A new inlier identification scheme for robust estimation problems

***Abstract*— Common goal of many computer vision and robotics algorithms is to extract geometric information from the sensory data. Due to the presence of the sensor noise and errors in matching or segmentation, the available data are often corrupted with outliers. In such instances, the problem of estimation of parametric models needs to be tackled by robust estimation methods. In the presence of large fraction of outliers sampling based methods are often employed to tackle the task. When the fraction of the outliers is significant and the parametric model is complex, the traditionally used RANSAC algorithm requires large number of samples, prior knowledge of the outlier ratio and additional, difficult to obtain, inlier threshold for hypothesis evaluation.**

To tackle these problems we propose a novel and efficient sampling based method for robust estimation of model parameters from redundant data. The method is based on the observation that for each data point, the properties of the distribution of the residuals with respect to the generated hypotheses reveal whether the point is an outlier or inlier. The problem of inlier/outlier identification can then be formulated as a classification problem. The proposed method is demonstrated on motion estimation problems from with large percentage of outliers (70%) on both synthetic and real data and estimation of planar models from range data. The method is shown to be of an order of magnitude more efficient than currently existing methods and does not require prior knowledge of an outlier ratio and inlier threshold.

I. Introduction

Many computer vision and robotics algorithms strive to extract geometric information from the sensory data. The sensory data range from images, laser range data, or ultrasound and the geometric information is typically represented by a parametric model. Different models include planar surfaces to be estimated and segmented from the range data, motion/pose estimation problems from image correspondences or range scans. In most scenarios the data, in addition to sensor noise, are corrupted with significant fraction of outliers, due to either measurements errors, mismatches in correspondences or errors in segmentation. This rules out the applications of traditional least squares methods for estimation. The need for robust estimation methods has been widely acknowledged in both computer vision and robotics communities.

Many efforts have been made to obtain provably robust estimators. However their breakdown points [1] are usually low and they are very costly to implement in practice. Estimators such as LMedS and LTS [9] can tolerate only 50% of outliers. Although it is desirable to design estimators with a solid theoretical footing and provable breakdown points, they often have small bearing on practical problems, which can be tackled reliably. Many of the practical problems have been successfully approached either by sampling based methods (e.g. RANSAC) or Hough Transform, which can empirically tolerate high fractions of outliers.

Our work is motivated by the class of sampling based methods, similar in the spirit to RANSAC [2] algorithm introduced by Fishler and Bolles [1] , which is widely adopted for various robust estimation problems in computer vision. When the fraction of the outliers is significant and the parametric model is complex, the traditional RANSAC algorithm requires large number of samples and additional, difficult to obtain, inlier threshold for hypothesis evaluation. In the basic algorithm, individual hypotheses generated by the sampling process are evaluated with respect to all data points and ranked based on the number of their inliers, searching for the best hypothesis. The number of needed samples is related to the fraction of outliers which is often not known a-priori. Although RANSAC can handle more then 50% of outliers, as the fraction of outliers increases it becomes prohibitively expensive.

The main contribution of this paper is a novel inlier identification scheme, where we propose to classify the data points directly based on the generated hypotheses. The proposed approach is very efficient, especially for data sets contaminated with large fractions of outliers and eliminates the need of predefined inlier scale (threshold) and prior knowledge of the outlier ratio which determines the number of needed samples .

In our work we are motivated and focus on the problem of the estimation of camera motion from correspondences between two widely separated views. This problem is of great relevance for vision based localization problems in large scale environments. In particular in the context of relative positioning and location recognition tasks, where the camera pose with respect to a known landmark or reference view has to be computed. As Figure 9 demonstrates, in large scale urban environments, the matching stage and search for correspondences, usually yields large number of incorrect correspondences. The need for robust estimation methods has been previously explored in this context by several authors [10], [8], [5], [7].

The rest of the paper is organized as follows. In Section 2 we briefly review the basic RANSAC algorithm and discuss its drawbacks. Related work and partial improvements over traditional RANSAC are discussed in Section 3. The proposed hypothesis evaluation and inlier/outlier identification scheme

[1]Breakdown point of an estimator corresponds to a smallest percentage of outliers, which can cause arbitrarily large values of the estimator.

[2]RANdom SAmple Consensus.

is described in Section 4 and demonstrated on synthetic data. In Section 5 we present experiments on real data and Section 6 concludes the paper.

