

NIH Public Access

Author Manuscript

IEEE Trans Pattern Anal Mach Intell. Author manuscript; available in PMC 2010 April 21

Published in final edited form as:

IEEE Trans Pattern Anal Mach Intell. 2010 January ; 32(1): 178–184. doi:10.1109/TPAMI.2009.148.

A Generalized Kernel Consensus-Based Robust Estimator

Hanzi Wang [Member, IEEE],

School of Computer Science, The University of Adelaide, Adelaide SA 5005, Australia. hanzi.wang@ieee.org

Daniel Mirota [Student Member, IEEE], and

Department of Computer Science, The Johns Hopkins University, 3400 N. Charles St., Baltimore, MD 21218. dan@cs.jhu.edu

Gregory D. Hager [Fellow, IEEE]

Department of Computer Science, The Johns Hopkins University, 3400 N. Charles St., Baltimore, MD 21218. hager@cs.jhu.edu

Abstract

In this paper, we present a new Adaptive-Scale Kernel Consensus (ASKC) robust estimator as a generalization of the popular and state-of-the-art robust estimators such as RANdom SAmple Consensus (RANSAC), Adaptive Scale Sample Consensus (ASSC), and Maximum Kernel Density Estimator (MKDE). The ASKC framework is grounded on and unifies these robust estimators using nonparametric kernel density estimation theory. In particular, we show that each of these methods is a special case of ASKC using a specific kernel. Like these methods, ASKC can tolerate more than 50 percent outliers, but it can also automatically estimate the scale of inliers. We apply ASKC to two important areas in computer vision, robust motion estimation and pose estimation, and show comparative results on both synthetic and real data.

Keywords

Robust statistics; model fitting; kernel density estimation; motion estimation; pose estimation

1 INTRODUCTION

Robust statistical techniques play an important role in many areas of computer vision. Applications include segmentation [1], [11], [28], [29], optical flow calculation [2], [15], fundamental matrix estimation [24], [28], as well as many others. The key point of robust methods is to extract the parameters of an underlying structural model while tolerating the influence of outliers, which do not obey the assumed model.

A number of robust estimation techniques (e.g., [3], [4], [7], [13], [17], [18], [19], [24], [26], [27], [29]) have been proposed during the last decades. The Maximum-likelihood estimators (M-estimators) [7], the Least Median of Squares (LMedS) estimator [17], and the RANdom SAmple Consensus (RANSAC) estimator [4] are some of the most popular robust estimators that have been widely used in many areas. However, M-estimators and LMedS cannot deal with data involving more than 50 percent outliers; RANSAC can tolerate more than 50

© 2010 IEEE

Recommended for acceptance by B. Triggs.

For information on obtaining reprints of this article, please send tpami@computer.org, and reference IEEECS Log Number TPAMI-2009-01-0005.

percent outliers, but it requires the user to specify an error tolerance which is not known a priori in many practical environments. The value of this tolerance has a significant influence on the performance of RANSAC [25], [28]. M-estimator SAmple Consensus (MSAC) [24] improves the performance of RANSAC by modifying its cost function, but it also requires a user-specified error tolerance. Minimum Unbiased Scale Estimator (MUSE) [13], MINmize the Probability of RANdomness (MINPRAN) [19], Adaptive Least Kth order Squares (ALKS) [11], Residual Sample Consensus (RESC) [29], and Adaptive Scale Sample Consensus (ASSC) [28] can tolerate more than 50 percent outliers. However, MUSE is biased in the scale estimation, and thus, it needs a lookup table to correct the estimated scale. MINPRAN and ALKS are computationally expensive and will break down for multistructure data with a large percentage of outliers. The RESC estimator requires the user to tune many parameters in its objective function. ASSC assigns equal weights to all inliers, and thus, is less effective in estimating the model parameters. Recently, Chen and Meer [3] modified the cost function of the M-estimators to create projection-based M-estimators (pbM-estimators). Rozenfeld and Shimshoni [18] and Subbarao and Meer [20] proposed some variants to improve the performance of pbM. Lavva et al. [10] applied a modified pbM to 3D range image segmentation. All of these modifications [3], [10], [18], [20] are based on the projection pursuit paradigm [8], [9].

1.1 Contributions of the Paper

In [28], the first author developed the ASSC estimator which uses the ratio of the number of inlier samples to the scale of inliers as its objective function. We now describe a new robust estimator, Adaptive Scale Kernel Consensus (ASKC). ASKC assumes that, with the correct model parameters, the mode at the origin in the residual space has the maximum kernel density. Thus, ASKC estimates the parameters maximizing the kernel density value at the origin of the residual space. As we employ the nonparametric kernel density estimation techniques in the objective function of ASKC, ASKC is based on kernel consensus and is more effective than ASSC and RANSAC which are based on sample consensus. Moreover, ASKC provides a generalized framework for robust model fitting, from which a set of new robust estimators can be derived by employing different kernels. ASKC unifies a set of popular robust estimators, such as RANSAC, ASSC, and Maximum Kernel Density Estimator (MKDE) [26].

To show the ability of ASKC, we apply it to two important areas in computer vision: robust motion estimation and robust pose estimation. We compare the performance of ASKC with those of several other popular robust methods (LMedS, RANSAC, MSAC, RESC, pbM, and ASSC). Experimental results on both synthetic and real data show our method can achieve better results than other methods.

