# MAXIMUM KERNEL DENSITY ESTIMATOR FOR ROBUST FITTING

*Hanzi Wang*
Center for Computer-Integrated Surgical System and Technology
Department of Computer Science
The Johns Hopkins University, Baltimore, MD, 21218

**ABSTRACT**

Robust model fitting plays an important role in many computer vision applications. In this paper, we propose a new robust estimator — Maximum Kernel Density Estimator (MKDE) based on the nonparametric kernel density estimation technique. It can be viewed as an improved version of our previously proposed Quick Maximum Density Power Estimator (QMDPE) [15]. Compared with QMDPE, MKDE does not require running the mean shift algorithm for each candidate fit. Thus, the computational complexity of MKDE is greatly reduced while the accuracy of MKDE is comparable to QMDPE and outperforms that of other popular robust estimators such as LMedS and RANSAC. We evaluate MKDE in robust line fitting and fundamental matrix estimation. Experiments show that MKDE has achieved promising results.

***Index Terms*** — machine vision, robustness, model fitting, kernel density estimation, algorithms

## 1. INTRODUCTION

Robust regression techniques have been widely used in many computer vision applications such as range image segmentation [15, 19], fundamental matrix estimation [10], optical flow calculation [1, 18], visual tracking [17], etc. The key behind those robust approaches is in that the approaches can resist the influence of outliers, which do not belong to the model to fit. In real situations, data are often noisy and contaminated by outliers such as mismatches, wrong segmentation, multiple structures. One main task in computer vision community is to find a robust estimator which can tolerate highly contaminated data while the computational complexity of the estimator is reasonably low.

There are a lot of robust estimators having been proposed in the literature (e.g., [2, 4, 6, 8-10, 12, 15, 19]). The maximum-likelihood estimators (M-estimators) [6] minimize the sum of symmetric, positive-definite functions of the residuals with a unique minimum at zero. The Least Median of Squares (LMedS) estimator [9] minimizes the median of squared residuals. However, both the M-estimators and the LMedS estimator can not tolerate more than 50% outliers. When data involve more than 50% outliers, both methods break down.

Great efforts have been made to seek for highly robust estimators which can tolerate more than 50% outliers. Random SAmple CONsensus (RANSAC) [4] is such a robust estimator which, given a correct estimate of the scale of inliers, can be robust to more than 50% outliers. Unlike M-estimators and LMedS which find a model by minimizing their objective functions, RANSAC maximizes its objective function, i.e., the number of samples within the given scale of inliers (or *error tolerance*). However, the performance of RANSAC largely depends on the given scale or *error tolerance*. MUSE [8], RESC [19] and MINPRAN [12] can deal with more than 50% outliers as well. However, MUSE needs a lookup table for the scale estimator correction and it can not handle extreme outliers; RESC needs the user to tune many parameters in compressing a histogram. MINPRAN [12] is computational expensive and can not effectively deal with multiple structures.

Recently we have proposed a robust estimator QMDPE [15] which uses nonparametric density estimation and density gradient estimation techniques. In this paper, we, based on QMDPE, propose a new robust estimator — MKDE (Maximum Kernel Density Estimator) which uses the nonparametric kernel density estimation technique. Compared with QMDPE, the procedure of MKDE is simplified and the computational efficiency of MKDE is improved. MKDE is usually 2 to 4 times faster than QMDPE while it achieves comparable accuracy as QMDPE. Compared with RANSAC, MKDE is less sensitive to the user-specified error tolerance (or scale) and has similar computational efficiency.

## 2. METHODOLOGY

We begin with simply reviewing the QMDPE algorithm proposed in [15]. Then we propose MKDE and the complete procedure of MKDE.

### 2.1. Review of QMDPE

Let $y_i$ is a response variable; and $(x_{i1},...,x_{ip})$ are the explanatory variables. The classical linear model can be described as followings:

$$y_i = x_{i1}\theta_1 + ... + x_{ip}\theta_p + e_i \quad (i = 1,...,n) \tag{1}$$

where $\theta = (\theta_1 \cdots \theta_p)'$ are the regression coefficients. The error term $e_i$ is usually assumed to be normally distributed with $N(0, \sigma)$.

