to a given value, which is typically 1 pixel, but depending on the quality of the matches, it can produce a very small number of inliers, which largely increases the number of iterations. On the other hand, if the threshold is too large respect to the accuracy of the matches, then the estimated fundamental matrix loses accuracy. Hence, setting the threshold is problem dependent and might need to be adjusted in each case. Besides, as the number of inliers decreases the number of required iterations increase exponentially, making RANSAC unpractical for situations with approximately less than 50% of inliers.

The median has always been considered a good measure when finding LMSE solutions. Having this in mind, MSAC [3] uses the following discriminant function,

$$f(r_i) = \begin{cases} r_i, & r_i \leq t \\ t, & r_i > t \end{cases} \quad (8)$$

Inliers are scored according to their fitness to the model, while outliers are given a constant weight. The solution that produces the smallest median of residuals is picked. As will be seen later, MSAC produces very similar results respect to RANSAC and has the same disadvantages as it also requires a threshold that is application dependent and a similar computational cost.

LMedS [4], although similar to previous robust algorithms, is based on the idea of minimizing the median of the residuals, but the threshold to determine which matches are inliers is determined automatically as a value proportional to the median of the residuals produced for each subset. The LMedS algorithm typically produces better results than RANSAC or MSAC, specially when the inliers contain low levels of noise, because in such cases RANSAC or MSAC might have thresholds set too large and LMedS adapts more gracefully to such levels of noise.

PROSAC is an improvement of RANSAC which takes into account additional information of the quality scores of the matches to largely reduce the number of iterations of the fundamental matrix estimation process. Matches are ordered from best to worse quality and the algorithm starts picking subsets of 7 points among the points having the largest quality. If a suitable solution having a minimum number of inliers is not found, the algorithm continues by selecting matches with lower quality using a given growing function that determines the new minimum of quality score required to pick new subsets. The algorithm continues decreasing the quality until subsets of points can be picked among all points ranging from largest to lowest quality if a solution is not found first. As the algorithm starts picking matches with the largest quality, it usually finds solutions in a few iterations and becomes much faster than RANSAC. PROSAC produces results with quality similar to RANSAC and can be up to 100 times faster. As PROSAC needs a threshold to determine which matches are inliers, it suffers from the same problem as RANSAC in that regard.

3. PROMEDS

Because LMedS is able to automatically set the threshold to determine inliers while at the same time producing slightly better accuracy than RANSAC or MSAC we propose in this document the combination of the PROSAC and LMedS algorithms to get the best of both algorithms in terms of speed and quality of the estimation. We call this algorithm PROMedS. PROMedS behaves exactly as PROSAC except that it picks the fundamental matrix that minimizes the median of the residuals as LMedS does instead of picking the one that

