IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received July 19, 2018, accepted August 21, 2018, date of publication August 30, 2018, date of current version September 21, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2867915

RES-Q: Robust Outlier Detection Algorithm for Fundamental Matrix Estimation

SUSHIL PRATAP BHARATI[1], FENG CEN[2], AJAY SHARDA[3], AND GUANGHUI WANG[1,4], (Senior Member, IEEE)

[1]Department of Electrical Engineering and Computer Science, University of Kansas, Lawrence, KS 66045, USA
[2]Department of Control Science and Engineering, College of Electronics and Information Engineering, Tongji University, Shanghai 200092, China.
[3]Biological and Agricultural Engineering, Kansas State University, Manhattan 66506, KS, USA
[4]Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China

Corresponding author: Guanghui Wang (ghwang@ku.edu)

This work was supported in part by the Kansas NASA EPSCoR Program under Grant KNEP-PDG-10-2017-KU, in part by the United States Department of Agriculture (USDA) under Grant USDA 2017-67007-26153, in part by the General Research Fund of the University of Kansas under Grant 2228901, and in part by the National Natural Science Foundation of China under Grant 61573351.

ABSTRACT Detection of outliers present in noisy images for an accurate fundamental matrix estimation is an important research topic in the field of 3-D computer vision. Although a lot of research is conducted in this domain, not much study has been done in utilizing the robust statistics for successful outlier detection algorithms. This paper proposes to utilize a reprojection residual error-based technique for outlier detection. Given a noisy stereo image pair obtained from a pair of stereo cameras and a set of initial point correspondences between them, reprojection residual error and 3-sigma principle together with robust statistic-based Q_n estimator (RES-Q) is proposed to efficiently detect the outliers and estimate the fundamental matrix with superior accuracy. The proposed RES-Q algorithm demonstrates greater precision and lower reprojection residual error than the state-of-the-art techniques. Moreover, in contrast to the assumption of Gaussian noise or symmetric noise model adopted by most previous approaches, the RES-Q is found to be robust for both symmetric and asymmetric random noise assumptions. The proposed algorithm is experimentally tested on both synthetic and real image data sets, and the experiments show that RES-Q is more effective and efficient than the classical outlier detection algorithms.

INDEX TERMS Fundamental matrix, stereo vision, robust statistics, outliers detection.

I. INTRODUCTION

The popularity of stereo cameras for recognition and localization in intelligent vehicles is increasing tremendously. Application of computer vision algorithms in object tracking [2], [3] and 3D reconstruction using stereo cameras for depth estimation [32], 3D lane detection [5], traffic sign recognition [7], pedestrian detection and tracking with night vision [40], driver assistance [9] have been extensively studied in the field of computer vision. Stereo vision, along with the sensor fusion, has been used in a number of security and safety applications as the trend of autonomous and intelligent vehicles is increasing. In order to successfully exploit the benefits of stereo cameras, one needs to estimate the fundamental matrix accurately from the noisy matching point correspondences in the stereo image pair. Thus, the obtained fundamental matrix can be successfully used to reproduce 3D reconstruction of the scene for further analysis, such as determining the distance to other objects ahead of the autonomous vehicle, buildings or obstacles ahead, traffic lights, poles, and pedestrians. However, noises in the images captured by these stereo cameras are inevitable due to unpredictable disturbances in the camera system. For example, an intelligent vehicle might be running through a rough terrain or the stereo cameras might be affected due to the presence of dust, rain, fog, irregular illumination, reflectance, occlusion, and other surrounding electromechanical or electromagnetic interferences. Similarly, instances like inconsistent feature extraction, depth discontinuities or cyclic patterns also add to this. Thus, a real-time and robust algorithm is required to filter the correct matching points from the noisy measurement for an accurate fundamental matrix estimation which helps in a precise 3D reconstruction of the surrounding environment.

The estimation of the fundamental matrix from feature point correspondences between a stereo image pair of the same scene is one of the most important steps in the field

of 3D computer vision. It is considered to be a vital step since the fundamental matrix stores all the geometric information for the relative transformation between an image pair and helps in the projective reconstruction of a scene. In general, a stereo image pair is taken with the help of two cameras with different orientations. However, the two-view images can also be taken using a single camera with appropriate motion dynamics such as rotation with translation. It is shown in [10] that an estimation of such a 3×3 fundamental matrix is governed by the epipolar geometry between the camera orientations.

It would have been a lot easier to estimate the fundamental matrix if there were no erroneous matching point correspondences between the image pair. Pragmatically, a number of mismatched points are present in two views and a reliable outlier detection algorithm should be employed to filter the outliers efficiently. The matching point pair is considered to be an outlier if it violates the epipolar constraint or comes from different 3D coordinates. The epipolar constraint is merely a geometrical constraint between the stereo image pair. It can be understood as if a feature point on one image is matched with the feature point on the other image, then both the points must lie on their respective epipolar lines. The epipolar lines are determined using the point coordinates and the fundamental matrix. Thus, the main goal of an outlier detection algorithm is to validate the match points against such constraints and also check if the point pairs come from the same 3D coordinates. Those pairs of matching points that follow these constraints are termed as inliers, otherwise, they are called outliers.

A number of algorithms have been proposed in the past to estimate the fundamental matrix from the matching point correspondences. These algorithms can be grouped into three separate categories: linear, iterative and robust estimation approach [1], [42]. In the linear estimation approach, each pair of correspondences is first fit into a fundamental matrix equation and then grouped together to form a homogeneous linear system. Next, the seven-point [10] or eight-point technique [10], [11] is used to compute the fundamental matrix. The pros of using this method are its simplicity and computational efficiency. However, the estimation of the fundamental matrix cannot be guaranteed to be exact in the presence of outliers among the point correspondences.

In the second category, an iterative method to minimize the non-linear distance function with the synchronized updates of the fundamental matrix can be used to estimate the required fundamental matrix from the matching correspondences. In this method, the geometric distance between each matched feature points and their corresponding epipolar lines (determined iteratively using the feature coordinates and the estimated fundamental matrix) is computed and summed over all the matched points. Thus, the goal of the iterative method is to minimize the sum and update the fundamental matrix in each iteration. The advantage of this approach over linear estimation approach is its meaningful geometric explanation. However, one of its cons is that the approach is still not void of noise sensitivity and performs poorly in the presence of outliers.

To alleviate the adversities of outliers, some robust estimation approaches such as RANSAC [8], LMedS [23], M-estimator [27], MLESAC [31], Guided-MLESAC [28], PROSAC [6], ARRSAC [21] and MAPSAC [29] have been proposed in the past few decades. These methods can be utilized to correctly identify and remove the outliers for an accurate fundamental matrix estimation. Such methods show improved data noise tolerance but are computationally expensive than the earlier discussed approaches. In our early studies [41], [34], a new iterative algorithm to eradicate possible outliers for an accurate fundamental matrix estimation using reprojection error, instead of the algebraic error, was developed to enhance the computational efficiency as well as match the performance of the robust estimation approaches. The paper assumed that the image noise as well as the reprojection residual error follows a Gaussian distribution. Similarly, it was shown in the paper that the outliers yield very large reprojection errors and could be successfully removed using three-sigma principle, i.e., the point correspondences lying outside three times the sigma (variance) in the Gaussian distribution of their reprojection residual error are treated as outliers. One of the major limitations of the paper was the assumption of the Gaussian noise which may not always be valid in the practical scenarios. Moreover, the paper extensively used Median Absolute Deviation (MAD) estimator to calculate the standard deviation of the reprojection residual error which is only 37% efficient in removing the outliers for the Gaussian noise model [26].

Therefore, it is very important to find an algorithm that is robust to general noise models in the stereo image pair for accurate fundamental matrix estimation and 3D reconstruction. Similarly, on the quest to find such algorithms one should also make sure that the algorithm performs with reliable efficiency. Otherwise, the unnoticed outliers would severely impact the accuracy of the estimated fundamental matrix and the intelligent vehicle shall possess inaccuracies which could be risky.