The residual $r_i$ for the $i$'th set of observed data is the difference between the estimated value $\hat{y}_i$ ( $\hat{y}_i = x_{i1}\hat{\theta}_1 + ... + x_{ip}\hat{\theta}_p$ ) and the actually observed value $y_i$:

$$r_i = y_i - x_{i1}\hat{\theta}_1 - ... - x_{ip}\hat{\theta}_p \quad (2)$$

The purpose of regression techniques is to estimate the regression coefficients by minimizing (such as M-estimators, LMedS) or maximizing (e.g., RANSAC, QMDPE etc.) an objective function of residuals of the data. QMDPE assumes when a model is correctly fitted, the absolute residual value corresponding to a local peak $\lambda_p$, which is obtained by running a mean shift procedure [3] with initial position of zero in residual space, should be as small as possible; and the probability density $\hat{f}(\lambda_p)$ at $\lambda_p$ should be as high as possible. QMDPE can be written as:

$$\hat{\theta} = \arg\max_{\theta} O(\lambda_p, \theta) \quad (3)$$

$$O(\lambda_p, \theta) = \frac{\left(\hat{f}_\theta(\lambda_p)\right)^\nu}{\exp(|\lambda_p|)} \quad (4)$$

where $\nu$ is a constant factor that adjusts the relative influence of the probability density $\hat{f}_\theta(\lambda_p)$ to the residual at the point $\lambda_p$.

### 2.2. MKDE

QMDPE has to run the mean shift algorithm for each candidate to find $\lambda_p$ which corresponds to a local peak. Thus, the computational efficiency of QMDPE is lowered. In this section, we modify QMDPE and derive the MKDE method.

Unlike LMedS and M-estimators which assume inliers occupy an absolute majority of the data, we assume that inliers occupy a relative majority, with a Gaussian-like distribution, of the data points. When a model is correctly fitted, the residuals of inliers should be as close to zero as possible. Thus, the probability density at the origin point ($\mathcal{P}_0$) with zero residual value in residual space should be as high ass possible. Let $\hat{f}^*$ be the probability density at the point $\mathcal{P}_0$ in residual space. MKDE can be written as:

$$\hat{\theta} = \arg\max_{\theta} \hat{f}_\theta^* \quad (5)$$

Let $\{r_i\}_{i=1,...,n}$ be a set of $n$ residual points, the kernel density $\hat{f}^*$ with kernel $K$ and bandwidth $h$ is defined as follows:

$$\hat{f}^* = \frac{1}{nh}\sum_{i=1}^{n} K(\frac{r_i}{h}) \quad (6)$$

We employ the Epanechnikov kernel [11] which yields the minimum mean integrated square error (MISE):

$$K_e(x) = \begin{cases} \frac{1}{2}c_d^{-1}(d+2)(1-x^T x) & \text{if } x^T x < 1 \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

where $c_d$ is the volume of the unit d-dimensional sphere. The dimension of residual space is 1.

Compared with QMDPE in equations (3) and (4), MKDE is different in that (1) there is no need to estimate the position of local peak $\lambda_p$. Thus, we do not need to run the mean shift algorithm for each candidate fit in MKDE; (2) MKDE is theoretically solid. The factor $\nu$ in QMDPE is not required in MKDE. Experiments in Section 3 show that MKDE is very robust to outliers. MKDE can achieve comparable (or better) performance to QMDPE while it is two to four times faster than QMDPE.

Like QMDPE, we employ a hypothesize-and-select scheme in which we randomly sample enough subsets $\eta$ and choose the best candidate that yields the highest score (i.e., $\hat{f}^*$). Let $P$ be the probability that at least one "clean" $p$-subset is chosen; $\varsigma$ be the percentage of outliers involved in the data. We have:

$$\eta = \frac{\log(1-P)}{\log[1-(1-\varsigma)^p]} \quad (8)$$

Our work is different to the work of Chen and Meer [2] and its variants [10, 13] in that the latter ones place emphasis on the projection pursuit paradigm while MKDE considers the kernel density of the mode in residual space. The methods in [2, 10, 13] need to seek for the mode by maximizing the density in the projection space, which maximizes the projection index. In contrast, MKDE seek for the mode by maximizing the kernel density of the mode in residual space. Thus, MKDE is computationally efficient.

