Ensemble method for robust motion estimation

Wei Zhang and Jana Košecká
George Mason Univerity
4400 University Dr. Fairfax, VA 22030 USA
{wzhang,kosecka}@cs.gmu.edu

Abstract

The core of the traditional RANSAC algorithm and its more recent efficient counterparts is the hypothesis evaluation stage, with the focus on finding the best, outlier free hypothesis. Motivated by a non-parametric ensemble techniques, we demonstrate that it proves advantageous to use the entire set of hypotheses generated in the sampling stage. We show that by studying the residual distribution of each data point with respect to the entire set of hypotheses, the problem of inlier/outlier identification can be formulated as a classification problem. We present extensive simulations of the approach, which in the presence of a large percentage ($> 50\%$) of outliers, provides a repeatable and, an order of magnitude more efficient method compared to the currently existing techniques. Results on wide-baseline matching and fundamental matrix estimation are presented.

1. Introduction

The basic RANSAC algorithm introduced by Fischler and Bolles [3] is one of the most frequently used techniques for robust estimation problems in computer vision. The essence of the RANSAC algorithm is the generation of multiple hypotheses by means of sampling. Given a predetermined number of samples M, hypothesis model parameters are estimated for each sample, followed by finding the support for each hypothesis. To the best of our knowledge, the existing works following the standard RANSAC scheme focus on improving either the efficiency [5, 6] or the accuracy and the reliability of the hypothesis evaluation stage [9, 8, 10]. Additional works dealing with the issues of model selection and handling of the dominant plane are also valuable extensions useful, in the context of motion estimation problems [1]. We will briefly review the some of the representative works in these categories and refer the reader to the references therein.

The standard RANSAC uses the number of inliers as the main means of scoring the generated hypothesis. 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. To overcome the drawbacks of this scoring strategy, the MLESAC approach introduced by [9] proposed to evaluate the hypotheses by their maximum likelihood. For inliers the Gaussian model has been adopted and outliers were assumed to be distributed uniformly around 0, with the interval proportional to the size of the search window used in tracking. While this improved the scoring of individual hypotheses, the main idea behind the evaluation scheme was the same. The MLESAC idea was extended further to exploit the matching score to assess the probability that a feature is valid correspondence and to use the score to guide the matching process [8]. To deal with the problem of threshold selection the authors in [10] proposed an automatic scale selection method for estimation of the scale of inlier noise by analyzing the distribution of residuals of each hypothesis. The inlier scale was estimated using an iterative mean shift algorithm for locating the modes in the residual distribution. Although the approach was capable of handling a 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 the iterative scale estimation scheme has not been addressed.

Another class of improvements to the traditional RANSAC scheme was focused on the improvements of the efficiency of the hypothesis evaluation scheme. In [5] authors introduced a pre-evaluation $T_{d,d}$ test, which enabled them to evaluate only a fraction of the data points and hence gain additional savings. This scheme, however, did not affect favorably the number of required hypothesis samples and still relied on the presence of an outlier free hypothesis. The real-time constraints led to the development of the preemptive

RANSAC scheme proposed in [6], where the preemptive score was used to sequentially remove bad hypotheses until only the best one remained or the time budget ran out. The scheme was tested on synthetic data with 20% outliers and real video sequences. The fractions of outliers were relatively small due to the fact that the correspondences were obtained by tracking techniques.

Presented work is motivated by difficult wide-baseline matching/correspondence problems with a large number of outliers and the need for efficient and reliable robust estimation methods. Instead of relying on the existence of an outlier free hypothesis, we demonstrate that the inliers can be identified directly by characterizing the distribution of the residuals of each data point with respect to all hypotheses. The presented method, similarly to RANSAC, relies on the generation of a number of hypotheses by sampling. We exploit the entire ensemble of hypotheses and show that the distributions of residuals of inliers and outliers are qualitatively different and can be used for reliable classification of data points as inliers and outliers. Due to the fact that the presented method does not rely on the presence of an outlier free hypothesis, the number of required samples for high outlier ratios is of an order of magnitude less than the traditional RANSAC approach. In our earlier recent work [11] we introduced the main idea behind the proposed inlier estimation scheme, focusing on wide-baseline matching and fundamental matrix estimation and surface fitting from the laser range data. This paper extends the previous work in several important ways. First, we demonstrate the sensitivity of the method to different fractions of outliers, different separation of outliers and inliers given commonly encountered error distributions and different number of samples. We also point out the relationship of the procedure to commonly used non-parametric re-sampling methods used in statistics such as bootstrap, jackknife [2, 7]. At last, we demonstrate the applicability of the method to robust motion estimation problems from wide-baseline correspondences with high outlier ratios.