2 THE PROPOSED METHOD

The problem of model fitting can be described as follows: given a set of data points $\{(x_i, y_i)\}_i = 1, ..., N$ (with the explanatory variables $(x_i = (x_{i1}, ..., x_{id}) \in \mathbb{R}^d$ and the response variable $y_i \in \mathbb{R}^1$) and a model **M** with $\mathbf{M}(\mathbf{\theta}, \mathbf{X}) \mapsto \mathbb{R}^N$, where model parameters are $\mathbf{\theta} = (\theta_1, ..., \theta_d)^t \in \mathbb{R}_d$ and data vector is $\mathbf{X} = [(x_1, y_1), ..., (x_N, y_N)]^t \in \mathbb{R}^{N \times d}$, compute an estimate $\mathbf{\theta} \wedge \text{from } \mathbf{X}$.

The classical linear regression (LR) model can be written as follows:

$$y_i = x_{i1}\theta_1 + \dots + x_{id}\theta_d + e_i = x_i\theta + e_i \quad (i=1,\dots,N).$$
(1)

The measurement error term e_i is usually assumed to be independent and identically distributed, and the (identical) distribution is a zero-mean Gaussian $G(0, \sigma^2)$. In many real situations, both the regression parameters θ and the scale of noise σ are unknown and must be estimated.

With an LR model, a *d*-dimensional data point x_i can be mapped to the 1D residual space. Given an estimate $\hat{\theta}$, the residual r_i corresponding to x_i is written as:

$$r_{i,\widehat{\theta}} = y_i - x_{i1}\widehat{\theta}_1 - \dots - x_{id}\widehat{\theta}_d \quad \text{and } \mathbf{M}(\widehat{\theta}, \mathbf{X}) \mapsto \mathbf{R}_{\widehat{\theta}},$$
(2)

where $\mathbf{R}_{\hat{\theta}} = [r_{1,\hat{\theta}}, \dots, r_{N,\hat{\theta}}].$

2.1 Kernel Density Estimation

Given an estimate $\hat{\theta}$ and a set of residuals $R_{\hat{\theta}} = \{r_{1,\hat{\theta}}, ..., r_{N,\hat{\theta}}\}$, the *fixed bandwidth* kernel density estimate at γ with a kernel *K*(.) and a bandwidth *h* is computed as follows:

$$\widehat{\mathbf{f}}_{h}(\gamma; \mathbf{R}_{\widehat{\theta}}) = \frac{1}{Nh} \sum_{i=1}^{N} K\left(\frac{\gamma - r_{i,\widehat{\theta}}}{h}\right).$$
(3)

An alternative is to select a different bandwidth $h = h(\hat{\theta}) \equiv h_{\uparrow}$ for a different parameter estimate \hat{h} . The *variable bandwidth* kernel density estimate can be written as:

$$\widehat{f}_{h_{\widehat{\theta}}}(\gamma; \mathbf{R}_{\widehat{\theta}}) = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{h_{\widehat{\theta}}} K\left(\frac{\gamma - r_{i,\widehat{\theta}}}{h_{\widehat{\theta}}}\right).$$
(4)

We will utilize radial basis function (RBF) kernels which satisfy the following condition:

$$K(\beta) = c_k k(\|\beta\|^2) \quad k(\beta) \ge 0 \quad \text{for} \quad 0 \le \beta \le 1,$$
(5)

where $k(\beta)$ is the profile of the kernel. c_k is a normalization constant to make $K(\beta)$ integrate to one. Note: The value of c_k changes for different kernels.

There are many kernels available in the literature. In this paper, we only consider two popular RBF kernels, the Epanechnikov kernel $K_E(\beta)$ and the normal kernel $K_G(\beta)$:

$$K_{E}(\beta) = \begin{cases} c_{k}(1 - ||\beta||^{2}), & ||\beta|| \le 1, \\ 0, & ||\beta|| > 1, \end{cases}$$

with $K_{E}(\beta) = \begin{cases} 1 - \beta, & 0 \le \beta \le 1, \\ 0, & \beta > 1, \end{cases}$ (6)

$$K_{G}(\beta) = c_{k} \exp\left(-\frac{1}{2}||\beta||^{2}\right), \text{ with } k_{G}(\beta) = \exp\left(-\frac{\beta}{2}\right) \quad \beta \ge 0.$$
⁽⁷⁾

Although this paper investigates the properties of ASKC with the Epanechnikov kernel (henceforth ASKC1) and the normal kernel (henceforth ASKC2), our method can easily employ an arbitrary kernel.

2.2 Bandwidth Estimation

The bandwidth h (or $h\theta$) is a crucial parameter in kernel density estimation. Let $\hat{\sigma\theta}$ be a scale estimate of inliers for a given R_{θ} . An oversmoothed bandwidth estimate is written as follows [21], [22], [25]:

$$\widehat{\boldsymbol{h}}_{\theta} = \left[\frac{243\int_{-1}^{1} K(\beta)^{2} d\beta}{35N\left(\int_{-1}^{1} r^{2} K(\beta) d\beta\right)^{2}}\right]^{1/5} \widehat{\boldsymbol{\sigma}}_{\theta}.$$
(8)

It is recommended that the bandwidth is set as $c_h \hat{h}_{\theta}$ (0 < c_h < 1) to avoid oversmoothing ([25, p. 62]). To estimate the scale of inliers $\hat{\sigma}_{\theta}$, we can employ robust scale estimators such as the median [17], the Median Absolute Deviation (MAD) [16], the robust *k* scale estimator [11], or the two-step scale estimator (TSSE) [28].