### 2.3. The procedure of MKDE

The complete procedure of the MKDE algorithm can be described as follows:

**Step 0:** Input a set of data points, the bandwidth $h$, and a repetition number $\eta$.

**Step 1:** Randomly choose a $p$-subset.

**Step 2:** Estimate the model parameters by the chosen $p$-subset.

**Step 3:** Calculate the residuals of all data points to the estimated parameters.

**Step 4:** Calculate a score for the p-subset, according to equations (6) and (7).

**Step 5:** Repeat step (1) to step (4) $\eta$ times.

**Step 6:** Output the parameters with the highest score.

**Step 7:** (Optional) Once the parameters of the model are estimated, a robust scale estimator (TSSE [16]) can be employed to refine the scale of inliers.

## 3. EXPERIMENTS

We evaluate our method in robust line fitting and fundamental matrix estimation. We also compare the performance of MKDE with those of LMedS, RANSAC, and QMDPE. In all experiments, the four methods are implemented 20 times and the averaged results are used.

### 3.1. Robust line fitting

First, we test the influence of outliers on the performance of the four methods. We generate a one-step signal (y = Ax + B) with total 1000 data points: line 1: x: (0–65), A = 0, B = 70, the number of data points will be decreased with the increase of randomly distributed outliers so that outlier percentage $\varsigma$ changes from 0 to 85%; line 2: x: (65–100), A = 0, B = 20, the number of data points is fixed at 100. Data points on both lines are corrupted by Gaussian noise with zero mean and unit standard variance.

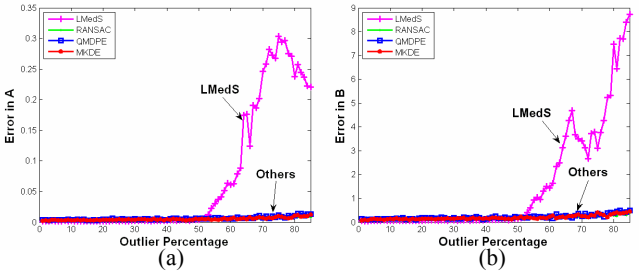



Fig. 1. Breakdown plot for the four methods: (a) and (b) estimation errors in A and B vs. the percentage of outliers. RANSAC, QMDPE and MKDE are given the correct scale estimate of inliers.

As shown in Figure 1, the LMedS estimator achieves good results when the outlier percentage is less than 50% but it begins to break down when the data involve more than 50% outliers. In contrast, RANSAC, QMDPE and MKDE with the correct scale estimate have achieved accurate results even when the data contain more than 80% outliers.

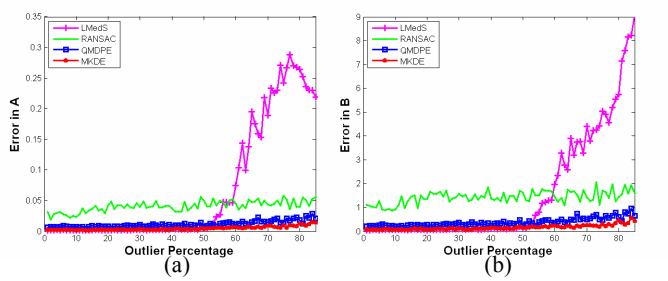



Fig. 2. Breakdown plot for the four methods: (a) and (b) estimation errors in A and B vs. the percentage of outliers. RANSAC, QMDPE and MKDE are given a scale which is 5 times of the correct scale.

However, in real situations, the accurate scale estimate of inliers is not always available. In some cases, we can only give an approximate estimate of the scale. It is interesting to evaluate the influence of the given scale on the performance of the methods. In figure 2 and Table 1, we use a scale which is 5 times lager than the true scale. The performance of LMedS does not depend on the input scale and LMedS achieves similar results as in Figure 1. Among RANSAC, QMDPE and MKDE, RANSAC achieves inaccurate results throughout the tests. This means that the performance of RANSAC is affected when the scale is biased. Both QMDPE and MKDE obtain good results for all tests while MKDE achieves relatively more accurate results than QMDPE.

| | LMedS | RANSAC | QMDPE | MKDE |
|---|---|---|---|---|
| Mean Error in A | 0.0675 | 0.0408 | 0.0116 | 0.0047 |
| Mean Error in B | 1.5740 | 1.4235 | 0.3766 | 0.1588 |

Table 1. Mean error in A and B for the four methods. MKDE achieves the best results among the four methods.