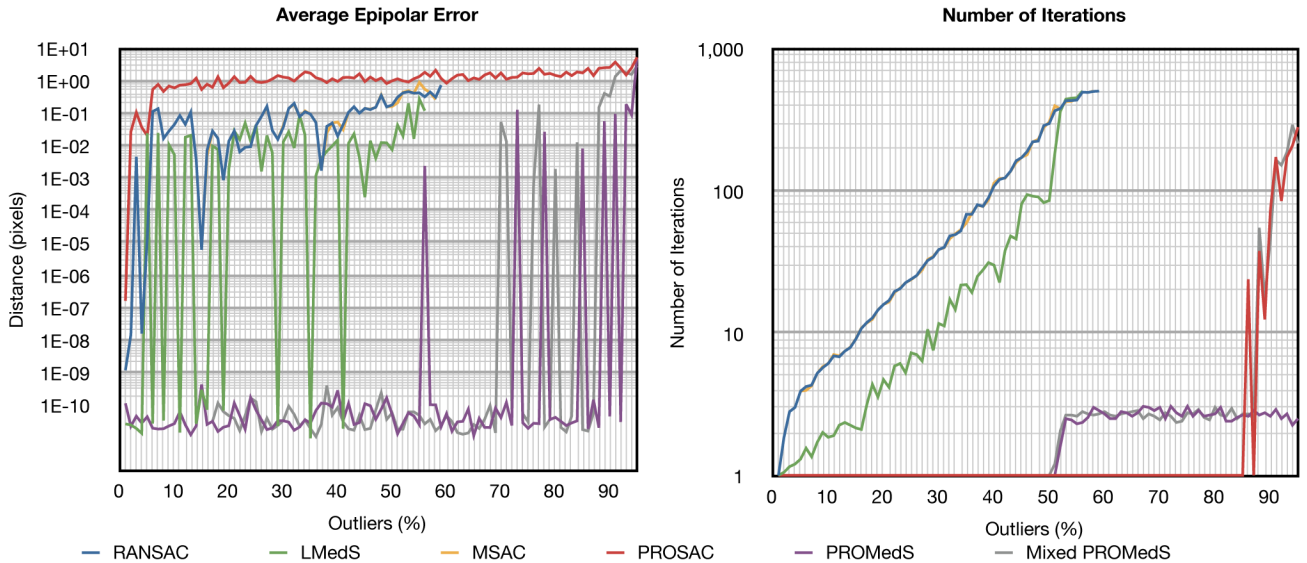


Figure 1. Error in fundamental matrix estimation respect to ground truth [6] and number of iterations for different fundamental matrix robust estimation methods depending on the outliers percentage. Outliers contain Gaussian noise with $\sigma = 5$ pixels, inliers do not have noise. PROMedS and Mixed PROMedS are able to adapt to the lack of inlier noise and obtains much better accuracy than PROSAC with approximately the same computational cost (graphs are shown in logarithmic scale). For large levels of outliers (i.e. greater than 85%) PROMedS iterates less than Mixed PROMedS

maximizes the number of inliers as RANSAC or PROSAC do.

$$t = \text{median}(r_0, r_1, \dots, r_i, \dots, r_N) \quad (9)$$

where r_i is the residual of the i -th sample.

Inliers are determined as in LMedS using the median of the residuals as a reference (9). With these simple modifications, PROMedS is able to estimate the fundamental matrix much more accurately than PROSAC when inliers contain low levels or noise, at a similar computational cost, as will be seen in section 4. However, in practical situations, inliers might contain a certain level of noise (as in Fig. 2), and depending on the amount of that noise PROMedS might provide worse accuracy than PROSAC. To alleviate this situation, we would like PROMedS to make extra iterations in order to provide a more accurate solution, as PROSAC does.

For that reason we proposed the Mixed PROMedS, which is a variation of PROMedS also taking a threshold on the residuals to determine when samples are inliers, just as RANSAC and PROSAC do. Mixed PROMedS only consider as inliers samples having residuals below the provided threshold (typically 1 pixel) or samples below the median of residuals, whichever is stricter in number of inliers. As in PROMedS, Mixed PROMedS picks the fundamental matrix that minimizes the median of residuals.

4. RESULTS

To derive the experimental results, we have computed a random cloud of 3D points consistent of 100 points and two random cameras located at similar points of views. After that, we have projected the 3D points using both cameras to obtain pairs of matching points on their respective image planes. To determine with accuracy the behavior of each algorithm we randomly set a percentage of the projected points as outliers by adding to them an amount of Gaussian noise with $\sigma = 5$ pixels. In order to test algorithms that required quality scores, we set quality scores for each match as follows,

$$q_i = \frac{1}{1 + n_i} + e_i \quad (10)$$

where q_i is the quality score, n_i is the amount of noise on the projected matches and e_i is a given amount of error on the quality score itself, because quality measures of matches contain a certain amount of noise. Even if scores were totally random, PROSAC would behave similarly to RANSAC and PROMedS would behave as LMedS in terms of computational cost and accuracy. Because as the percentage of outliers increases, the number of iterations on robust methods increase exponentially, we have limited the number of iterations to 500 to make their computation feasible. Data is not available in figures of this document when it exceeds this limit, although additional measures can be checked at [7] where it can be seen that when algorithms exceed this limit their accuracy worsens as the number of outliers grow.

We repeated the process of data generation 20 times to average the results provided for each algorithm and then continued repeating again everything by increasing the amount of outliers until a 95% of outliers was reached. The data tested on each algorithm was exactly the same at each run to ensure that results were comparable, and the process was repeated to average the results to obtain more meaningful measures.

Figures 1-2 show comparative measures of different robust methods for fundamental matrix estimation. Measures show average epipolar error [6] (i.e. distance of epipolar lines in estimated fundamental matrix to matched points in their corresponding ground-truth epipolar lines) to determine the accuracy of the estimation, and number of iterations to account for the computational cost of each algorithm.