II. RANSAC ALGORITHM

The essence of the RANSAC algorithm is the generation of multiple hypotheses by means of sampling of the data. Given the minimal number of data points p needed to estimate the model and the fraction of the outliers ϵ, we can compute the probability ρ that given m samples, at least one of the samples is outlier free:

$$\rho = 1 - (1 - (1 - \epsilon)^p)^m. \quad (1)$$

In order to achieve a desired probability (confidence) ρ of an outlier free hypothesis and provided that the outliers fraction ϵ is known, we can compute from the above equation the required number of samples:

$$M = \left\lceil \frac{ln(1-\rho)}{ln(1-(1-\epsilon)^p)} \right\rceil. \quad (2)$$

Given the determined number of samples M (calculated based on Equation 2), hypothesis model parameters are estimated for each sample, followed by finding the support (typically number of inliers) for each hypothesis. Alternatively, a stopping criterion can be used to terminate the sampling if sufficient percentage of inliers has been encountered. It has been shown in [5] that the stopping times for the two strategies mentioned above differ only by a multiplicative factor. In the second stage the hypothesis with the largest support is chosen, and all its inliers are used to refine the model parameters. More detailed description of the RANSAC algorithm can be found in many papers such as [13].

The larger the sample size p, it is likely that the sample is outlier free and more samples are needed to achieve a target confidence. For illustration we show the number of samples needed to estimate fundamental matrix model for displacement between two views. The fundamental matrix has 9 elements, but only 7 degrees of freedom. When data set contains 50% of outliers, to estimate the fundamental matrix using linear 8-point algorithm, 766 samples are needed to assure 95% confidence that one outlier free sample is obtained. The number of required samples goes to 1177 for 99% confidence. As pointed out by [11], the theoretical number of samples is wildly optimistic. In practice, the number of samples required to reach a good hypothesis is around an order of magnitude more. The experiments in [5] also validated this rule. The actual number of samples needed for 99% confidence is on the order of 5000 (our simulations confirm this), which means around 5000 hypotheses need to be evaluated. As shown in Table I, when $\epsilon = 0.7$, the number of required samples is 45658. Consequently, the number of hypotheses to be evaluated will be on the order of 10^5. For each hypothesis, standard RANSAC algorithm computes the residual for every data point. Hence the computation increases linearly with the number of data points. Most of the related work tries to alleviate the efficiency problems related to large number of samples, expensive hypothesis evaluation stage and inlier threshold selection in various ways.

Outlier percentage ϵ	30%	40%	50%	60%	70%
7-point algorithm	35	106	382	1827	13696
8-point algorithm	51	177	766	4570	45658

TABLE I

THE THEORETICAL NUMBER OF SAMPLES REQUIRED TO ENSURE 95% CONFIDENCE THAT AT LEAST ONE OUTLIER FREE SAMPLE IS OBTAINED.

III. RELATED WORK

Chum and Matas [5] suggested to improve the efficiency of the standard RANSAC by a pre-evaluation, called $T_{d,d}$ test. It exploits the fact that for erroneous model, only a small number of data points needs to be evaluated. If d randomly selected points pass the $T_{d,d}$, test the hypothesis is not considered further. This enables the authors to increase the efficiency of the hypothesis evaluation stage, but the number of samples remains still large. In [6], the authors proposed to select sample sets of adjacent points based on the assumption that inliers will tend to be closer to one another than outliers and therefore increasing the probability of an outlier free hypothesis. Guided sampling by quality of matches was proposed by [11] increased the chance of sampling 'good' correspondences more often and hence generate good hypotheses. Torr and Zisserman [12] have noticed that simple evaluation of the hypotheses by their inlier count is faulty, since it treats all the inliers equally (error terms for the inliers are constant). Consequently, if the threshold T on the residual errors which is used for classifying the data points as inliers and outliers is not set appropriately, the final model estimate will be poor. They suggested using log likelihood of the solution as the support instead of number of inliers. Nister [7] has demonstrated a preemptive RANSAC scheme which can run in real time. The preemptive score was used to sequentially remove bad hypotheses, until only the best hypothesis is left or time budget is used out. The scheme was tested on synthetic data with 20% outliers. In real experiments the points were tracked between individual frame of the video sequence and contained small fraction of outliers. Additional speed up was obtained by the use of the 5-point algorithm method assuming that the camera is calibrated in advance. The issue of threshold selection for inlier identification have been addressed recently by [14]. They proposed an automatic scale selection methods for estimation of the scale of inlier noise by analyzing the distribution of residuals of each hypothesis and hence avoiding the threshold selection stage. The inlier scale was estimated, using iterative mean shift algorithm for locating the modes in the residual distribution. Although the approach was capable of handling large percentage of outliers ($\approx 85\%$) on simple line fitting examples, the efficiency related to the required number of samples and additional overhead caused by iterative scale estimation have not been addressed.