2. Motivation

In order to deal with the presence of outliers in regression problems, the standard RANSAC algorithm tackles this issue by means of sampling. Given a minimal number of points k, needed to estimate the model parameters, S subsets of samples of k points are first drawn from the data and used to estimate an initial set of S hypotheses. These hypotheses are subsequently evaluated in the search for the *best*, outlier free hypothesis. The points are then classified as inliers if their residual, with respect to the *best* hypothesis, is below some threshold. Standard analysis assumes, that given the minimal number of data points k needed to estimate the model and the fraction of the outliers ϵ, we can compute the probability ρ that given s samples at least one of the samples is outlier free as:

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

Table 1. The number of samples required to ensure 95% confidence that at least one outlier free sample is obtained.

$$\rho = 1 - (1 - (1 - \epsilon)^k)^s. \quad (1)$$

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:

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

Table 1 shows the number of samples needed for different outlier ratios so as to assure 95% confidence of an outlier free sample. When the data set contains 50% of outliers and 8-point algorithm is used, 766 samples are needed. The number of required samples goes to 1177 for 99% confidence. As pointed out by [8], the above number of samples is wildly optimistic. In practice, the number of samples required to reach a good hypothesis is of an order of magnitude more. The experiments in [5] also validate this rule. The actual number of samples needed for 99% confidence is in the order of 5000. 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. One reason for this is that Equation 1 is only an approximation of the true probability, which is valid when number of initial matches $n >> k$. The exact analysis shows that if we have n data points and m outliers the probability of good sample is the number of good subsets $\binom{n-m}{k}$, divided by the number of all possible subsets, $\binom{n}{k}$ which gives

$$p = \frac{(n-k)(n-k-1)\dots(n-k-m+1)}{n(n-1)\dots(n-m+1)}. \quad (3)$$

For example, for 200 data points and $\epsilon = 0.3$ with $k = 8$, we get $p = 0.054$. The number of good solutions g out of s possible samples is a random variable with binomial distribution

$$\binom{s}{g} p^g (1-p)^{s-g}.$$

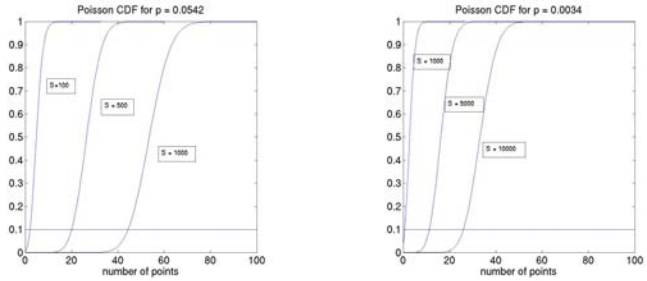


Figure 1. The cumulative distribution of Poisson approximation to binomial distribution for different samples sizes, given $N = 200$ matches. (left) The outlier ratio $\epsilon = 0.3$ yields a probability $p = 0.0542$ of a good hypothesis; (right) the outlier ratio $\epsilon = 0.5$, gives $p = 0.0032$.

If $p < 0.1$ and s is large, this can be approximated by Poisson distribution

$$P_\gamma(g) = \frac{\gamma^g e^{-\gamma}}{g!},$$

where $\gamma = Sp$ is the expected number of successes over S trials. Note that the estimates in Table 1 are quite optimistic. Figure 1 (right) shows that with $\epsilon = 0.5$ almost $S = 1000$ samples are needed to obtain the probability of an outlier free sample around 0.9. Similar relationship is shown for $\epsilon = 0.3$ and $S = 100, 500, 1000$ (left).

The earlier improvements of RANSAC for motion estimation, which considered parametric noise models [8, 9], typically considered scenarios where correspondences were attained by means of feature tracking in the video sequence. In such settings, the percentage of the outliers rarely exceeded 30%. Outliers were typically caused by occlusions or a drift accumulated during tracking and the range of errors was related to the size of the search window in tracking, which was around 5-10 pixels. Our work is motivated by recent advances in wide baseline matching and motion estimation from widely separated views and is most advantageous in this context. Here the number of potential correspondences is usually higher (on the order of 10^2-10^3), but in certain environments the percentage of outliers easily exceeds 50%. While the inlier errors are related to the localization of individual keypoints and are on the order of 1-2 pixels, the outlier errors are usually much larger, making the two error distributions well separated.