In this paper, we propose a reprojection residual error based iterative approach for robust estimation of the fundamental matrix. As a substantial extension of our previous study [41], [37], the proposed approach successfully exploits both the Gaussian, as well as the other generalized noise model, which are often present in the image pair captured by the cameras. Our approach exploits the reprojection residual error as a confidence measure of the fundamental matrix. Similarly, we will also utilize an efficient robust statistics based estimator Q_n [22], which performs much better than the MAD estimator in the Gaussian as well as other noise distributions for outlier elimination to estimate an accurate fundamental matrix for 3D computer vision tasks. Compared to the state-of-the-art outlier detection techniques, our approach is more efficient and accurate, as shown in Fig. 1 and demonstrated via extensive experiments on both synthetic and real image datasets.

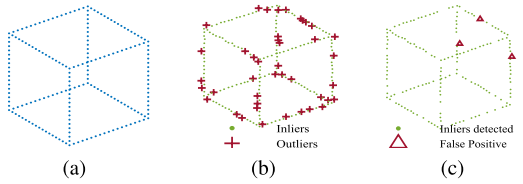

FIGURE 1. RES-Q rejects most of the outliers successfully. (a) Original simulated cube. (b) Cube with inliers and 20% outliers. (c) Resulting cube with detected inliers where most of the outliers present earlier are rejected after running through RES-Q. (Best viewed in color and digitally zoomed.)

The rest of the paper is organized as follows. Some earlier research work on outlier detection methods and the fundamental matrix estimation are presented in Section II. Section III discusses our proposed approach in details together with necessary theoretical understandings. Experimental findings are demonstrated in Section IV and V. Finally, a short conclusion is drawn in Section VI.

II. RELATED WORK

Researchers have proposed several robust outlier detection methodologies to correctly identify and eliminate the outliers in a given noisy data points for the fundamental matrix estimation. One of the most celebrated and widely used robust outlier detection methods is RANSAC [8]. In RANSAC, the fundamental matrix is initially guessed using the minimal set of points. Next, each matching feature points is tested against a hypothesized model to choose an inlier set which gives an error below a set threshold and a new fundamental matrix is determined for successive iteration. By the end of the algorithm, RANSAC chooses an inlier set with the maximum number of data points. However, RANSAC requires a preset threshold, which is hard to determine. A very low threshold might result in a very few number of inliers depending on the accuracy of the match points. Similarly, if the threshold is set to some higher value with respect to the quality of the match points, the fundamental matrix cannot be accurately estimated. These inconsistencies make RANSAC unsuitable for the scenarios where there are approximately less than 50% inliers. Moreover, the lower the number of inliers, the higher the computational cost of RANSAC due to the increased number of iterations.

One of the RANSAC based methods which improved on the computational cost is PROSAC [6]. This improvement is possible as PROSAC keeps track of the quality measures of the point correspondences while determining the fundamental matrix. This technique is absent in RANSAC. In addition, the matches are sorted in their non-increasing order of quality scores and subsets of seven points are progressively chosen unless a minimum number of inliers are found to estimate the fundamental matrix. Since the algorithm starts with the match points having the best quality scores, PROSAC converges in fewer iterations than RANSAC. Though PROSAC is comparatively faster than RANSAC, it undergoes the same limitations as in RANSAC because it requires a preset threshold to vote the match points as inliers.

The other method MLESAC [31], instead of maximizing the number of inliers, utilizes a median-based approach and a different cost function, unlike RANSAC. Inliers are assigned a fitness score whereas the outliers are assigned a constant weight. Next, expectation maximization is performed to maximize the maximum likelihood estimates of the normal distribution for the inliers and the uniform distribution for the outliers. The solution that produces the least median of residuals is considered to be the set of inliers. However, MLESAC possesses the same limitations as in RANSAC and also requires the noise parameters as a prior knowledge.

Similarly, MAPSAC [29] tried to improve MLESAC using posterior estimation maximization of the fundamental matrix and the matching point correspondences under Bayesian statistics. MAPSAC uses a new evaluation technique to determine the consistency of the solution. Guided MLESAC [28] replaced the random search in MLESAC or RANSAC with a guided search and significantly reduced the required number of iterations. ARRSAC [21] removed the outliers in real-time thus making the algorithm computationally inexpensive than RANSAC. To do so, ARRSAC claimed a new framework for the real-time robust estimation by utilizing the efficient adapting techniques for faster performance over a wide range of the inlier ratios.

In order to solve the problem of automatically determining the preset threshold for separating inliers from outliers, a value proportional to the median of the residuals was used as a threshold in LMedS [23]. LMedS also elegantly depicted the use of median as a robust outlier detection estimator [13], [14] and accurate fundamental matrix estimator.

Recently, the machine learning algorithms, as well as the neural networks, have been extensively used for efficient outlier detection and robust fundamental matrix estimation. A one-class support vector machine-based pre-selection algorithm for matching the correspondences obtained using SIFT [15], together with the maximization of a soft decision objective function to refine the obtained inlier set, was used to estimate the fundamental matrix in [39]. A robust estimation technique least trimmed squares (LTS) regression was used in [25] to track the outliers which deviate away from the majority linear model. This technique was successfully applied to remove the outliers resulting from the occlusion. Cluster-based outlier removing methodology was proposed in [38] that solved outlier detection problems in complex datasets with an ample of clusters and varied densities. An unsupervised boosting approach was tested in [24] to improve the accuracy of the ensemble outlier detection algorithms. Moreover, autoencoder ensembles were used in [4] for unsupervised outlier detection. However, one of the major problems with such approaches is that they are extremely sensitive to noise and often require a lot of datasets to train. Similarly, as the model hyper parameters increase, there is a significant impact on their performance speed which makes them unsuitable for real-time applications like intelligent vehicles. Therefore, a robust statistics-based

model is required in order to perform with minimal payload and apt accuracy.

In addition, some robust feature matching and registration algorithms have been studied in the literature, including vector field consensus (VFC) [18], regularized VFC [17], locally linear transforming (LLT) [19], and L2E estimator [16]. More robust estimation techniques can be found in [22], [30], [20], and [35]. The motivation to use such robust estimators is that they are simplistic, computationally inexpensive and robust as demonstrated by their influence function and 50% breakdown point in [23]. In our paper, we propose to employ Q_n estimator as it is experimentally proven to be the most efficient for Gaussian as well as non-Gaussian noises.

III. PROPOSED APPROACH

A. FUNDAMENTAL MATRIX ESTIMATION AND ALGEBRAIC ERROR

Given a stereo image pair, the intrinsic projective geometry between the pair of images is captured by its epipolar geometry. Epipolar geometry between two-view images is dependent only on the camera's intrinsic parameters and relative pose. Since it does not depend on the scene structure, the fundamental matrix $\mathbf{F}$ is able to encapsulate the epipolar geometry between the stereo image pair. The fundamental matrix is mathematically represented as a rank 2 matrix of size 3×3.

The fundamental matrix can be estimated from a set of point correspondences between a pair of images. Provided a stereo image pair A and B such that $\mathbf{x}_i \in A$ and $\mathbf{x}'_i \in B$ be two i^{th} matching point correspondences represented using homogeneous coordinates, $\mathbf{F}$ should satisfy the following equation [10].

$$\mathbf{x}_i'^T \mathbf{F} \mathbf{x}_i = 0 \tag{1}$$

Each matching correspondence point pair contributes to a single linear constraint. Since $\mathbf{F}$ is a 3×3 rank 2 matrix, defined up to a scale, eight pairs of matching point correspondences are sufficient to determine $\mathbf{F}$ using the eight-point linear algorithm [10], [11]. However, least square estimation technique can also be utilized if there are more than eight point pair correspondences available.