IV. The proposed scheme

We are motivated by motion estimation problem from two widely separated views, given image correspondences. In this problem the model to be estimated is complex and the data often contain significant fraction of outliers. The presence of the outliers is particularly pervasive in large scale outdoor urban environments and it is due to the often significant viewpoint change, illumination changes and ambiguities due to the repetitive structures inherent to buildings. The sets of correspondences often contain more than 50% outliers. As the Table 1 indicates using the traditional RANSAC sampling techniques would be prohibitively time consuming, in addition to the issues of inlier threshold selection. Even though the automated threshold selection methods [14] can overcome some of the difficulties, they introduce an additional overhead and do not affect the number of samples favorably, nor is the number of samples known ahead of time.

Note that inlier identification is at the core of RANSAC algorithm. The final model parameters are then estimated based on the identified inliers. The basic premise of the sampling based algorithms, is the generation of many hypothesis which would guarantee with some confidence that an outlier free hypothesis is encountered in the set. As shown in Table 1, this depends on the complexity of the model and fraction of the outliers, which is not known ahead of time. The preemptive RANSAC [7] is the only exception which uses a fixed number of samples (500-800), assuming outlier percentage is around 20% and calibrated setting with 5-point algorithm. The essence of the preemptive RANSAC scheme is to still try to find the good hypotheses. Although this method has been show to work well with video sequence (and hence low outlier ratios), it has not been extended to data containing more outliers.

In the presented approach, we instead of evaluating the goodness of individual hypothesis generated by the sampling process, we evaluate the residuals of each data point with respect to all hypotheses. The proposed method is based on the observation, that for each data point the properties (higher order statistics) of the distribution of the residuals with respect to the generated hypotheses reveal whether the point is an outlier or inlier. The problem of inlier/outlier identification can then be formulated as a classification problem. The presented approach relies on the continuity of the hypothesis space populated by hypotheses generated by sampling and does not require per se a presence of an outlier free hypothesis. Hence the large number of samples is not necessary. The approach in addition to its efficiency does not require prior knowledge of the outliers percentage and doesn't need any threshold for identification of inlier's support of the hypothesis. We demonstrate the performance of the proposed method on the problem of motion estimation, with varying outlier percentages (up to 70%) and show that we can correctly identify the inliers over varying fractions of outliers with fixed number of samples. In the next section we will describe the approach and justify it on a simple example. Extensive simulations and experiments on real images are presented in Section 4.

A. Inlier identification procedure

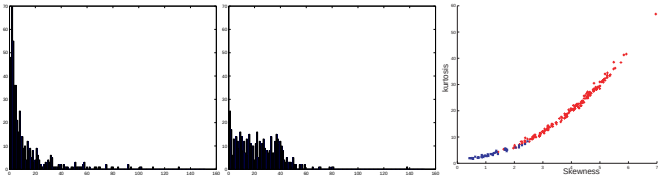
We will describe the proposed method on an example of estimation of the epipolar geometry between two views. Given a set of correspondences $\{\mathbf{x}_i, \mathbf{x}'_i\}_{i=1}^{C}$ between two views of the same scene, our goal is to estimate the fundamental matrix F. Similarly as in the standard RANSAC scheme we first use sampling to generate a set of hypotheses, (*i.e.* fundamental matrices). This is achieved by sampling the set of correspondences by selecting 8-point samples and estimating F using 8-point algorithm with normalization. At this stage our method dramatically departs from the previously proposed approaches. Instead of evaluating/scoring each hypothesis, we look at the data points directly. For each data point (e.g. correspondence) we study the distribution of the errors with respect to all hypotheses. For a hypothesis F_j instead of considering residual error $(r_j^i)^2 = (\mathbf{x}_i^T F_j \mathbf{x}'_i)^2$ we use the so called Sampson distance which approximates the geometric distance [3] of the point to the epipolar line and is defined as:

$$(r_j^i)^2 = \frac{(\mathbf{x}_i^T F_j \mathbf{x}'_i)^2}{(F_j\mathbf{x}_i)_1^2 + (F_j\mathbf{x}_i)_2^2 + (F_j^T\mathbf{x}'_i)_1^2 + (F_j^t\mathbf{x}'_i)_2^2} \qquad (3)$$