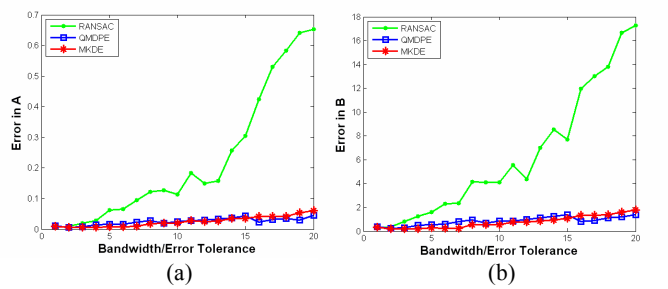



Fig. 3. Influence of the scale on the performance of RANSAC, QMDPE and MKDE. (a) and (b) estimation errors in A and B vs. the scale.

In figure 3, we generate a signal, which is similar to the signal used in figure 1, with 75% outliers. We change the value of the scale from 1 to 20. As we can see, the performance of RANSAC is greatly affected by the scale value. In comparison, the performance of QMDPE and MKDE is much less sensitive to the change in the scale value.

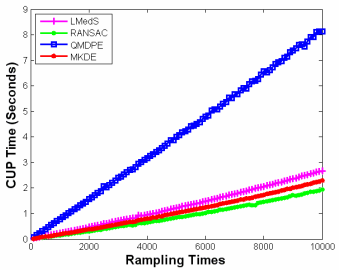


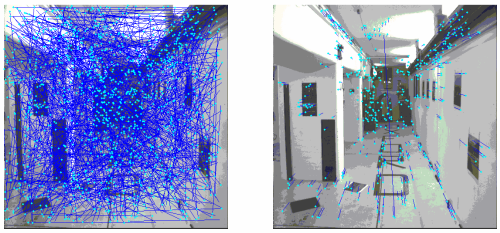

Fig. 4. Processing time of the four methods vs. the sampling times $\eta$. All methods are performed in MATLAB code.

Figure 4 shows the processing time of the four methods with respect to the number of sampling times. From figure 4 we can see that the computational time of QMDPE is the most expensive while RANSAC is the fastest among the four methods. The processing time of MKDE is slightly slower than that of RANSAC but it is relatively faster than LMedS. Generally speaking, MKDE is about 2 to 4 times faster than QMDPE.

### 3.2. Fundamental matrix estimation

We also apply our method to estimate the fundamental matrix between multiple views. We use two frames of the Corridor sequence (bt.000 and bt.001) from http://www.robots.ox.ac.uk/~vgg/data/. We employ SIFT matching algorithm [7] to detect the matches between the two images. There are 612 matches detected. We add 700 random matches to increase the percentage of outliers so that the outlier percentage is more than 50%. Image points are normalized by using the Hartley's algorithm [5] before we employ the 7 points algorithm [14] to solve for candidate fits.

To evaluate the performance of the methods, we use the following four error measures: Sum of absolute differences in the left epipole (SADL); Sum of absolute differences in the right epipole (SADR); Mean of absolute reprojection errors (MARE); Standard variance of reprojection errors (SVRE).



(a) (b)

Fig. 5. One of the 20 tests. (a) the input matches (SIFT matches and random matches); (b) the matches selected by MKDE.

Figure 5 illustrates that MKDE is robust to mismatches (i.e., outliers) and can select correct matches (note: there is one random match is selected by MKDE as it satisfies the epipolar constraint). Table 2 shows that LMedS totally breaks down because it can not tolerate more than 50% outliers. MKDE achieves the most accurate results in SADL and SADR. In the measure of MARE and SVRE, MKDE performs relatively better than QMDPE but slightly worse than RANSAC.