In order to evaluate the correctness of the estimated fundamental matrix $\mathbf{F}$, we must determine an error measurement scheme that aids to further correction of the fundamental matrix estimation technique. Errors in the fundamental matrix estimation occur due to erroneous matching point pairs. One of the most widely used error measurement criteria is an algebraic error. An algebraic error e_{alg} can be calculated as

$$e_{alg}(i) = \mathbf{x}_i'^T \mathbf{F} \mathbf{x}_i \tag{2}$$

Since the matching point correspondences are faulty, they do not satisfy equation (1) and result in some error as calculated in equation (2).

B. REPROJECTION RESIDUAL ERROR

Though the calculation of algebraic error is rather simple, it does not contain any clear geometric meaning. Also, equation (1) provides a necessary but insufficient condition for correct matches as the point in $\mathbf{x}_i \in A$ may have the corresponding match points $\mathbf{x}'_i$ throughout the epipolar line defined by $\mathbf{F}\mathbf{x}_i$ in B. Therefore, in this paper, we propose to use the reprojection residual error to overcome such ambiguity.

First, we estimate the fundamental matrix $\mathbf{F}$ from available matching point correspondences between the stereo image pair. Next, a pair of camera matrices $\mathbf{P_1}$ and $\mathbf{P_2}$ are recovered from $\mathbf{F}$ for both the cameras using the following equation as described in [10].

$$\mathbf{P_1} = \left[\mathbf{I} \mid \mathbf{0}\right] \tag{3}$$

$$\mathbf{P_2} = \left[\left[\mathbf{e}'\right]_\times \mathbf{F} + \mathbf{e}'\mathbf{q}^T \mid \beta \mathbf{e}'\right] \tag{4}$$

where $\mathbf{e}'$ is the epipole in the second image such that $\mathbf{e}'^T\mathbf{F} = 0$, $\mathbf{q}$ is an arbitrary three dimensional vector and β is any nonzero scalar ($\beta \neq 0$).

Based on the recovered projection matrices, a triangulation algorithm [12] is used to obtain a perspective 3D reconstruction of the matching correspondences. Thereafter, we use $\mathbf{P_1}$ and $\mathbf{P_2}$ to reproject the reconstructed 3D points back to the respective image planes. Let $\hat{\mathbf{x}}_i$ and $\hat{\mathbf{x}'}_i$ be the reprojected image point pairs for $\mathbf{x}_i$ and $\mathbf{x}'_i$ respectively, the reprojection residual for point i in the two images can be calculated as $\mathbf{r}_i = \begin{bmatrix} \Delta u_i \\ \Delta v_i \end{bmatrix}$ and $\mathbf{r}'_i = \begin{bmatrix} \Delta u'_i \\ \Delta v'_i \end{bmatrix}$, where Δu and Δv are the reprojection errors along the two coordinate axes, respectively. Then, the reprojection residual error can be defined as the following vector by collecting the residuals of all the corresponding matches and their reprojections in the two images.

$$e_{rr} = \left\{\mathbf{r}_i, \mathbf{r}'_i \mid i = 1, \ldots, N\right\} \tag{5}$$

where N is the total number of correspondences.

C. OUTLIER DETECTION AND REMOVAL POLICY

Since erroneous matching point correspondences result in an inefficient fundamental matrix estimation, it is necessary to determine a suitable outlier detection policy to eliminate these faulty matches. In this section, we will discuss the outlier detection and removal policies for a symmetric Gaussian noise model as well as a generalized random noise model to successfully eradicate the outliers present in the image pairs. Before moving on to these cases, it is vital to understand that the reprojection residual error undergoes similar noise distribution as present in the ordinary image [41] as shown in the simulation result for a Gaussian inlier noise and random outlier noise in Fig. 2. The reprojection residual error plot is simply a histogram plot of the vector as defined in equation (5).

1) GAUSSIAN NOISE MODEL

One of the most widely adopted noise models in the field of computer vision and image processing is the Gaussian

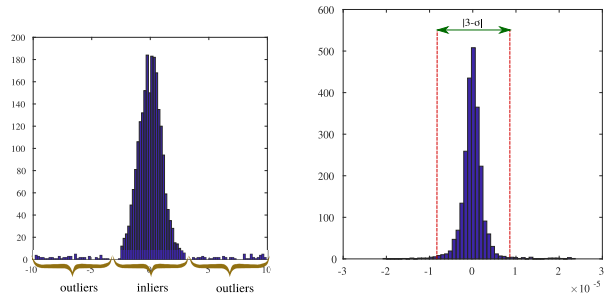

FIGURE 2. (a) Histogram of the added Gaussian noise (inliers) along with large standard deviation noise (outliers). (b) Histogram of the reprojection residual error that follows Gaussian distribution and 3-σ range for inliers classification.

noise model. The credit can also be given to the central limit theorem. Gaussian model has a very simplistic representation as it can be analyzed using only two parameters: mean μ and variance σ^2. The probability density function of a Gaussian distribution comprising of data points y is given by the following equation

$$f(y|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y-\mu)^2}{2\sigma^2}} \tag{6}$$

Since the outliers always tend to have a larger standard deviation σ, through extensive experimentations we have found that the reprojection residual error of the outliers also tend to have a larger standard deviation in comparison to that of the inliers, as demonstrated in Fig. 2. This finding, supported by our earlier research work [34], [41], helps us to correctly estimate and remove the outliers using 3-σ principle. According to the Gaussian distribution, 99.7% of the inliers should be within 3-σ of the mean. This principle explains that the matching points with the reprojection residual error larger than three times the variance (also known as sigma scaling factor (σ_f) calculated from $e_{rr}(i)$ for all the corresponding point pairs) can be classified as outliers. Thus, 3-σ can be considered as a suitable threshold to filter the inliers from the outliers effectively. Hence, the only policy needed to determine now is how to find the required standard deviation from the calculated residual errors.

We have based our approach on robust statistics [23] as they are proven to be the most reliable outlier detection method. Although the median absolute deviation (MAD) was extensively used in our previous work [34], [41], the study [22] suggests that the Q_n estimator is a more robust statistic that experimentally outperforms MAD. Thus, we decided to adopt the Q_n estimator in this paper which can be computed using the following equation.

$$Q_n = \sigma = c\{|x_a - x_b|; a < b\}_{(m)} \tag{7}$$

where x_a and $x_b \in x$ (data points), c is a constant factor and $m = \binom{v}{2} \approx \binom{n}{2}/4$, $v = \binom{n}{2} + 1$. Thus, the formula suggests that we need a constant factor c to scale for the different noise distributions and m^{th} order statistic of the $\binom{n}{2}$ interpoint distances. The constant factor c for the Gaussian noise model is determined to be 2.2219.

Once the Q_n estimator is calculated, the outliers can now be easily classified using

$$\mathcal{O} = \{(x_i, x_i') : |e_{rr}(i)|_{std} > 3\sigma\} \tag{8}$$

where $\mathcal{O}$ is an outlier set and $|e_{rr}(i)|_{std}$ is a standardized (mean = 0) reprojection residual error for x_i and x_i'.

2) GENERALIZED NOISE MODEL

Since the noises in the stereo image cannot be always assumed to follow a symmetric Gaussian distribution in the real-world, it is essential to determine an outlier detection policy for asymmetric noise distributions as well. Thus, we experimented our algorithm with the stereo images populated with random noises corresponding to a lower standard deviation (typically within ± 3) as inliers and a larger standard deviation (typically ± 4 to ± 10) as outliers. Now, the theoretical understanding follows the same as discussed earlier because the inlier distribution is bounded and the outliers can be successfully filtered using 3-σ principle.

D. OUTLINE OF RES-Q ALGORITHM

The implementation details of the above proposed robust algorithm RES-Q is summarized as below.