where $(F\mathbf{x})_k^2$ represents the square of the *k*-th entry of the vector $F\mathbf{x}$. Figure 1(a) and Figure 1(b) shows typical error distributions with respect to all generated hypotheses for a data containing 20% outliers. The data was generated using a total of 200 3D points projected into two views related by general motion. Note that the residual histograms of the inliers and outliers are very different. The inliers typically have strong peaks close to 0, while the outliers don't. We will use this observation for classification of the points to inliers and outliers based on n^{th} order statistics of their residual distribution. The outliers histogram of residuals, can also have high count in the first bin, because some hypotheses are generated using the samples which contain the outlier itself. For this reason the 1^{st} bin was set to 0 prior to computation of the statistics. The strong peak of the inlier's error distribution comes from two sources: the inlier can be included in several samples and it can be expected that several good hypotheses yielding a low residual error are included in the hypotheses set. In this example the probability that an 8-point sample is outlier free is $0.8^8 \approx 0.168$, the expect number of outlier free samples is approximately $0.168 \times 500 = 84$. [3] The number of samples used to generate the hypotheses is set to be $N = 500$. Based on our experiments, 500 samples approximate sufficiently the residual histogram, which is the foundation for inlier identification. Considering the size of the image plane is 400×600, the error histogram has 150 bins, representing the Sampson error ranging from 0 to 149 (large enough to capture the detail of the error distribution). We disregard errors greater than 149.

1) Features for characterizing the distributions: In order to characterize the qualitative differences between the distribu-

[3] The number of outlier free samples obeys a binomial distribution with N trials and the probability of success is the probability that a sample is outlier free.



(a) (b) (c)

Fig. 1. Error distribution for a true inlier (a) and a true outlier (b), $\epsilon = 0.2$. (c): plot of skewness vs. kurtosis computed for all residual distributions of the 200 data points (red '+' represents inliers, while blue 'x' represents outliers.)

tions of inliers and outliers depicted in Figure 1, several order statistics can be used. Most commonly used are the lower order statistics such as mean, standard deviation, skewness and kurtosis. Our experiments show that the skewness and kurtosis are very discriminative for the two kinds of residual histograms. Skewness γ measures the asymmetry of the data around the sample mean μ

$$\gamma = \frac{E(x-\mu)^3}{\sigma^3}. \tag{4}$$

Skewness of the normal distribution (or any perfectly symmetric distribution) is zero. If the value of skewness is positive, the data are spread out more to the right of the mean than to the left. Kurtosis β is the degree of peakedness of a distribution, which in our case measures how outlier prone a distribution is. Kurtosis is defined as:

$$\beta = \frac{E(x-\mu)^4}{\sigma^4}. \tag{5}$$

For the two histograms shown in Figure 1, the kurtosis and skewness for the inlier are 24.4 and 4.6, while for the outlier they are much smaller: 7.6 and 1.7, respectively. These characteristics capture the fact that inlier's histogram of residuals has much stronger peak than that of an outlier and can be used as feature for further classification.

We can plot the values of skewness and kurtosis for each data point in 2D, as Figure 1(c) shows. Note that the kurtosis and skewness are correlated, thus it's not necessary to use the two statistics together. In our case only kurtosis is used for identifying the inliers, making the classification more efficient. From the plot, we can see that the inliers and outliers have different values of skewness and kurtosis. Hence they can be easily separated, either by k-means clustering algorithm or we can simply rank the points in the order of decreasing kurtosis and consider the top k to be inliers. Notice that the true inliers have kurtosis with much larger variance than that of true outliers. Consequently, some true inliers will be misclassified as outliers after the grouping. This however will not cause a problem for the model estimation, because enough true inliers are identified. Also, a small number of true outliers might be included in the identified inliers set. The standard RANSAC can be applied for this inliers set. The computational demands are very low, since the outlier percentage small in this case[4] with no more than 10% outliers as our experiments show.

The inlier identification procedure for the case of fundamental matrix estimation is summarized below.

Algorithm 1 Inliers identifications procedure

1) Randomly select N 8-point samples and generate N fundamental matrix hypotheses $\{F_j\}, j = 1, 2 \ldots, N$.
2) For each correspondence (data point), compute its Sampson error [3] r_i^j with regard to each hypothesis.
3) For each correspondence, estimate its residual distribution by constructing histogram of N residuals associated with it. The histogram is used to evaluate whether the correspondence is an inlier.
4) For the C histograms, of residuals compute the value of kurtosis β_k to characterize each of them. In this stage each correspondence is represented by a point in the 1D kurtosis space.
5) Use k-means clustering algorithm to cluster the data into two groups, which are identified inliers and outliers or simply rank the points by their kurtosis value.

Note that the proposed scheme doesn't need a predefined threshold for inliers. The RANSAC schemes require a threshold T to determine whether a data is inlier. As mentioned in [12], the choice of the threshold is a sensitive parameter and can affect the performance dramatically. Without the need for the predefined T makes the proposed scheme very flexible to handle different data, which shows clear advantage over standard RANSAC scheme.