In this paper, we use a TSSE-like procedure to estimate the scale of inliers. TSSE employs a mean-shift procedure and a mean-shift valley (MSV) procedure to estimate the scale for multiple-mode data. Our implementation operates as follows: First, we use a robust k scale estimator to compute an initial scale estimate

$$\widehat{\sigma}_0 = |r_\theta|_k / \Psi^{-1}[(1+k)/2], \tag{9}$$

where $|r_{\theta}|_k$ is the $(k \times N)$ th residual of the ordered absolute residuals. $\Psi^{-1}[.]$ is the argument of the normal cumulative density function. For example, when the *k* value is set to 0.1 or 0.2, at least 10 or 20 percent of the data points are included in the shortest window.

Once we have an initial scale estimate, we refine the scale estimate by running TSSE. However, in addition to the Epanechnikov kernel employed in [28] for the mean shift and the mean-shift valley procedures, we may employ the normal kernel (or other kernels) as well. In the case that the normal kernel is used, mean-shift mode-seeking can be written as:

$$\vec{\mathbf{m}}_{h_{\theta},G}(\gamma) = \left[\sum_{i=1}^{N} r_{i,\theta} exp\left(\frac{\|\gamma - r_{i,\theta}\|^2}{h_{\theta}}\right) \left| \sum_{i=1}^{N} exp\left(\frac{\|\gamma - r_{i,\theta}\|^2}{h_{\theta}}\right) \right] - \gamma.$$
(10)

The sequence of successive locations $\{\gamma_i\}_{i=1,...}$ of the MSV procedure is:

$$\gamma_{i+1} = \gamma_i + \zeta \left(\gamma_i - \sum_{i=1}^N r_{i,\theta} exp\left(\frac{\|\gamma_i - r_{i,\theta}\|^2}{h_{\theta}}\right) \Big/ \sum_{i=1}^N exp\left(\frac{\|\gamma_i - r_{i,0}\|^2}{h_{\theta}}\right) \right), \tag{11}$$

where ζ is an adjustment factor which keeps the MSV vectors consistent in direction.

Fig. 1 shows that when the $\hat{\theta}$ estimate is correct (i.e., the parameters of the red horizon line are used in Fig. 1a), the kernel density at the origin point is higher (Fig. 1b). And the ratio of

the kernel densities at the detected peak and valley is higher. In contrast, when the model parameter estimate $\hat{\theta}$ is incorrect, the kernel density estimate at the origin is low and the ratio of the kernel densities at the detected peak and valley is lower (Fig. 1c).

2.3 The ASKC Estimator

We consider the kernel density value at the origin (\mathcal{O}) in the residual space as our objective function. Given a set of residuals $\mathbf{R}_{\theta} = \{r_{i,\theta}\}_{i=1,...,N}$ subject to θ , we define the objective function of ASKC based on (4) as:

$$F_{h_{\theta}}(R_{\theta}) = \widehat{f}_{h_{\theta}}(\mathscr{O}; R_{\theta}) = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{h_{\theta}} K\left(\frac{r_{i,\theta}}{h_{\theta}}\right).$$
(12)

The ASKC estimator can be written as:

$$\widehat{\theta} = \arg\max_{\theta} F_{h_{\theta}}(R_{\theta}) = \arg\max_{\theta} \frac{1}{N} \sum_{i=1}^{N} \frac{1}{h_{\theta}} K\left(\frac{r_{i,\theta}}{h_{\theta}}\right).$$
(13)

In the fixed bandwidth case, (13) can be rewritten as follows (which we termed the MKDE in [26]):

$$\widehat{\theta} = \arg\max_{\theta} F(R_{\theta}) = \arg\max_{\theta} \widehat{f}(\mathscr{O}; R_{\theta}) = \arg\max_{\theta} \frac{1}{Nh} \sum_{i=1}^{N} K\left(\frac{r_{i,\theta}}{h}\right).$$
(14)

In [26], we have shown that the scale estimate of inliers or the bandwidth has a much weaker influence on the performance of MKDE than that of RANSAC, while the computational efficiency of MKDE is comparable to that of RANSAC.

To calculate the solution of (13) (or (14) for the fixed bandwidth case), we employ a sampling procedure to produce a set of candidates. We can employ either a random sampling scheme [17] or a guided sampling technique [23].

Fig. 2 gives the detailed procedure of the ASKC estimator. The purpose of using residuals from \mathbf{M}_i (i.e., the data other than the sample candidate) rather than X is to avoid influence of the sampled subset on the residual distribution. An additional TSSE-like procedure in step 6 may refine the scale estimate for heavily contaminated data. To improve the computational efficiency, it is unnecessary to run the TSSE-like procedure for all samples. In our case, we

only run the TSSE-like procedure for the samples with high ASKC scores $\widehat{f}_{\widehat{h}_{\theta}}^* (\geq 0.5 \widehat{f}_{\widehat{h}_{\theta}}^{\max})$.