Algorithm 1 Robust Outlier Detection Using RES-Q

Input: stereo images
Output: optimal fundamental matrix

1. Estimate an initial fundamental matrix $\mathbf{F_{int}}$ using all the points in the stereo image
2. Estimate the camera parameters $\mathbf{P_1}$ and $\mathbf{P_2}$ from $\mathbf{F_{int}}$
3. 3D reconstruct the point pairs using triangulation
4. Reproject the 3D cloud points using $\mathbf{P_1}$ and $\mathbf{P_2}$
5. Calculate the reprojection residual error e_{rr} from reprojection in step 4 with input image pairs
6. Use 3-σ principle to remove the outliers (use equation 8)
7. Re-estimate the fundamental matrix $\mathbf{F_{est}}$ using detected inliers and repeat steps 3 - 7 one more time to refine the inliers
8. Final $\mathbf{F_{est}}$ is the optimal fundamental matrix for the input stereo images

IV. EVALUATIONS ON SYNTHETIC DATA

Our algorithm is evaluated on a computer-simulated synthetic data and compared qualitatively as well as quantitatively against similar algorithms in the literature. During the simulation, we generated nearly 1200 space points to form a cube of $50 \times 50 \times 50$ as shown in Fig. 3. Then, an image pair of size 800×800 is produced using camera matrices $\mathbf{P_1}$ and $\mathbf{P_2}$. During the simulation, the focal length of the camera is set to 800 and is kept 60 units apart from the cube and angular rotations of the camera are set to [45, 30, 75]

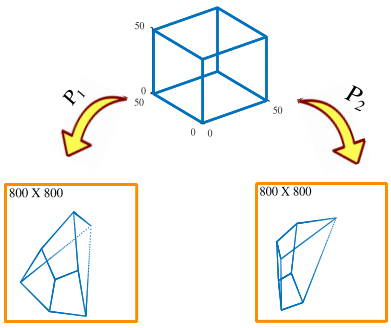

FIGURE 3. Simulated 3D cube and projected stereo image using camera parameters P_1 and P_2 respectively.

and [45, 15, 0] to generate a pair of stereo images. Inliers were simulated by adding Gaussian noise of 0 mean and 3 pixels standard deviation whereas outliers were randomly introduced by adding noise with the larger standard deviation (up to ± 10). To best observe the performance of the algorithms on different amount of random noises, we varied the outlier percentage from 5% to 25% during the experiments. Every algorithm was executed 500 trials for each outlier percentage value to best generalize the performance of the compared algorithms.

A. MEAN REPROJECTION RESIDUAL ERROR

We analyzed our algorithm with the other state-of-the-art algorithms and plotted the mean reprojection residual error plot to determine the merits of the algorithms based on the different reprojection residual error at different outlier levels. In Fig. 4(a) and Fig. 5(a), we plot the mean reprojection residual error curve, where the algorithm is executed to determine the final set of inliers using final estimated fundamental matrix and the mean reprojection residual error is calculated for these inliers with their corresponding points in an input image over 500 trials. If the fundamental matrix determined by the algorithm is precise, the algorithm shall be robust to the outliers and effectively detect the inliers with less error. Thus, it can be verified that the robust algorithms should show less mean reprojection residual error. In the curve, we can see that both the proposed algorithm and the MAD [41] show minimal reprojection residual error compared to the other algorithms. LTS [25] though starts with a lower mean reprojection residual error, further increases significantly when the percentage of the outliers in an input image is increased thus making it less robust to heavier noises. A similar pattern can be observed in RANSAC [8], MSAC [29], and LMedS [23] except for the fact that they are not as robust as LTS, MAD and our approach. The gradual increase in the curve can be accounted for the fact that as the outlier percentage increases, the algorithm finds harder to determine the accurate fundamental matrix which would result in erroneous inliers.

B. MEAN STANDARD DEVIATION OF REPROJECTION RESIDUAL ERROR

Mean standard deviation of the reprojection residual error is plotted by calculating the average standard deviation of the reprojection residual error for each percentage of the outliers added over 500 independent trials. Since the standard deviation measures the variation of the reprojection residual error from its mean over multiple trials, it determines the stability, reliability and reproducibility of the algorithm. Hence, if the algorithm tends to show the constant mean standard deviation of the reprojection residual error over different levels of outliers, we can conclude that the performance of the algorithm is noise independent and static. The experimentation plots in Fig. 4(b) and Fig. 5(b) clearly demonstrate that MAD and our approach have almost constant mean standard deviation for the different outlier percentages. LTS and LMedS tend to deviate from their previous mean reprojection residual errors when the outlier level increases, thus making them noise sensitive. Similarly, RANSAC and MSAC clearly are not able to cope with the changing outlier ratios in an input image. Ours and MAD show constant plot because these algorithms remove the outliers by adaptively choosing σ using 3-σ principle, unlike other methods which use a fixed threshold for every outlier levels, thus making them unable to choose the true inliers with the varying percentages of outliers.

C. MEAN PRECISION

To measure the relevancy of the inliers detected by the algorithms, we need to determine a measure that accurately scores the algorithm based on a number of true inliers/outliers detected. For example, if a particular algorithm classifies a handful amount of inliers, we cannot claim the algorithm to be a reliable algorithm merely based on the number of inliers detected, unless we verify the authenticity of the detected inliers with the ground truth. This is where the precision of the algorithm comes into play, which carefully scores the algorithm up if the detected inliers match the ground truth and thumbs down otherwise. Let TP denote true positives (detected inliers/outliers are true inliers/outliers) and FP denote false positives (detected inliers are true outliers), then the precision of the algorithm is calculated as $\frac{TP}{TP+FP}$. Thus, it can be deduced from the definition that the higher the precision, the better the algorithmic performance.

The precisions of different algorithms are plotted in Fig. 4(c) and Fig. 5(c). It is evident that our approach has higher precision compared to the other approaches for the most of the outlier percentages. LTS tends to show the highest precision for lower levels of outliers, while the precision drastically decreases in a near linear fashion for higher outlier levels. This can be accounted for the fact that as the number of outliers increases, LTS finds it hard to segregate the true inliers from outliers. This can also be observed in Fig. 4(a) and Fig. 5(a) where mean reprojection residual error for LTS gradually increases due to the fact that the outliers being detected as the inliers. MAD shows considerable precision but does not perform better than our approach. This is one of the major advantages of using RES-Q over MAD algorithm. However, we observed that MAD performs slightly better than RES-Q at higher outlier level, i.e., 25%. RANSAC and MSAC perform very similar but have lower precisions.

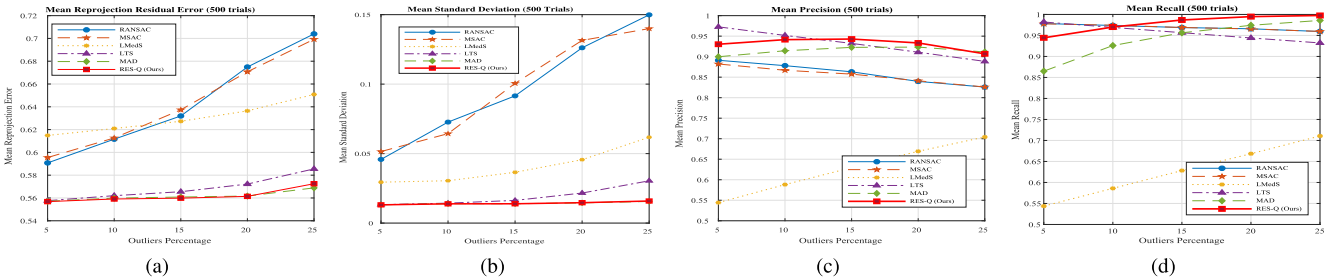

FIGURE 4. Synthetic experimentation plots (inliers modeled as a random symmetric noise) comparing algorithms for 500 independent trials. (a) Mean reprojection residual error. (b) Mean standard deviation of reprojection residual error. (c) Mean precision. (d) Mean recall. (Best viewed in color).