The average epipolar error explained in [6] is a more reliable method to compare fundamental matrices than comparing the error of estimated epipoles. This is specially true in situations where epipoles might be located far away (e.g. epipoles are at infinity in rectified images) and machine precision might decrease for such high values. For that reason we use that method to compare the accuracy of the estimations in the detailed results provided at [7].

Dataset	RANSAC	PROSAC	PROMedS
Aloe	0.28 / 54	4.7e-13 / 1	1.26e-7 / 1
Art	0.007 / 60	6.5e-4 / 2	0.001 / 2
Baby1	0.003 / 40	1.5e-10 / 2	5.4e-5 / 2
Baby2	0.02 / 44	1.7e-14 / 1	7.6e-13 / 1
Baby3	0.007 / 56	1.34e-12 / 1	3.2e-6 / 1
Bowling1	0.01 / 71	4.5e-14 / 1	3.1e-13 / 1
Bowling2	0.009 / 59	2.0e-12 / 1	3.0e-7 / 1
Cloth1	0.05 / 32	0.07 / 1	0.009 / 2
Dolls	0.006 / 49	3.12e-13 / 1	4.53e-10 / 1
Flowerpots	0.008 / 61	8.8e-5 / 6	2.1e-4 / 5
Lampshade1	0.01 / 71	2.6e-14 / 1	1.6e-11 / 1
Lampshade2	0.02 / 66	0.007 / 1	0.007 / 2
Laundry	0.003 / 68	0.003 / 8	0.02 / 3
Midd1	0.04 / 66	1.2e-13 / 1	3.2e-13 / 1
Monopoly	0.004 / 54	2.15e-14 / 1	6.49e-12 / 1
Plastic	0.003 / 63	2.8e-13 / 1	1.12e-7 / 1
Reindeer	0.004 / 65	0.005 / 1	0.004 / 2
Rocks1	0.002 / 40	4.1e-12 / 1	2.3e-6 / 1
Rocks2	0.005 / 52	3.8e-14 / 1	7.42e-11 / 1
Wood1	0.007 / 51	5.2e-12 / 1	1.1e-7 / 1
Wood2	0.002 / 51	1.43e-10 / 3	0.009 / 4

Table 1. Each cell contains average epipolar error vs number of iterations for datasets [8]. Larger set of results is available in [7]

Among all robust methods, Mixed PROMedS and PROMedS look the most promising in terms of accuracy and computational cost, closely followed by PROSAC with similar computational cost but worse accuracy. Although RANSAC or MSAC achieve better accuracy for low percentage of outliers, their computational cost is much higher, increasing exponentially as the outliers increase, becoming unpractical when outliers are larger than approximately 50%.

The experiments were done by considering inliers without noise, as in Fig. 1, and then repeated with a small quantity of gaussian noise on them (i.e. $\sigma = 0.25$ pixels), as in Fig. 2. The latter was done because typically correspondences are found up to a certain accuracy (typically 1 pixel) unless matching methods with subpixel accuracy are used, but even in those cases a certain amount of noise will still be present.

Results shown in figures 1-2 show that Mixed PROMedS is the most capable to adapt to the amount of noise on the inliers, getting up to 1 million times more accuracy than PROSAC when inliers contain no noise and up to twice the accuracy when inliers contain some noise at a similar cost than PROSAC.

PROMedS was a second close, getting similar results to Mixed PROMedS when inliers contain no noise, but degrading considerably when noise is present, giving worse results than PROSAC but having a better computational cost and needing no additional adjustments as the inlier threshold is estimated automatically. Both algorithms excel when the number of outliers is very high, providing good estimations.

Results obtained in our synthetic set of data were confirmed against real data sets obtained from Middlebury's stereo site [8]. We computed the ground truth of the fundamental matrix using the ground truth disparity from the datasets (up to 1 pixel of accuracy), and then computed Harris corners [9] on each pair of images to obtain a set of possible matches. Several algorithms for Fundamental matrix estimation were tested against this matches corroborating our synthetic results. Table 1 shows some results of average epipolar error for some sets, but the complete set of results for both real and synthetic data is available at [7].

In summary, results in Table 1 show that PROMedS performs similarly in accuracy and computational cost to PROSAC, however, PROMedS sets the inlier threshold automatically whereas in PROSAC is a problem dependent parameter. Besides, results in figures 1-2 show that depending on the amount of noise on the inliers, PROMedS can obtain much better accuracy than PROSAC.