2.4 Unifying Other Robust Estimators

Consider the RANSAC estimator [4],

$$\widehat{\theta} = \arg \max_{\theta} \mathbf{n}_{\theta},$$

and the ASSC estimator [28],

IEEE Trans Pattern Anal Mach Intell. Author manuscript; available in PMC 2010 April 21.

(15)

$$\widehat{\theta} = \arg \max_{\theta} (\mathbf{n}_{\theta} / \widehat{S}_{\theta}), \tag{16}$$

where \mathbf{n}_{θ} is the number of data points within an error tolerance (for RANSAC) or the scale of inliers (for ASSC) and S^{\wedge}_{θ} is the estimated scale of inliers.

From the above equations, we can see that RANSAC and ASSC are actually special cases of ASKC with the uniform kernel:

$$K_{U}(\beta) = \begin{cases} c_{k} & ||\beta|| \le 1, \\ 0, & \text{otherwise,} \end{cases} \quad \text{with } k_{U}(\beta) = \begin{cases} 1, & 0 \le \beta \le 1, \\ 0, & \beta > 1. \end{cases}$$
(17)

More specifically, RANSAC is one special case of ASKC with a fixed bandwidth and the uniform kernel. Replacing h_{θ} in (13) with \hat{S}_{θ} in (16), ASSC is another special case of ASKC with a variable bandwidth and the uniform kernel.

The effectiveness of the uniform kernel is low because it weights all inliers equally. In comparison, ASKC with the Epanechnikov kernel or the normal kernel (or other kernels such as triangular kernel or quartic kernel) is more effective than ASSC and RANSAC.

2.5 Differences from the pbM-Estimators

The pbM-estimators [3] and our method share some similarities because both employ the kernel density estimate techniques. However, the differences between our method and pbM are significant:

- 1. pbM places emphasis on the projection pursuit paradigm and works in the projection space along the direction of parameter vector. ASKC considers the kernel density of the mode in the residual space.
- 2. The methods in [3] find the optimal projection direction by an inner optimization and an outer optimization, i.e., they seek the mode by maximizing the density in the projection space, which, in turn, maximizes the projection index. In contrast, ASKC directly seeks the mode by maximizing the kernel density at the origin in the residual space. Our method is thus more computationally efficient (see Table 1).
- **3.** pbM employs MAD in the bandwidth estimation. However, MAD can break down when outliers occupy more than 50 percent, which will lead to oversmoothing in the kernel density estimate. ASKC uses a more robust TSSE-like scale estimator to estimate the bandwidth.
- **4.** After finding the mode with the maximum density in the projection space, pbM searches for two local minima (one on each side of the mode) by a heuristic procedure and the points with values within the two minima are treated as inliers. ASKC searches for one local minimum in the residual space and the local minima is adaptively found by a mean-shift valley procedure.

3 ASKC FOR MOTION ESTIMATION AND POSE ESTIMATION

Estimating relative camera motion from two views with a calibrated camera, and determining camera poses relating 2D images to known 3D geometry are two classical problems in computer vision. There are some existing solutions in the literature for motion estimation (e.g, [5], [14]) and pose estimation (e.g., [12]). In this section, we show a way to apply ASKC to robust motion estimation and pose estimation.

3.1 Motion Estimation

Considering a set of 3D points $\mathbf{P} = {\mathbf{p}_i}_{i=1,...,N} = {(x_i, y_i, z_i)^t}_{i=1,...,N}$ observed by a calibrated camera from two camera positions, the points \mathbf{P} are projected to two image

locations $\mathbf{Q} = {\mathbf{q}_i}_{i=1,...,N}$ and $\mathbf{Q}^* = {\mathbf{q}_i^*}_{i=1,...,N}$. The pairs ${(\mathbf{q}_i, \mathbf{q}_i^*)}_{i=1,...,N}$ form a set of 2D-2D feature matches and we have [6]:

$$(\mathbf{q}_{i}^{*})^{t} K_{2}^{-t} \mathbf{E} K_{1}^{-1} \mathbf{q}_{i} = 0,$$
(18)

where **E** is the essential matrix. K_1 and K_2 are, respectively, the intrinsic camera matrices corresponding to the two images. The essential matrix can be written as $\mathbf{E} = [\mathbf{T}]_{x}\mathbf{R}$, where $[\mathbf{T}_{x}]$ is the cross-product matrix of the translation vector **T** and **R** is the rotation matrix.

Given the camera matrices and a set of feature matches, **E** can be estimated using the nonlinear five-point algorithm [14]. The camera motion (**R** and **T**) can be recovered from **E** by Singular Value Decomposition (SVD) [6] and **T** can only be estimated up to a scale factor (i.e., $||T^*|| = 1$).