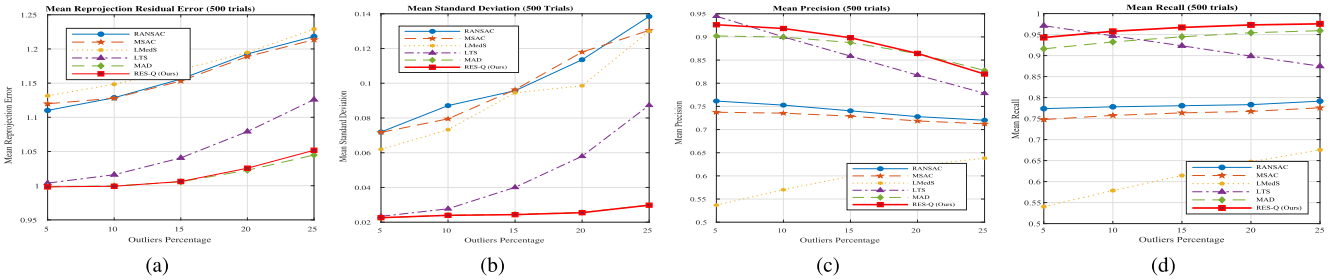

FIGURE 5. Synthetic experimentation plots (inliers modeled as a random asymmetric noise) comparing algorithms for 500 independent trials. (a) Mean reprojection residual error. (b) Mean standard deviation of reprojection residual error. (c) Mean precision. (d) Mean recall. (Best viewed in color).

Again, this limitation is due to having a constant threshold for every outlier levels and no adaptive outlier rejection strategy. In addition, LMedS plot shows that LMedS is not a suitable algorithm for lower outlier percentage and is not able to outperform the compared algorithms for any higher outlier percentages. It can be concluded that, for varying outlier levels, our approach is more precise to use with the least concern of detecting false positives.

D. MEAN RECALL

While precision score determines how precise the algorithm is on classifying correct inliers/outliers with respect to the ground truth, it does not fully help us to understand if we have correctly predicted all the inliers as well as outliers present. This necessitates the evaluation of the algorithms on recall. Recall helps to understand how the algorithms performs in determining all the inliers/outliers correctly. Mathematically, let TP denote true positives (detected inliers/outliers are true inliers/outliers) and FN denote false negatives (detected outliers are true inliers), then the recall of the algorithm is calculated as $\frac{TP}{TP+FN}$. Thus, it can be deduced from the definition that an algorithm with higher recall is more reliable.

The corresponding recall curves are plotted in Fig. 4(d) and Fig. 5(d). The plots clearly depict that our algorithm has higher mean recall compared to other competitive algorithms. Although LMedS and RANSAC perform better at 5% outlier level, RES-Q maintains higher recall at various outlier levels in overall. MAD performs better at higher outlier levels than the others, but RES-Q clearly outperforms all the compared algorithms. It is also important to note that RES-Q has maximum attainable recall, i.e., 1 at 25% outlier level for symmetric noise and near 100% recall for asymmetric noise. Thus, RES-Q is not only precise but also accurately predicts most of the inliers/outliers irrespective of the noise patterns.

E. COMPUTATIONAL COMPLEXITY ANALYSIS

Let $\mathbf{N}$ be the number of iterations performed, $\mathbf{D}$ be the number of data points consisting of both inliers and outliers, $\mathbf{F}$ be the number of features (for example, 8 features for estimating the fundamental matrix using the normalized eight point algorithm [10], [11]), $\mathbf{C_{fit}}$ be the time complexity for determining the fundamental matrix, $\mathbf{M_L}$ be the time complexity for maximum likelihood operations to determine the most likely set of inliers in MSAC and $\mathbf{T}$ be the computational cost for triangulation, we can compare the total worst case scenario computational costs for different algorithms. Since RANSAC fits the detected inliers in every iteration to estimate the fundamental matrix to see the fit according to a preset threshold, its complexity is $\mathcal{O}(\mathbf{N}(\mathbf{C_{fit}} + \mathbf{D}))$. MSAC performs on a common ground with RANSAC except for calculating a maximum likelihood fit for the final estimation of the inlier set, its computational complexity is $\mathcal{O}(\mathbf{N}(\mathbf{C_{fit}} + \mathbf{D}) + \mathbf{M_L})$, slightly higher than RANSAC. Both LMedS and LTS have the same order of complexity $\mathcal{O}(\mathbf{NF^2D})$ because of their iterative machine learning algorithmic approach. However, MAD and RES-Q, based on the previously proposed Algorithm 1, calculate the fundamental matrix twice, performs triangulation twice and computes the reprojection residual error twice which is dependent on $\mathbf{D}$, it has a complexity of $\mathcal{O}(\mathbf{C_{fit}} + \mathbf{DlogD} + \mathbf{T})$. A key observation here is that MAD and RES-Q have their order of complexities

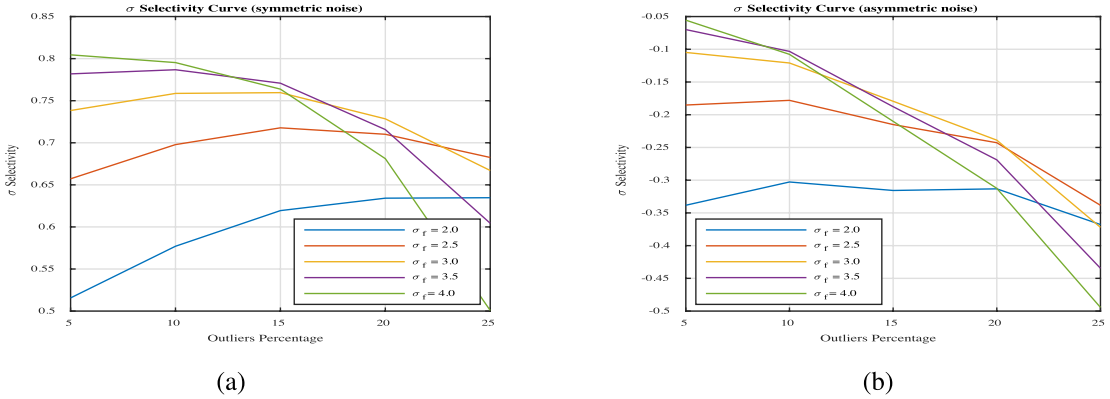

FIGURE 6. RES-Q sigma selectivity versus various outlier levels for (a) symmetric noise and (b) asymmetric noise.

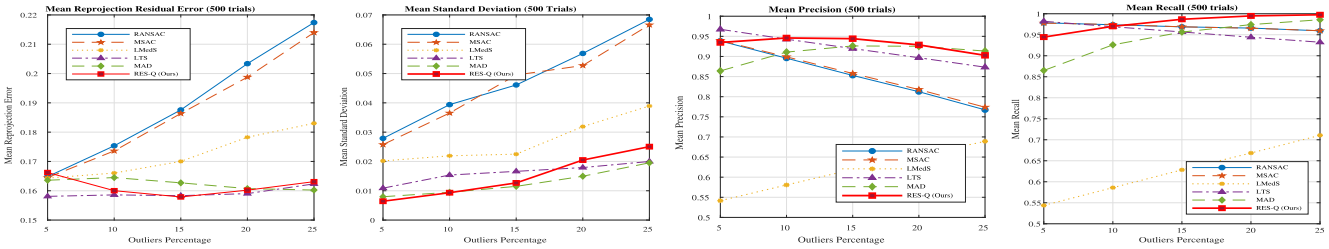

FIGURE 7. Real data experimentation plots comparing algorithms for 500 independent trials on Merton College II dataset. (a) Mean reprojection residual error. (b) Mean standard deviation of reprojection residual error. (c) Mean precision. (d) Mean recall. (Best viewed in color).

independent to the number of iterations **N** thus making them computationally more efficient than the classical approaches.