B. Asymptotic running time analysis

Note the steps 3, 4, and 5 of Algorithm 1 require extra computation compared to standard RANSAC. Given N samples and C correspondences, constructing the histograms takes $O(N \times C)$ and computing the value of kurtosis takes $O(N \times C)$ multiplications; k-means clustering in one dimension is very efficient. Together, the computation time they require are less than the second hypothesis evaluation stage of standard RANSAC which requires $O(N \times C)$ matrix multiplication. In our experiments the number of samples N was set to be 500. We have also evaluated the sensitivity of our methods with respect to the number of samples and obtained repeatable performance for varying outlier ratio when the number of samples varied between (400 - 1000). Note that this is an improvement of an order of magnitude compared to the works [5]. Just for comparison, the standard RANSAC requires $O(M \times C)$ matrix multiplication to evaluate all the hypotheses. Without knowing outlier percentage a-priori, M has to be set conservatively, e.g. $M = 30000$ to handle 60% outliers [15]. Hence the presented approach is much more efficient than standard RANSAC, especially when the outlier percentage is high.

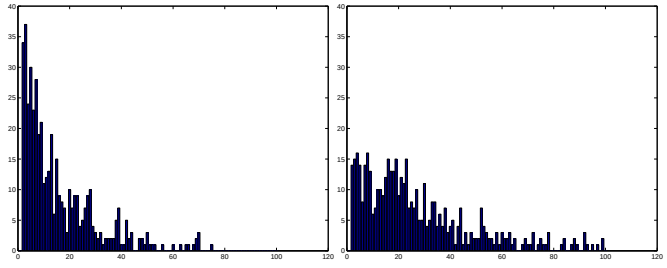
C. Justification based on synthetic data

We have shown in Section IV-A conceptual example that inliers can be identified directly. In the following section

[4]This step is not used in our inlier identification scheme, we emphasize in the inlier identification scheme.

we will demonstrate the feasibility of our approach based on a synthetic experiments. Set of 200 correspondences was generated by projecting 200 random cloud of 3D points, placed 1000 units of focal length in front of the camera, with the depth variation of 2000. The two views were related by general motion of translation around x-axis and rotation around y-axis of the camera frame. All the correspondences are corrupted by Gaussian noise (standard deviation was 1 pixel). 200 random correspondences were uniformly distributed in the image plane, yielding an outlier ratio of $\epsilon = 0.5$.

As Figure 2(a) and Figure 2(b) show, error distribution for inlier and outliers are rather different in this case ($\epsilon = 50\%$). This can be explained as follow: the residual distribution of each point is a mixture of two distributions. Residuals of wrong hypotheses are approximately random because wrong hypotheses are computed based on combination of inliers and outliers; while residuals of correct hypotheses are coherent and close together, because correct hypotheses are computed based on inliers only. Thus residual distributions of inliers are well peaked unimodal distribution, where the mode is close to 0. On the other hand, for the outliers the distribution is more spread out and has multiple modes.



(a) (b)

Fig. 2. Error distribution for a true inlier (a) and a true outlier (b), $\epsilon = 0.5$.

Figure 3 shows kurtosis of all the 400 data points (correspondences). For better visibility, the data are organized as 200 inliers followed by 200 outliers. Note that the inliers and outliers have quite different kurtosis. Inlier identification by k-means clustering is shown in Figure 4, depicting the actual projections of points on simulated image plane. The true inliers are represented by "x", the 138 identified inliers are circled and 2 false positives (outliers identified as inliers) are included and colored red. Other outliers are not shown for better visibility. With the true positive rate of $138/200 = 68\%$ and false positive rate $2/200 = 1\%$, the inlier identification performs fairly well with this heavily contaminated data set.

As the percentage of outliers increases, it can be expected that peak of inliers' error histogram becomes lower and eventually undistinguishable from the outlier. It's interesting to see to what extent our approach can tolerate outliers. We tried to study the separability of inliers from data containing different percentage of outliers. The number of inliers is fixed to be 200 obtained by projecting 200 random 3D points into two widely separated views, while the number of outliers varies for desired percentage. Figure 5 illustrates the changing of kurtosis. The

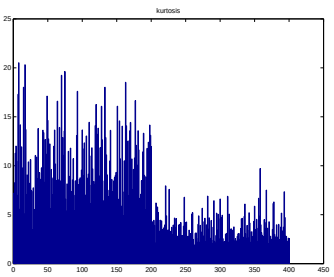


Fig. 3. The kurtosis of the 400 error distributions.