Letting $\overline{\mathbf{q}}_i = K_1^{-1} \mathbf{q}_i$ and $\overline{\mathbf{q}}_i^* = K_2^{-1} \mathbf{q}_i^*$, we define the residual of a corresponding pair as:

$$r_i(\mathbf{T}, \mathbf{R}; \overline{\mathbf{q}}_i, \overline{\mathbf{q}}_i^*) = (\overline{\mathbf{q}}_i^*)^t [\mathbf{T}]_{\mathbf{x}} \mathbf{R} \overline{\mathbf{q}}_i.$$
⁽¹⁹⁾

Assuming a unique solution exists, the robust estimate of the camera motion can be written as:

$$[\widetilde{\widetilde{\mathbf{T}}}, \widehat{\mathbf{R}}] \propto \arg\max_{\mathbf{T}, \mathbf{R}} \sum_{i=1}^{N} \frac{1}{h_{(\mathbf{T}, \mathbf{R})}} K\left(\frac{r_i(\mathbf{T}, \mathbf{R}; \overline{\mathbf{q}}_i, \overline{\mathbf{q}}_i^*)}{h_{(\mathbf{T}, \mathbf{R})}}\right).$$
(20)

3.2 Pose Estimation

The pose estimation problem can be described as: Given a set of 2D–3D correspondences $\{(\mathbf{q}_i, \mathbf{p}_i)\}_{i=1,...,N'}$, estimate the camera pose (rotation **R** and translation vector **T**) which maps the 3D reference points to 2D image coordinates. Let

$$\mathbf{R} = (r_1^t, r_2^t, r_3^t)^t \in SO(3) \text{ and } \mathbf{T} = (\mathbf{t}_x, \mathbf{t}_y, \mathbf{t}_z)^t \in \mathbf{R}^3.$$
(21)

We have [12]:

$$u_i = \frac{r_1^{T} \mathbf{p}_i + \mathbf{t}_x}{r_3^{T} \mathbf{p}_i + \mathbf{t}_z}; v_i = \frac{r_2^{T} \mathbf{p}_i + \mathbf{t}_y}{r_3^{T} \mathbf{p}_i + \mathbf{t}_z}.$$
(22)

Let $(u_i^{\dagger}, v_i^{\dagger})^{\prime}$ be an observed image point. Then, the residual of the corresponding pair is defined as:

$$r_i(\mathbf{T}, \mathbf{R}; \mathbf{q}_i, \mathbf{p}_i) = \left| u_i^{\dagger} - \frac{r_1^{t} \mathbf{p}_i + \mathbf{t}_x}{r_3^{t} \mathbf{p}_i + \mathbf{t}_z} \right| + \left| v_i^{+} - \frac{r_2^{t} \mathbf{p}_i + \mathbf{t}_y}{r_3^{t} \mathbf{p}_i + \mathbf{t}_z} \right|.$$
(23)

4 EXPERIMENTS

In this section, we evaluate the performance of the ASKC estimator employing the Epanechnikov kernel (ASKC1) and the normal kernel (ASKC2) in model fitting (including 2D line fitting and 3D plane fitting), motion estimation, and pose estimation. We compare the performance of ASKC1/ASKC2 with those of several other popular robust estimators (LMedS, RANSAC, MSAC, RESC, and ASSC). We also compare with pbM¹ in line/plane fitting and motion estimation where the application of pbM is straightforward. As the value of the error tolerance is not known a priori, we use a median scale estimator for RANSAC and MSAC (suggested by [24]). In Section 4.3, we also test the performance of RANSAC with a manually chosen optimal scale (termed as RANSAC2). The scale of inliers (σ) for the synthetic data used in this section is set to 0.2.

4.1 Line and Plane Fitting

We generate a four-line signal containing 50 data points in each line and 300 randomly distributed outliers. Thus, the total number of the data points is 500 and there are 90 percent outliers (including pseudo-outliers and random outliers) for each line at the beginning.² We apply the robust estimators to sequentially fit and extract all the four lines. The results (except for the result by RANSAC due to space limits) are shown in Fig. 3. LMedS, MSAC and RANSAC cannot successfully fit/extract any line. RESC and pbM correctly fit one line but fail to extract any line because the scale is wrongly estimated. ASSC extracts three lines but fails in one. ASKC1/ASKC2 successfully fit/extracts all four lines.

We sequentially fit and extract four 3D planes with the comparative robust estimators. Among the total 500 data points, the number of points in each of the four planes is 45. The number of random outliers is 320. Thus, the outlier percentage for each plane is 91 percent at the beginning.

Fig. 4 shows that only the proposed ASKC1/ASKC2 correctly fits and extracts all the four planes. In contrast, ASSC succeeds on two planes but wrongly fits/extracts two planes. MSAC and RESC correctly fit one plane but fail to extract any plane due to incorrect scale estimates. LMedS, RANSAC, and pbM fail to fit and extract all of the four planes.

4.2 Time Complexity

It is interesting to investigate the computational efficiency of the robust estimators. We compare the time complexity of the above methods in Table 1. All methods were coded in MATLAB except for pbM which was coded in C++. We use 3,000 random samples (1,000 for pbM) in line fitting and 6,000 random samples (2,000 for pbM) in plane fitting. We only compare the time to fit and extract the first structure (line/plane) because some methods break down after this step.