F. DETERMINATION OF SIGMA SCALING FACTOR

We experimented with several sigma scaling factor (σ_f) versus outlier percentages to best determine the inliers with an apt precision. Since the precision of determining as well as eliminating the outliers should be higher for an algorithm and the reprojection residual error should be lower for better performance, we define sigma selectivity as a logarithmic ratio of precision p to reprojection residual error e_{rr}. The selectiveness of σ is mathematically calculated as $\sigma_{sel} \frac{log(p)}{log(e_{rr})}$. Fig. 6 depicts sigma selectivity against different levels of outliers. It can be observed that lower sigma scaling factors (2.0 and 2.5) has less selectivity for the most of the outlier levels. In addition, higher scaling factors (4.0 and 3.5), though have higher selectivity at lower outlier levels, quickly drop down with the increase of outlier levels. Thus, it can be safely concluded that sigma scaling factor of 3.0 is the most appropriate value as it has steady selectivity across different outlier percentages and is chosen both for symmetric and asymmetric noise in our experiments for the RES-Q approach.

V. EVALUATIONS ON REAL DATA

We extensively evaluated the RES-Q algorithm against other competing algorithms in the real data sets. For this purpose, four datasets: Lion [36], Fountain [33], Merton College I and Merton College II from Oxford Multiple View VGG dataset[1] were chosen. Lion dataset consists of 3204 matching point correspondences and 5% random outliers were added for the experimentation. Fountain dataset consists of 4219 matching point correspondences and 10% outliers were added. Similarly, Merton College I and Merton College II consist of 383 and 344 matching point correspondences and 20% and 25% outliers were added respectively. All the stereo image pairs in the datasets have a resolution of 1024 × 768 and the calibrated camera parameters were provided along with the datasets. For the generalization purpose, the algorithms were executed independently for 500 times on each dataset and the results are reported.

A. MEAN REPROJECTION RESIDUAL ERROR

In Fig. 7(a), we plot the mean reprojection residual error curve. This curve determines the number of correct inliers detected as the reprojection residual error for the true positives contributes less to the reprojection residual error, whereas, that of the false positives contributes significantly. Thus, a good outlier rejection algorithm should have lower mean reprojection residual error. In the plot, we can observe that RES-Q, LTS and MAD have considerably lower mean reprojection residual error than that of RANSAC,

[1] http://www.robots.ox.ac.uk/vgg/data/data-mview.html

MSAC or LMedS. Also, RES-Q and LTS outperform MAD in most of the outlier levels. Though LTS did not perform as good as MAD and RES-Q in the synthetic experiments, the real dataset experiments show that LTS could be a suitable choice whereas MAD may deviate from the synthetic experimental findings. Based on the experiments, we can conclude that RES-Q performs competitively better in both of the experiments, and thus, proving its robustness at several outlier levels.

B. MEAN STANDARD DEVIATION OF REPROJECTION RESIDUAL ERROR

To observe the noise stability, outlier level adaptability and reproducibility of the compared algorithms, the standard deviation of the reprojection residual errors over several outlier percentage values is evaluated over 500 independent trials to plot the mean standard deviation of the reprojection residual errors in Fig. 7(b) as explained in the synthetic experiments in the previous section. The plot clearly indicates that the mean standard deviation of the reprojection residual errors is almost constant for RES-Q, MAD and LTS compared to RANSAC, MSAC or LMedS. This observation also matches with the plots in the synthetic experiments. Though MAD and RES-Q performed almost similar in the synthetic experiments as their plots overlapped, real data experiments show that RES-Q could be unstable to noises compared to MAD and LTS for higher outlier levels.

C. MEAN PRECISION

The precision plot over varying outlier percentages helps to determine the true analysis of the performance of the various competing algorithms. Since the precision plot carefully accounts for the ground truth match with the set of classified inliers by the various algorithms and penalizes the algorithm for every misclassification, the higher the precision curve, the more reliability of an algorithm. In the precision plot Fig. 7(c), RES-Q clearly outperforms all the competing algorithms. LTS, though has a slightly higher precision for the lowest outlier level (5%) (also similar findings in the synthetic experiments), the precision falls below RES-Q and MAD approaches for higher outlier levels. This can also be verified by the increasing mean reprojection residual errors for LTS in Fig. 7(a) due to false classifications. MAD shows a good precision for higher outlier levels but is not able to perform with the utmost precision compared to LTS or RES-Q at lower outlier levels. This was also outlined in the precision plot for synthetic experiments and argued to be the key disadvantage of using MAD over RES-Q. RANSAC and MSAC due to their inadaptive outlier rejection strategy cannot outperform adaptive algorithms. LMedS does not seem to perform with much precision at any outlier levels. Thus, RES-Q is a better choice for precise outlier rejection scheme for a robust fundamental matrix estimation.

To further demonstrate the performance of RES-Q, the histogram distribution of the reprojection residual error in the Lion dataset before and after running RES-Q is plotted in Fig. 8. As discussed earlier, the reprojection residual errors follow Gaussian distribution and RES-Q is able to reject most of the outliers using 3-σ principle for the fundamental matrix estimation which can also be verified in Fig. 8.

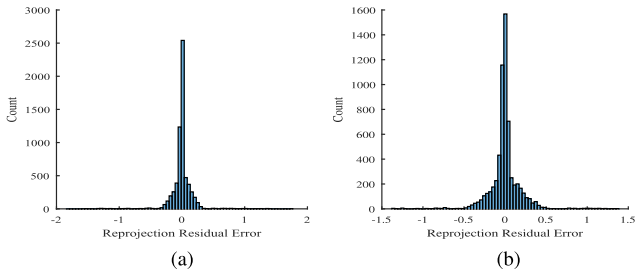

FIGURE 8. Histogram of reprojection residual error on Lion dataset with 10% outlier before (a) and after (b) RES-Q.

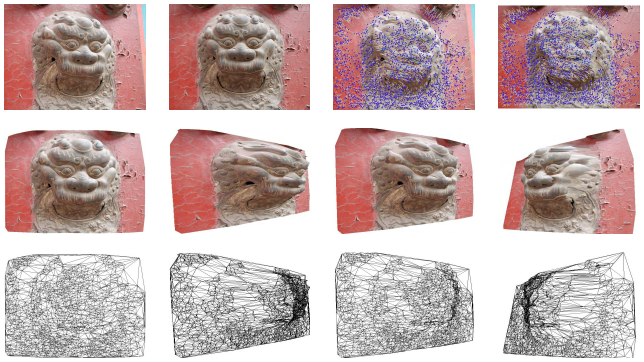

FIGURE 9. Reconstruction results of Lion dataset with 5% outlier. (First row) left two are original stereo images and the right two are overlaid with the matching features obtained by the RES-Q. (Second row) the reconstructed VRML model of the scene shown from different viewpoints with texture mapping. (Last row) the corresponding triangulated wireframe of the VRML models.

D. MEAN RECALL

Although mean precision significantly helps to compare the algorithms on how precise are their inlier/outlier predictions, mean recall is also an important metric to evaluate the algorithms on the accuracy of their predictions. In the recall plot Fig. 7(d), RES-Q is found to outperform all the compared algorithms in majority of the outlier level percentages. RANSAC and MSAC cannot maintain a good recall at higher outlier levels due to their inadaptable algorithmic approach. LMeds shows an overall poor performance. RANSAC, MSAC and LTS have a good recall at 5% outlier level, which is similar to the findings in the synthetic experiment, however, they are not able to keep up with increasing outlier levels. It is evident that RES-Q outperforms all the competing algorithms and has a 100% recall at 25% outlier level. Therefore, RES-Q can be concluded to be more reliable at different outlier ratios.