| | LMedS | RANSAC | QMDPE | MKDE |
|---|---|---|---|---|
| SADL | 157.7179 | 7.9077 | 7.9282 | 6.7878 |
| SADR | 158.4915 | 8.2237 | 8.3003 | 7.0998 |
| MARE | 1.9415 | 0.1173 | 0.1252 | 0.1216 |
| SVRE | 3.5441 | 0.1462 | 0.1543 | 0.1508 |

**Table 2.** Error measurement of the four methods.

## 4. CONCLUSIONS

The proposed MKDE algorithm employs the nonparametric kernel density estimation technique in determining the merit of model fit. The performance of MKDE is compared with those of popular robust estimators (LMedS, RANSAC) and recently proposed robust estimator QMDPE. MKDE is simple and computationally efficient, and it is 2 to 4 times faster than QMDPE. The scale estimate of inliers has a weaker influence on the performance of MKDE than that of RANSAC. MKDE is very robust to outliers and can tolerate more than 50% outliers. Experiments show that MKDE can successfully deal with data involving more than 80% outliers.

## 5. ACKNOWLEDGEMENT

The author thanks Gregory Hager for his valuable comments. This work is supported by NIH 1R21EB005201 - 01A1.

## 6. REFERENCES

1. Bab-Hadiashar, A. and D. Suter, "Robust Optic Flow Computation". IJCV. Vol. **29,** No.1, pages 59-77, 1998.
2. Chen, H. and P. Meer, "Robust Computer Vision through Kernel Density Estimation". ECCV, pages 236-250, 2002.
3. Comaniciu, D. and P. Meer, "Mean Shift: A Robust Approach towards Feature Space A Analysis". IEEE Trans. PAMI. Vol. **24,** No.5, pages 603-619, 2002.
4. Fischler, M.A. and R.C. Rolles, "Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography". Commun. ACM. Vol. **24,** No.6, pages 381-395, 1981.
5. Hartley, R. and A. Zisserman, Multiple View Geometry in Computer Vision. 2nd ed. 2004. Cambridge University Press.
6. Huber, P.J., "Robust Statistics". New York, Wiley. 1981.
7. Lowe, D.G., "Distinctive Image Features from Scale-Invariant Keypoints". IJCV. Vol. **60,** No.2, pages 91-110, 2004.
8. Miller, J.V. and C.V. Stewart. "MUSE: Robust Surface Fitting Using Unbiased Scale Estimates". CVPR, p.300-306, 1996.
9. Rousseeuw, P.J. and A. Leroy, "Robust Regression and outlier detection". John Wiley & Sons, New York. 1987.
10. Rozenfeld, S. and I. Shimshoni. "The Modified pbM-estimator method and a runtime analysis technique for the RANSAC family". CVPR, Vol. I, p.1113-1120, 2005.
11. Silverman, B.W., Density Estimation for Statistics and Data Analysis. 1986. Chapman and Hall. London.
12. Stewart, C.V., "MINPRAN: A New Robust Estimator for Computer Vision". IEEE Trans. PAMI. Vol. **17,** No.10, pages 925-938, 1995.
13. Subbarao, R. and P. Meer. "Heteroscedastic projection based M-estimators". Workshop on Empirical evaluation Methods in Computer Vision, 2005.
14. Torr, P. and D. Murray, "The Development and Comparison of Robust Methods for Estimating the Fundamental Matrix". IJCV. Vol. **24,**271-300, 1997.
15. Wang, H. and D. Suter, "MDPE: A Very Robust Estimator for Model Fitting and Range Image Segmentation". IJCV. Vol. **59,** No.2, pages 139-166, 2004.
16. Wang, H. and D. Suter, "Robust Adaptive-Scale Parametric Model Estimation for Computer Vision". IEEE Trans. PAMI. Vol. **26,** No.11, pages 1459-1474, 2004.
17. Wang, H. and D. Suter, "A Consensus-Based Method for Tracking: Modelling Background Scenario and Foreground Appearance". PR. Vol. **40,** No.3, pages 1091-1105, 2007.
18. Ye, M. and R.M. Haralick. "Two-Stage Robust Optical Flow Estimation". CVPR, Vol. 2, p.623-628, 2000.
19. Yu, X., T.D. Bui and A. Krzyzak, "Robust Estimation for Range Image Segmentation and Reconstruction". IEEE Trans. PAMI. Vol. **16,** No.5, pages 530-538, 1994.