As shown in Table 1, ASKC1/2 is a little slower than LMedS, RANSAC, and MSAC but significantly faster than RESC, ASSC, and pbM. The reason that ASKC is faster than ASSC is because of step 5 in Fig. 2. ASKC is about 20–100 times faster than pbM (without considering that pbM was coded in C++ while ASKC coded in MATLAB) because it is computationally expensive to implement the step of local simplex optimization in pbM

¹The code is available at: http://www.caip.rutgers.edu/riul/research/code/PBM/index.html.

²When the inliers of one line are removed from the data, the percentage of outliers for the remaining lines decreases.

IEEE Trans Pattern Anal Mach Intell. Author manuscript; available in PMC 2010 April 21.

which is not used in ASKC. Between ASKC1 and ASKC2, we can see than ASKC1 is slightly faster than ASKC2.

4.3 Motion Estimation

We evaluate the performance of the robust estimators in motion estimation on a set of endoscopic sinus images. The endoscopic sinus image data were collected on a cadaverous porcine specimen with a calibrated monocular endoscope. The ground truth of the endoscope motion was recorded by an external Optotrak tracking system (Northern Digital Corp., Waterloo, Ontario). The endoscopic images involve a number of challenges such as low texture, specularities, illumination changes, and motion blurring (see Figs. 5a and 5b). As a result, the obtained feature-based matches between image pairs contain a large number of outliers or mismatches (see Fig. 5c). Thus, it requires the methods to be highly robust to outliers.

To obtain the quantitative results, we apply the comparative methods to 130 endoscopic sinus image pairs. The distance between the positions of the endoscopic camera in each pair of images is larger than 1 mm. The averaged number of the matched features is 922. The approximate outlier percentage is about 61.5 to 84.8 percent. We use the median, the mean, and the standard variances of the estimate errors in translation and rotation to evaluate the performance of the methods. The translation error ($\mathbf{E}_{\mathbf{T}}$) is computed as $\mathbf{E}_{\mathbf{T}} = \|\mathbf{\tilde{T}} - \mathbf{\tilde{T}}_{true}\| \times 100$, where $\mathbf{\tilde{T}}$ (with $\|\mathbf{\tilde{T}}\| = 1$) is the estimated translation direction vector and \mathbf{T}_{true} (with $\|\mathbf{\tilde{T}}_{true}\| = 1$) is the ground truth translation direction vector. Similarly, we write the rotation error (\mathbf{E}_{τ}) as $\mathbf{E}_{\tau} = \|\mathbf{\tau} - \mathbf{\tau}_{true}\| \times 100$, where τ is a quaternion representation of the rotation matrix \mathbf{R} .

Table 2 shows the quantitative results. As we can see, ASKC1/ASKC2 achieves the most accurate results among the comparative methods (including RANSAC with a manually chosen optimal scale). And between ASKC1 and ASKC2, ASKC2 generally outperforms ASKC1 in the translation and rotation estimation. LMedS, MSAC, and RANSAC with the median scale estimator achieve the worst results because the median is not robust to data with more than 50 percent outliers. The results of ASSC are better than those of RESC but less accurate than those of ASKC1/ASKC2. ASKC2 with the five-point algorithm [14] achieves better results than ASKC2* with the eight-point algorithm [6]. When the eight-point algorithm is employed for both pbM and ASKC2*, ASKC2* achieves better results than pbM.

4.4 Pose Estimation

We generate a set of synthetic data with various percentages of outliers to evaluate our methods in pose estimation. Among the total 100 2D–3D correspondences, the percentage of outliers (i.e., mismatches) is increased from 1 to 80 percent with interval being 1 percent. We repeat the experiments 20 times. We report errors in terms of rotation, translation, and RMS reprojection errors of inliers. We use $\mathbf{E}_{\tau} = \|\tau - \tau_{true}\| \times 100$ for the rotation error

measure and $E_r = \frac{2||\mathbf{T} - \mathbf{T}_{true}||}{||\mathbf{T}|| + ||\mathbf{T}_{true}||} \times 100$ for the translation error measure. The RMS reprojection error is written as

$$\sqrt{\sum_{i \in \text{inliers}} |\mathbf{q}_i^{\bullet} - \widehat{\mathbf{q}}_i|^2} / N_{\text{inliers}}$$
,

where \mathbf{q}_i^{\bullet} and $\hat{\mathbf{q}}$ are, respectively, the real reprojection point and the recovered reprojection point. The mean and the standard variance of the overall estimate errors are shown in Table 3.

From Table 3, we can see that LMedS, MSAC, and RANSAC yield the largest mean and standard variance errors. The results of RESC are more accurate than LMedS, MSAC, and RANSAC, but less accurate than ASSC, ASKC1, and ASKC2. ASSC achieves less accurate results than ASKC1/ASKC2. The results of ASKC1 and ASKC2 are similar, while ASKC2 achieves slightly better results than ASKC1.

5 CONCLUSIONS

This paper has presented a new robust estimator (ASKC) which is based on the nonparametric kernel density estimation techniques. We also provide a generalized framework in which a set of new robust estimators can be derived by employing various kernels. We show that ASKC can be treated as a generalized form of several traditional and recently developed robust estimators such as RANSAC, ASSC, and MKDE. We apply ASKC to robust motion and pose estimation and test the proposed method on both synthetic and real data. Experiments show that ASKC has achieved better results than several other popular robust estimators.