E. QUALITATIVE EVALUATION RESULTS

In order to further verify the robustness of our approach, different outlier levels were added to the real datasets (5% to Lion, 10% to Fountain, 20% to Metron College I

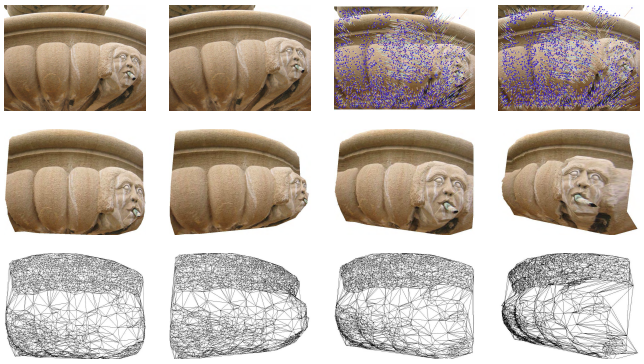

FIGURE 10. Reconstruction results of Fountain dataset with 10% outlier. (First row) left two are original stereo images and the right two are overlaid with the matching features obtained by the RES-Q. (Second row) the reconstructed VRML model of the scene shown from different viewpoints with texture mapping. (Last row) the corresponding triangulated wireframe of the VRML models.

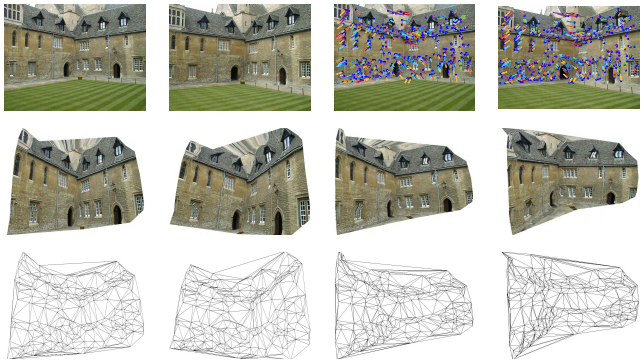

FIGURE 11. Reconstruction results of Merton College I dataset with 20% outlier. (First row) left two are original stereo images and the right two are overlaid with the matching features obtained by the RES-Q. (Second row) the reconstructed VRML model of the scene shown from different viewpoints with texture mapping. (Last row) the corresponding triangulated wireframe of the VRML models.

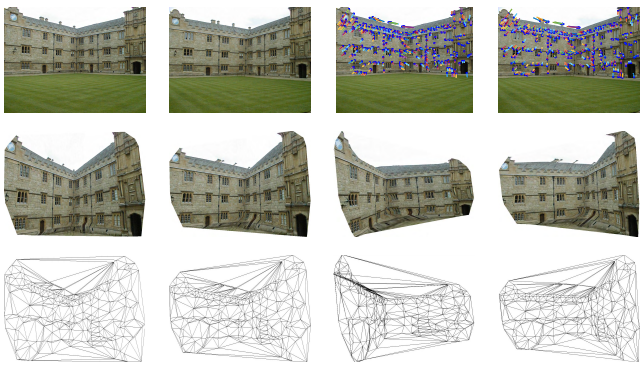

FIGURE 12. Reconstruction results of Merton College II dataset with 25% oultier. (First row) left two are original stereo images and the right two are overlaid with the matching features obtained by the RES-Q. (Second row) the reconstructed VRML model of the scene shown from different viewpoints with texture mapping. (Last row) the corresponding triangulated wireframe of the VRML models.

and 25% to Merton College II) and experimented with RES-Q to see the reconstructed VRML models and the corresponding wireframes. Figs. 9, 10, 11 and 12 show that RES-Q successfully discards the outliers and recovers the Euclidean structure of the scene for different levels of outliers. As the levels of outlier increase, the VRML models show some instances of false reconstruction, however, most of the parts are accurately reconstructed and visually plausible.

VI. CONCLUSION

In this paper, we have demonstrated that the robust statistics-based method can be used to determine and remove the outliers present in the corresponding matching points between a pair of stereo images. Through extensive synthetic and real data experiments, we have verified that the reprojection residual error based technique is more robust than the algebraic error based approaches. Furthermore, we have shown that Q_n estimator together with the 3-σ principle in our RES-Q algorithm can efficiently detect outliers at different levels and accurately estimate the fundamental matrix for successful reconstruction of the given scene. In addition, several experiments have been carried out to show that RES-Q performs equally well for both symmetric and asymmetric random noises. We have also demonstrated that one of the major advantages of using RES-Q over MAD is an improved precision as well as recall which is vital in practical applications. The greater precision and recall of RES-Q over MAD was further verified by experimenting them with various synthetic as well as real datasets against several outlier ratios.

REFERENCES

[1] X. Armangué and J. Salvi, "Overall view regarding fundamental matrix estimation," *Image Vis. Comput.*, vol. 21, no. 2, pp. 205–220, 2003.

[2] S. P. Bharati, S. Nandi, Y. Wu, Y. Sui, and G. Wang, "Fast and robust object tracking with adaptive detection," in *Proc. 28th Int. Conf. Tools Artif. Intell. (ICTAI)*, Nov. 2016, pp. 706–713.

[3] S. P. Bharati, Y. Wu, Y. Sui, C. Padgett, and G. Wang, "Real-time obstacle detection and tracking for sense-and-avoid mechanism in UAVs," *IEEE Trans. Intell. Veh.*, vol. 3, no. 2, pp. 185–197, Jun. 2018.

[4] J. Chen, S. Sathe, C. Aggarwal, and D. Turaga, "Outlier detection with autoencoder ensembles," in *Proc. SIAM Int. Conf. Data Mining*, 2017, p. 9.

[5] H.-Y. Cheng, B.-S. Jeng, P.-T. Tseng, and K.-C. Fan, "Lane detection with moving vehicles in the traffic scenes," *IEEE Trans. Intell. Transp. Syst.*, vol. 7, no. 4, pp. 571–582, Dec. 2006.

[6] O. Chum and J. Matas, "Matching with PROSAC–progressive sample consensus," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, vol. 1, Jun. 2005, pp. 220–226.

[7] A. de la Escalera, J. M. Armingol, and M. Mata, "Traffic sign recognition and analysis for intelligent vehicles," *Image Vis. Comput.*, vol. 21, no. 3, pp. 247–258, 2003.

[8] M. A. Fischler and R. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, no. 6, pp. 381–395, 1981.

[9] U. Handmann, T. Kalinke, C. Tzomakas, M. Werner, and W. V. Seelen, "An image processing system for driver assistance," *Image Vis. Comput.*, vol. 18, no. 5, pp. 367–376, 2000.

[10] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge, U.K.: Cambridge Univ. Press, 2003.

[11] R. I. Hartley, "In defense of the eight-point algorithm," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 6, pp. 580–593, Jun. 1997.

[12] R. I. Hartley and P. Sturm, "Triangulation," *Comput. Vis. Image Understand*, vol. 68, no. 2, pp. 146–157, 1997.

[13] V. J. Hodge and J. Austin, "A survey of outlier detection methodologies," *Artif. Intell. Rev.*, vol. 22, no. 2, pp. 85–126, 2004.

[14] C. Leys, C. Ley, O. Klein, P. Bernard, and L. Licata, "Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median," *J. Exp. Social Psychol.*, vol. 49, no. 4, pp. 764–766, 2013.