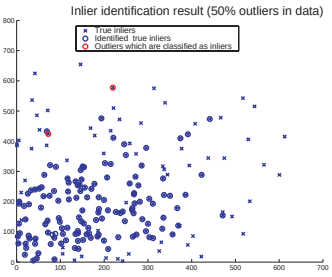


Fig. 4. The classification result based on kurtosis.

motion and 3D structure are set the same as in Figure 1. We can see that kurtosis value of outliers is always small, because they have no significant peaks. The kurtosis of inliers is much larger at first, meaning their error distributions do have strong peaks. Then it decreases as more outliers are added, because outliers would disperse the peaks. When fraction of outliers ϵ is less than 0.6, the mean of kurtosis computed based on inliers is above the 95% confidence interval of that of outliers. Therefore, the kurtosis of error histogram associated with inliers and outliers are statistically different, and inlier group obtained through k-means clustering is very unlikely to contain true outliers. When the outlier percentage increases further but no more than 0.7, the mean of inliers' kurtosis is close to the upper bound of that of outliers'. In this case, the inlier cluster obtained from k-means may contain some true outliers, but the percentage will be much lower than in the original data. As we mentioned before, an additional step of standard RANSAC on this inlier group can obtain model parameter with a small number of samples. When outlier percentage goes further to 0.75, inliers and outliers become indistinguishable. Figure 6 depicts the separation of inliers and outliers in the skewness/kurtosis space as the outlier ratio increases. The settings for the experiment were the same as in Figure 1. This indicates that the proposed approach can not tolerate more than 75% outliers. In theory the standard RANSAC does not have such limitation as long as enough samples are evaluated. When the outlier ratio ϵ is too high, the required number of samples is so large that it's impractical to work in practice. As mentioned in Section 2, when $\epsilon = 0.7$, the required number of samples is on the order of a half million, which is already too huge to work. So the proposed method has the same working range as standard RANSAC in practice, only that its much more efficient. Note that the limitation is obtained based on the estimation of fundamental matrix, which requires at least 7 data points. If the model to be estimated is simpler, for instance affine model which requires only 3 data points to estimate, more outliers can be tolerated. The reason is that the required number of samples would decrease dramatically in this case based on the relationship in Equation 2. We will demonstrate this using example of plane fitting.

1) Plane fitting in 3D space: Figure 7 shows 3D data with 500 points, 100 points lie in a plane and are corrupted by Gaussian noise ($\sigma = 1$) and 400 outliers are uniformly

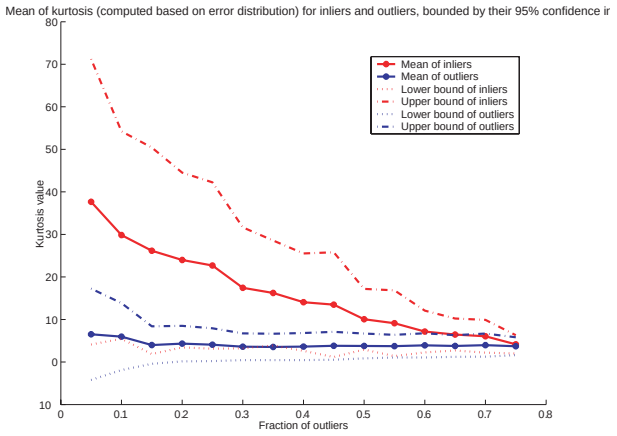
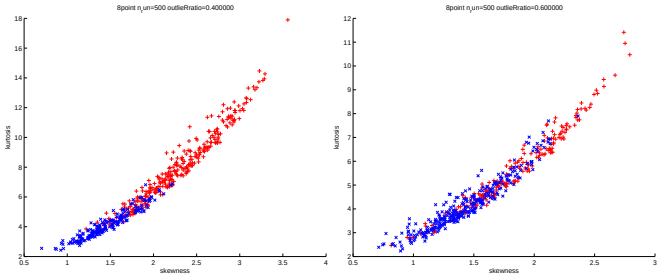


Fig. 5. This figure shows how the kurtosis value changes with different portion of outliers. Mean and 95% confidence interval of inliers' kurtosis are shown in red, mean and 95% confidence interval of outliers' kurtosis are shown in blue.



(a) $\epsilon = 0.4$. (b) $\epsilon = 0.6$.

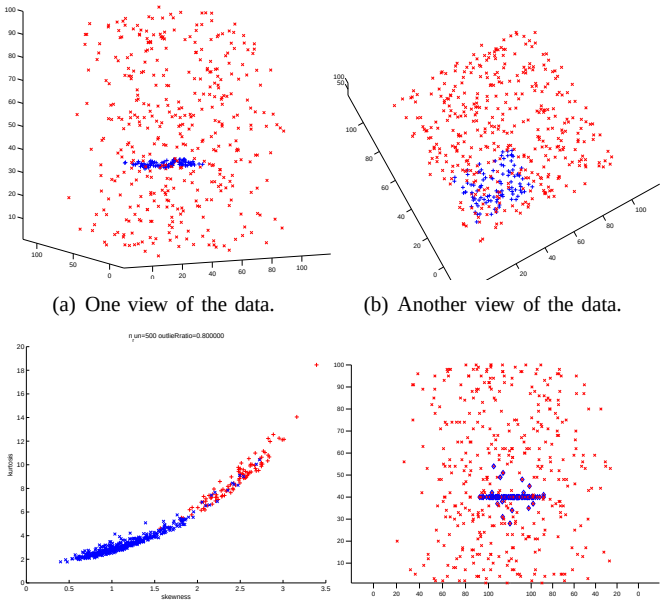
Fig. 6. Plots of kurtosis vs. skewness for different outlier percentage.