Acknowledgments

Parts of the work described in this paper appeared in [26], [27]. The authors thank the anonymous reviewers of those papers, and the *TPAMI* reviewers, Professors David Suter and Russell H. Taylor for their valuable comments and suggestions. The authors also thank Professor Darius Burschka for providing the endoscopic sinus image data. The work was supported by the US National Institutes of Health under grant number R21EB005201 and was done when H. Wang was at Johns Hopkins University.

REFERENCES

- Bab-Hadiashar A, Gheissari N. Range Image Segmentation Using Surface Selection Criterion. IEEE Trans. Image Processing 2006 July;vol. 15(no. 7):2006–2018.
- 2. Bab-Hadiashar A, Suter D. Robust Optic Flow Computation. Int'l J. Computer Vision 1998;vol. 29(no. 1):59–77.
- Chen H, Meer P. Robust Regression with Projection Based M-Estimators. Proc. IEEE Int'l Conf. Computer Vision 2003:878–885.
- 4. Fischler MA, Bolles RC. Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. Comm. ACM 1981;vol. 24(no. 6): 381–395.
- Li H, Hartley R. Five Point Algorithm Made Easy. Proc. Int'l Conf. Pattern Recognition 2006:630– 633.
- Hartley, R.; Zisserman, A. Multiple View Geometry in Computer Vision. second ed.. Cambridge Univ. Press; 2004.
- 7. Huber, PJ. Robust Statistics. Wiley; 1981.
- 8. Huber PJ. Projection Pursuit (with Discussion). Annals of Statistics 1985;vol. 13:435-525.
- Jones MC, Sibson R. What Is Projection Pursuit? (with Discussion). J. Royal Statistical Soc. A 1987;vol. 150:1–37.
- Lavva I, Hameiri E, Shimshoni I. Robust Methods for Geometric Primitive Recovery and Estimation from Range Images. IEEE Trans. Systems, Man, and Cybernetics, Part B 2008 June;vol. 38(no. 3):826–845.
- 11. Lee K-M, Meer P, Park R-H. Robust Adaptive Segmentation of Range Images. IEEE Trans. Pattern Analysis and Machine Intelligence 1998 Feb;vol. 20(no. 2):200–205.

- 12. Lu C, Hager GD, Mjolsness E. Fast and Globally Convergent Pose Estimation from Video Images. IEEE Trans. Pattern Analysis and Machine Intelligence 2000 June;vol. 22(no. 6):610–622.
- Miller JV, Stewart CV. MUSE: Robust Surface Fitting Using Unbiased Scale Estimates. Proc. IEEE Conf. Computer Vision and Pattern Recognition 1996:300–306.
- 14. Nistér D. An Efficient Solution to the Five-Point Relative Pose Problem. IEEE Trans. Pattern Analysis and Machine Intelligence 2004 June;vol. 26(no. 6):756–770.
- 15. Papenberg N, Bruhn A, Brox T, Didas S, Weickert J. Highly Accurate Optic Flow Computation with Theoretically Justified Warping. Int'l J. Computer Vision 2006;vol. 67(no. 2):141–158.
- Rousseeuw PJ, Croux C. Alternatives to the Median Absolute Deviation. J. Am. Statistical Assoc 1993;vol. 88(no. 424):1273–1283.
- 17. Rousseeuw, PJ.; Leroy, A. Robust Regression and Outlier Detection. John Wiley & Sons; 1987.
- Rozenfeld S, Shimshoni I. The Modified pbM-Estimator Method and a Runtime Analysis Technique for the RANSAC Family. Proc. IEEE Conf. Computer Vision and Pattern Recognition 2005:1113–1120.
- 19. Stewart CV. MINPRAN: A New Robust Estimator for Computer Vision. IEEE Trans. Pattern Analysis and Machine Intelligence 1995 Oct;vol. 17(no. 10):925–938.
- Subbarao R, Meer P. Heteroscedastic Projection Based M-Estimators. Proc. Workshop Empirical Evaluation Methods in Computer Vision. 2005
- 21. Terrell G. The Maximal Smoothing Principle in Density Estimation. J. Am. Statistical Assoc 1990;vol. 85:470–477.
- 22. Terrell G, Scott DW. Over Smoothed Density Estimates. J. Am. Statistical Assoc 1985;vol. 80:209–214.
- 23. Tordoff B, Murray DW. Guided Sampling and Consensus for Motion Estimation. Proc. European Conf. Computer Vision 2002:82–96.
- 24. Torr P, Murray D. The Development and Comparison of Robust Methods for Estimating the Fundamental Matrix. Int'l J. Computer Vision 1997;vol. 24(no. 3):271–300.
- 25. Wand, MP.; Jones, M. Kernel Smoothing. Chapman & Hall; 1995.
- 26. Wang H. Maximum Kernel Density Estimator for Robust Fitting. Proc. IEEE Int'l Conf. Acoustics, Speech, and Signal Processing 2008:3385–3388.
- 27. Wang H, Mirota D, Ishii M, Hager G. Robust Motion Estimation and Structure Recovery from Endoscopic Image Sequences with an Adaptive Scale Kernel Consensus Estimator. Proc. IEEE Conf. Computer Vision and Pattern Recognition. 2008
- Wang H, Suter D. Robust Adaptive-Scale Parametric Model Estimation for Computer Vision. IEEE Trans. Pattern Analysis and Machine Intelligence 2004 Nov;vol. 26(no. 11):1459–1474.
- 29. Yu X, Bui TD, Krzyzak A. Robust Estimation for Range Image Segmentation and Reconstruction. IEEE Trans. Pattern Analysis and Machine Intelligence 1994 May;vol. 16(no. 5):530–538.