5. REFERENCES

- [1] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, 2003.

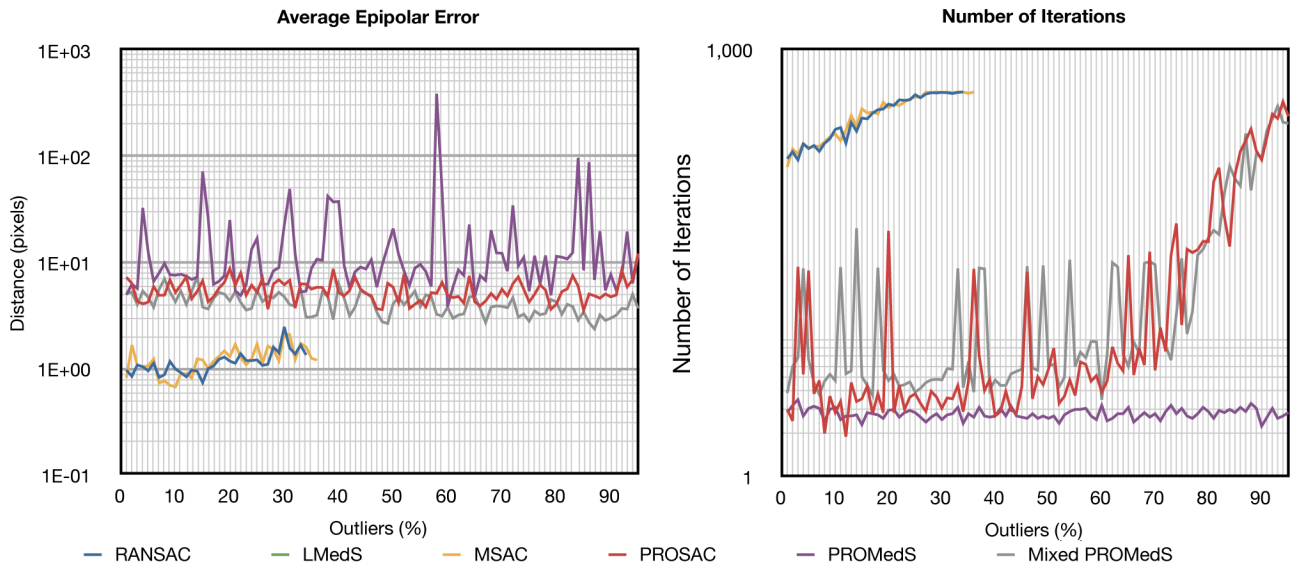


Figure 2. Error in fundamental matrix estimation respect to ground truth [6] and number of iterations for different fundamental matrix methods depending on the outliers percentage. Outliers contain Gaussian noise with $\sigma = 5$ pixels, inliers Gaussian noise with $\sigma = 0.25$ pixels. Mixed PROMedS is able to adapt to the level of inlier noise and slightly better accuracy than PROSAC with approximately the same computational cost (graphs are shown in logarithmic scale). PROMedS achieves slightly worse accuracy than PROSAC but with lower and more predictable computational cost. Data for LMedS is not shown because the 500 iterations limit is always reached, but the complete set of results can be seen at [7]

- [2] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Comm. Assoc. Comp. Mach.*, vol. 24, no. 6, pp. 381–395, 1981.
- [3] P. H. S. Torr and A. Zisserman, "MLESAC: A new robust estimator with application to estimating image geometry," *Computer Vision and Image Understanding*, vol. 78, 2000.
- [4] P.J. Rousseeuw and A.M. Leroy, *Robust Regression and Outlier Detection*, John Wiley & Sons, New York, 1987.
- [5] O. Chum and J. Matas, "Matching with PROSAC - progressive sample consensus," in *IEEE Conf. on Computer Vision and Pattern Recognition*, June 2005.
- [6] Z. Zang, "Determining the epipolar geometry and its uncertainty: A review," *Int. Journal of Computer Vision*, 1998.
- [7] "Promeds additional results," <http://imatge.upc.edu/web/?q=node/1386>.
- [8] "Middlebury stereo vision page," <http://www.middlebury.edu/stereo>.
- [9] C. Harris and M. Stephens, "A combined corner and edge detection," in *Proceedings of The Fourth Alvey Vision Conference*, 1988, pp. 147–151.