distributed in the space. 108 points are identified as inliers only 10% of which are false positives.

V. Experiments with real data

The proposed scheme was tested with real correspondences sets obtained from wide baseline matching. The putative correspondences were initiated based on matching of SIFT keypoints [4]. Two keypoints are set to be correspondence when the distance between the two SIFT keypoint descriptor is less than some threshold τ. We ran extensive experiments with correspondences sets containing different portion of outliers. We tested the methods in the domain of wide baseline matching between two views of urban scenes and/or buildings. In addition to large change of viewpoint between the views, these scene contain many repetitive structures, making the problem of finding correspondences by means of matching local feature descriptors highly ambiguous. Our focus is on the inlier identification capability of the proposed scheme. The identified inliers are not refined with additional RANSAC, so they might still contain few true outliers for severely contaminated data sets.

When percentage of outliers is low, our approach can identify inliers and outliers directly almost without mistake. The low percentage of outliers can also be handled by RANSAC without excessive computational overhead. We emphasize our

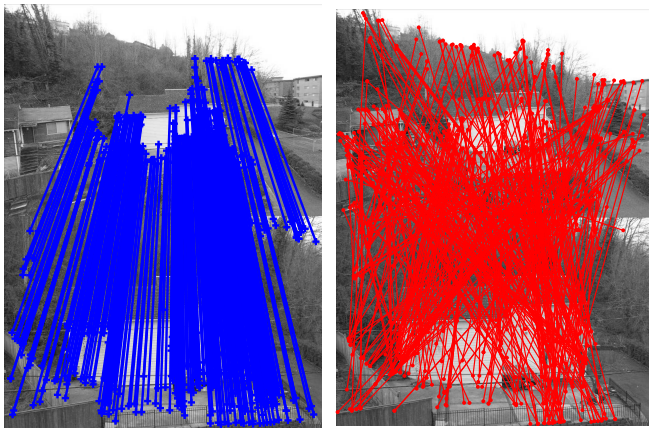


(a) One view of the data. (b) Another view of the data.

(c) Plot of skewness vs. kurtosis computed for all residual distributions. (d) Inlier/Outlier identification result.

Fig. 7. Fitting 3D plane with 80% outliers.

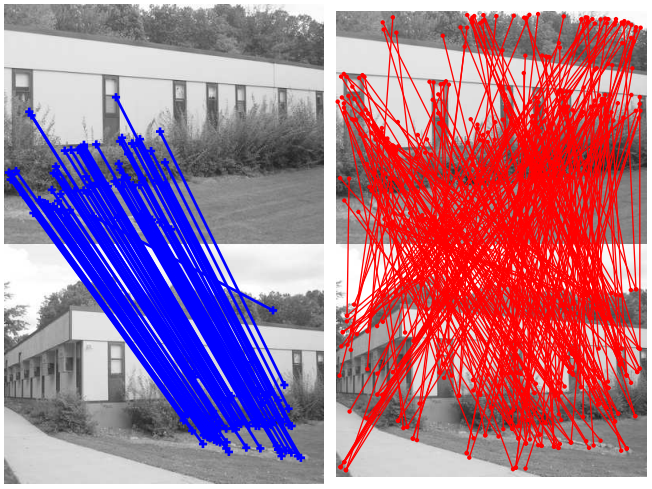
approach on correspondences sets with significant portion of outliers of more than 40%. Three examples are shown in Figures 9, 8 and 10. The identified inlier sets include most true inliers with very few outliers.



(a) identified inliers. (b) identified outliers.