Fig. 1.

The procedure of the TSSE-like scale estimator. (a) The input two-line data with 70 percent outliers. The detected peak and valley with correct parameters (b) and incorrect parameters (c).

Select a kernel for ASKC and set the number of samples *m*.

For *i* =1,...,*m*

- 1. Sample *d* values from the input data *X* yielding a subset M_i by [17] or [23] and calculate the model parameters $\hat{\theta}_i$ from M_i .
- **2.** Compute the residuals of the data points in $\overline{M}_i = X M_i$ by (2).
- **3.** Estimate an initial scale $\hat{\sigma}_0$ by (9) and an initial bandwidth \hat{h}_{θ} by (8).
- **4.** Calculate a coarse estimate of the ASKC score $\hat{f}_{\hat{h}_a}^*$ by (12).
- 5. (Optional) If the value of $\hat{f}_{\hat{h}_{\theta}}^{*}$ is less than a threshold (say, half of the current largest ASKC score $\hat{f}_{\hat{h}_{\theta}}^{\max}$), i=i+1 and go to step 1.
- 6. Run the TSSE-like procedure in the residual space. If the ratio of the kernel densities at the obtained local peak and local valley is large, estimate the scale of inliers and the corresponding bandwidth. Otherwise, i=i+1 and go to step 1.
- 7. Compute the ASKC score $\hat{f}_{\hat{h}_{a,i}}$ with the bandwidth from step 6 by (12).
- 8. If $\hat{f}_{\hat{h}_{\theta},i}$ is larger than $\hat{f}_{\hat{h}_{\theta}}^{\max}$, update $\hat{f}_{\hat{h}_{\theta}}^{\max} = \hat{f}_{\hat{h}_{\theta},i}$ and save the estimated parameters $\hat{\theta}_{i}$ and the inliers' scale estimate $\hat{\sigma}_{\theta_{i}}$.

End

Output the parameters $\hat{\theta}$ and the inliers' scale $\hat{\sigma}_{\theta}$ corresponding to the largest score $\hat{f}_{\hat{h}_{\theta}}^{\max}$.

Fig. 2. The ASKC estimation procedure.



Fig. 3. Line fitting and extraction by the robust estimators (result by RANSAC is not shown).

Wang et al.

Page 15







Fig. 5.

(a), (b) Two examples showing the challenges present in endoscopic sinus images. A snapshot showing the feature matches (c) and the matches selected by ASKC1 (d) on the left undistorted image.

TABLE 1

The Comparison on the Time Complexity of the Robust Estimators (in Seconds)

	LMedS	MSAC	RANSAC	RESC	pbM	ASSC	ASKC1	ASCKC2
Line fitting	0.50	0.57	09.0	4.08	18.63	3.95	0.67	0.82
Plane fitting	1.23	1.40	06.0	4.93	163.16	8.59	1.57	1.65

TABLE 2

NIH-PA Author Manuscript

Wang et al.

Quantitative Evaluation on 130 Pairs of Sinus Images

	Tra	nslation I	Error	R	otation E	ror
	Median	Mean	Std. Var.	Median	Mean	Std. Var.
LMedS	47.30	52.40	29.67	3.57	4.39	4.62
MSAC	32.50	40.71	28.13	6:59	8.06	6.28
RANSAC	29.38	36.59	26.85	11.38	13.34	10.57
RANSAC2	9.73	12.22	8.80	0.48	0.75	0.86
RESC	12.84	29.36	34.87	0.61	2.19	3.16
ASSC	9.61	12.33	14.73	0.41	0.83	1.99
pbM	47.18	52.98	36.99	2.40	3.41	3.10
ASKC2*	27.80	37.90	34.23	1.69	2.69	2.89
ASKC1	9.49	10.71	7.00	0.38	0.58	0.71
ASKC2	7.25	8.95	7.52	0.32	0.46	0.59

Because pbM employs the eight-point algorithm, we also test our ASKC2 with the eight-point algorithm (ASKC2*) for fairness.

Pose Estimation
Robust
easurement in
Error Me
Quantitative

	Transla	tion Error	Rotati	on Error	Reproje	ction Error
	Mean	Std. Var.	Mean	Std. Var.	Mean	Std. Var.
LmedS	36.33	57.10	10.59	26.67	25.50	42.24
MSAC	51.74	74.00	23.01	45.11	42.66	64.84
RANSAC	54.76	75.91	30.70	53.81	64.77	99.04
RESC	4.17	6.87	0.31	0.25	1.39	1.14
ASSC	0.41	0.72	0.04	0.02	0.16	0.06
ASKC1	0.37	0.64	0.04	0.02	0.15	0.06
ASKC2	0.36	0.55	0.04	0.02	0.15	0.06