[15] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, 2004.
[16] J. Ma, W. Qiu, J. Zhao, Y. Ma, A. L. Yuille, and Z. Tu, "Robust L2E estimation of transformation for non-rigid registration," *IEEE Trans. Signal Process.*, vol. 63, no. 5, pp. 1115–1129, Mar. 2015.
[17] J. Ma, J. Zhao, J. Tian, X. Bai, and Z. Tu, "Regularized vector field learning with sparse approximation for mismatch removal," *Pattern Recognit.*, vol. 46, no. 12, pp. 3519–3532, 2013.
[18] J. Ma, J. Zhao, J. Tian, A. L. Yuille, and Z. Tu, "Robust point matching via vector field consensus," *IEEE Trans. Image Process.*, vol. 23, no. 4, pp. 1706–1721, Apr. 2014.
[19] J. Ma, H. Zhou, J. Zhao, Y. Gao, J. Jiang, and J. Tian, "Robust feature matching for remote sensing image registration via locally linear transforming," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 12, pp. 6469–6481, Dec. 2015.
[20] R. A. Maronna, D. R. Martin, and V. J. Yohai, *Robust Statistics*. Chichester, U.K.: Wiley, 2006.
[21] R. Raguram, J.-M. Frahm, and M. Pollefeys, "A comparative analysis of RANSAC techniques leading to adaptive real-time random sample consensus," in *Computer Vision—ECCV*. Berlin, Germany: Springer, 2008, pp. 500–513.
[22] P. J. Rousseeuw and C. Croux, "Alternatives to the median absolute deviation," *J. Amer. Statist. Assoc.*, vol. 88, no. 424, pp. 1273–1283, Dec. 1993.
[23] P. J. Rousseeuw and A. M. Leroy, *Robust Regression and Outlier Detection*, vol. 589. Hoboken, NJ, USA: Wiley, 2005.
[24] M. Salehi, X. Zhang, J. C. Bezdek, and C. Leckie, "Smart sampling: A novel unsupervised boosting approach for outlier detection," in *Proc. Australas. Joint Conf. Artif. Intell.* Berlin, Germany: Springer, 2016, pp. 469–481.
[25] F. Shen, W. Yang, H. Li, H. Zhang, and H. T. Shen, "Robust regression based face recognition with fast outlier removal," *Multimedia Tools Appl.*, vol. 75, no. 20, pp. 12535–12546, 2016.
[26] R. G. Staudte and S. J. Sheather, *Robust Estimation and Testing*, vol. 918. Hoboken, NJ, USA: Wiley, 2011.
[27] C. V. Stewart, "Robust parameter estimation in computer vision," *SIAM Rev.*, vol. 41, no. 3, pp. 513–537, 1999.
[28] B. Tordoff and D. W. Murray, "Guided sampling and consensus for motion estimation," in *Computer Vision—ECCV*. Berlin, Germany: Springer, 2002, pp. 82–96.
[29] P. H. S. Torr, "Bayesian model estimation and selection for epipolar geometry and generic manifold fitting," *Int. J. Comput. Vis.*, vol. 50, no. 1, pp. 35–61, 2002.
[30] P. H. S. Torr and D. W. Murray, "The development and comparison of robust methods for estimating the fundamental matrix," *Int. J. Comput. Vis.*, vol. 24, no. 3, pp. 271–300, 1997.
[31] P. H. S. Torr and A. Zisserman, "MLESAC: A new robust estimator with application to estimating image geometry," *Comput. Vis. Image Understand.*, vol. 78, no. 1, pp. 138–156, 2000.
[32] W. van der Mark and D. M. Gavrila, "Real-time dense stereo for intelligent vehicles," *IEEE Trans. Intell. Transp. Syst.*, vol. 7, no. 1, pp. 38–50, Mar. 2006.
[33] G. Wang and Q. M. J. Wu, "Perspective 3-D Euclidean reconstruction with varying camera parameters," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 19, no. 12, pp. 1793–1803, Dec. 2009.
[34] G. Wang, J. S. Zelek, and Q. M. J. Wu, "Structure and motion recovery based on spatial-and-temporal-weighted factorization," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 11, pp. 1590–1603, Nov. 2012.
[35] G. Wang, "Robust structure and motion recovery based on augmented factorization," *IEEE Access*, vol. 5, pp. 18999–19011, 2017.
[36] G. Wang and Q. M. J. Wu, *Guide to Three Dimensional Structure and Motion Factorization*. Berlin, Germany: Springer, 2011.
[37] G. Wang, J. S. Zelek, Q. M. J. Wu, and R. Bajcsy, "Robust rank-4 affine factorization for structure from motion," in *Proc. IEEE Workshop Appl. Comput. Vis. (WACV)*, Jan. 2013, pp. 180–185.
[38] X. Wang, D. Shen, M. Bai, T. Nie, Y. Kou, and G. Yu, "Cluster-based outlier detection using unsupervised extreme learning machines," in *Proceedings of ELM*, vol. 1. Berlin, Germany: Springer, 2016, pp. 135–146.
[39] C.-B. Xiao, D.-Z. Feng, and M.-D. Yuan, "An efficient fundamental matrix estimation method for wide baseline images," *Pattern Anal. Appl.*, vol. 21, no. 1, pp. 35–44, 2018.
[40] F. Xu, X. Liu, and K. Fujimura, "Pedestrian detection and tracking with night vision," *IEEE Trans. Intell. Transp. Syst.*, vol. 6, no. 1, pp. 63–71, Mar. 2005.
[41] M. Zhang, G. Wang, H. Chao, and F. Wu, "Fast and robust algorithm for fundamental matrix estimation," in *Proc. Int. Conf. Image Anal. Recognit.* Berlin, Germany: Springer, 2015, pp. 316–322.
[42] Z. Zhang, "Determining the epipolar geometry and its uncertainty: A review," *Int. J. Comput. Vis.*, vol. 27, no. 2, pp. 161–195, 1998.

SUSHIL PRATAP BHARATI received the bachelor's degree from the Motilal Nehru National Institute of Technology, Allahabad, India. He is currently pursuing the master's degree with the University of Kansas. His research interests include real-time object detection and tracking, two view geometry and 3-D reconstruction, pattern recognition, autonomous robotics and broad applications of computer vision, and machine learning.

FENG CEN received the Ph.D. in computer application technology from Shanghai Jiao Tong University, Shanghai, China, in 2003.

He is currently an Associate Professor with the Department of Control Science and Engineering, Tongji University, Shanghai, China. His research interests include computer vision, image processing, and pattern recognition.

AJAY SHARDA received the Ph.D. degree in the field of precision Ag/machinery systems from Auburn University, AL, USA, in 2011. He was a Post-Doctoral Fellow with the Center for Precision and Automated Agricultural Systems, Washington State University, WA, USA, from 2012 to 2013. He was a tenured Assistant Professor of farm power and machinery with the College of Agricultural Engineering, Punjab Agricultural University, from 2003 to 2006. He joined the Department of Biological and Agricultural Engineering, Kansas State University, as an Assistant Professor of precision Ag and machinery system in 2013. He has published over 50 peer reviewed and conference paper. His research has focused on development, analysis and experimental validation of control systems for agricultural machinery systems, such as precision planting and liquid application technologies, and thermal infrared imaging applications with ground and aerial unmanned systems. He also holds several leadership positions in multiple professional organizations, including the Board Member of KARTA, the Vice-Chair of the MS-60 Unmanned Aerial Systems Committee, and the Secretary of the MS 54 Precision Agriculture Committee.

GUANGHUI WANG (M'10–SM'17) received the Ph.D. degree in computer vision from the University of Waterloo, Canada, in 2014. He is currently an Assistant Professor with the University of Kansas, USA. He is also with the Institute of Automation, Chinese Academy of Sciences, China, as an Adjunct Professor.

From 2003 to 2005, he was a Research Fellow and a Visiting Scholar with the Department of Electronic Engineering, The Chinese University of Hong Kong. From 2006 to 2010, he was a Research Fellow with the Department of Electrical and Computer Engineering, University of Windsor, Canada. He has authored one book, *Guide to Three Dimensional Structure and Motion Factorization* (Springer-Verlag). He has published over 100 papers in peer-reviewed journals and conferences. His research interests include computer vision, structure from motion, object detection and tracking, artificial intelligence, and robot localization and navigation. He has served as an associate editor and on the editorial board of two journals, an area chair or a TPC member of over 20 conferences, and a reviewer of over 20 journals.