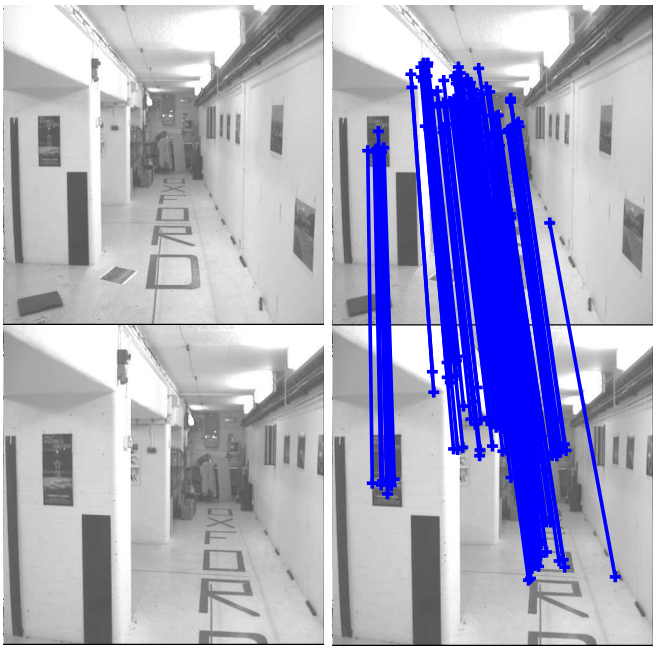
Fig. 8. 750 correspondences are initiated with around 50% outliers. 364 inliers are identified without false positive.

It is known that the distance threshold τ used for matching the SIFT keypoint descriptors affects the number of matches. Loose threshold results in many false correspondences. If the threshold is set too tight, hardly any matches could be found. Our work suggest a straightforward way to handle this: set a relatively loose threshold to obtain initial set of correspondences, and apply the proposed scheme to identify the true inliers.



(a) identified inliers. (b) identified outliers.

Fig. 9. 383 correspondences are initiated with approximately 60% outliers. 93 inliers are identified with only 1 false positive. Note the first left door in the left image corresponds to second left door in the right image.



(a) Pair of images. (b) identified inliers.

Fig. 10. Two frames of the widely used Corridor sequence (bt.001 and bt.006), obtained from http://www.robots.ox.ac.uk/ vgg/data/. Outlier percentage is over 50%. 134 inliers are identified with no false positive.

VI. Conclusion and future work

In this paper we proposed a new inlier identification scheme for robust estimation problems. We have demonstrated that it can efficiently handle data sets containing significant level of outliers. Inliers can be identified directly without looking for good hypothesis, thus avoiding the need for large number of samples, which is required for standard RANSAC algorithm. In addition to the efficiency of the proposed approach, we have also eliminated the need for sensitive threshold selection for outlier identification as well as need for prior knowledge about the percentage of outliers (which is needed when fixed number of samples is used in standard RANSAC). We would like to emphasize that the proposed method is especially suitable for data with large number of outliers as motivated and demonstrated in our application and often encountered in wide baseline matching. The proposed scheme is tested extensively with both synthetic and real data. We plan to refine the inlier identification step in future, by replacing the k-means clustering by its probabilistic version and hence obtaining the probability of being an inlier for each data point. We are also in the process of carrying out more extensive experiments with different distributions of outliers, in order to asses the generality of the presented method.

References

[1] M. A. Fischler and R. C. Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. In *Proceedings of the 4th European Conference on Computer Vision*, pages 683 – 695, 1996.

[2] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, ISBN: 0521540518, 2000.

[3] D. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 2004.

[4] Jiri Matas and Ondrej Chum. Randomized ransac with t(d,d) test. In *BMVC'02*.

[5] DR Myatt, PHS Torr, SJ Nasuto, and RJ Craddock. Napsac: High noise, high dimensional model parameterisation - its in the bag. In *BMVC*, pages 458–467, 2002.

[6] David Nister. Preemptive ransac for live structure and motion estimation. In *ICCV'03*, pages 199–206, 2003.

[7] Philip Pritchett and Andrew Zisserman. Wide baseline stereo matching. In *ICCV'98*.

[8] F. Schaffalitzky and A. Zisserman. Viewpoint invariant texture matching and wide baseline stereo. In *ICCV*, July 2001.

[9] Ben Tordoff and David W. Murray. Guided sampling and consensus for motion estimation. In *ECCV (1)*, pages 82–98, 2002.

[10] P. Torr and A. Zisserman. Mlesac: A new robust estimator with application to estimating image geometry. *CVIU*, 78:138–156, 2000.

[11] Philip H. S. Torr, Andrew Zisserman, and Stephen J. Maybank. Robust detection of degenerate configurations whilst estimating the fundamental matrix. pages 312–333, 1998.

[12] H. Wang and D. Suter. Robust adaptive-scale parametric model estimation for computer vision. *IEEE Trans. Pattern Anal. Mach. Intell.*, 26(11):1459–1474, 2004.

[13] H. Wang and D. Suter. Robust fitting by adaptive-scale residual consensus. In *ECCV*, pages 107–118, 2004.