3. The proposed scheme

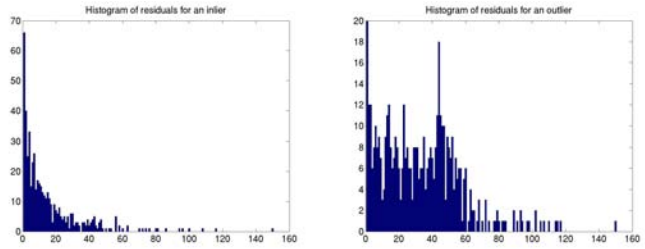
In the presented approach, instead of evaluating individual hypotheses generated in the sampling stage, we evaluate the residuals of each data point with respect to all hypotheses. We demonstrate that for inliers this distribution has a strong peak close to the origin, reflecting the fact that many inliers' residuals are small, while for outliers the distribution of residuals is more spread out. This property of residual distributions can be characterized by higher order statistics and enables us to formulate the problem of inlier/outlier identification as a classification problem. The presented approach relies on the properties of the hypothesis space populated by hypotheses generated by sampling and uses the entire ensemble of hypotheses to determine whether the point is an inlier or an outlier. Since the entire set of hypotheses is exploited, the method does not rely on a presence of an outlier free hypothesis. In the situations with a large fraction of outliers, the required number of samples is of an order of magnitude less then previously proposed methods. In addition to its efficiency, the presented approach does not require prior knowledge of the outliers percentage nor a threshold T for determining inliers' support of the hypothesis.

3.1. Inlier identification procedure

Given a set of correspondences $\{\mathbf{x}_i, \mathbf{x}'_i\}_{i=1}^C$ between two views, 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 an 8-point algorithm with normalization. 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 of the point to the epipolar line and is defined as:

$$(r_j^i)^2 = \frac{(\mathbf{x}_i^T F_j \mathbf{x}'_i)^2}{(F_j\mathbf{x}_i)_1^2 + (F_j\mathbf{x}_i)_2^2 + (F_j^T\mathbf{x}'_i)_1^2 + (F_j^t\mathbf{x}'_i)_2^2} \quad (4)$$

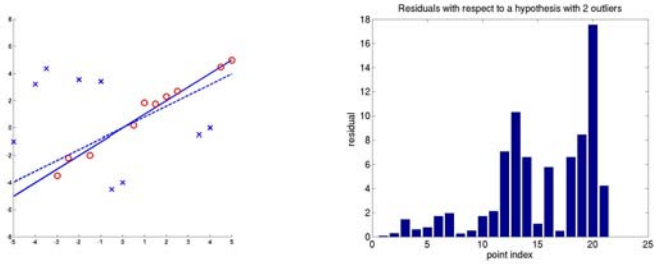
where $(F\mathbf{x})_k^2$ represents the square of the k-th entry of the vector $F\mathbf{x}$. Figure 2(a) and Figure 2(b) shows typical error distributions with respect to all generated hypotheses for a data containing 30% outliers. The data was generated using a total of 200 3D points in general position with depth variation of 1000 projected into two views related by general motion. The inliers were corrupted by a zero mean Gaussian noise and standard deviation of 2 pixels, while the inliers were assumed to be uniformly distributed in the interval $[-50, -20] \cup [20, 50]$ pixels. In this case, the error distributions of inliers and outliers are well separated. The residuals are computed with respect to 500 hypotheses generated in the sampling stage.



(a) (b)

Figure 2. Histogram of residuals for a true inlier (a) and a true outlier (b) with the outlier ratio, $\epsilon = 0.3$ and number of samples $S = 500$.

Note that the residual histograms of the inliers and outliers are very different. The inliers typically have a strong peak close to 0, while the outliers' distribution is more wide-spread. We will use this observation for classification of the points to inliers and outliers based on n^{th} order statistics of their residual distribution. Our approach exploits the fact that the residuals of the inliers are noticeably lower, even with respect to hypotheses which contain outliers. Hence, the hypotheses which contain some small number of outliers, contribute to the peak of the inliers' residual histogram. Figure 3 demonstrates this observation on a simple line fitting example. In this case the line fit was obtained using all the inliers and two outliers.



(a) (b)

Figure 3. (a) Line fitting example: the dashed line is the obtained estimate; (b) the residuals of all data points with respect to the line model with two outliers. Note that although the estimated line contains two outliers, the residuals of inliers (first 10 points) are notably smaller then the residuals of the outliers (last 10 points).

3.2. Distribution statistics

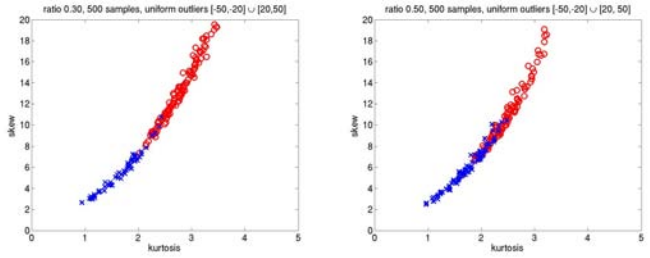
In order to characterize the qualitative differences between the distributions of inliers and outliers depicted in Figure 2, we compute the skewness and kurtosis of the two kinds of residual histograms. Skewness γ measures the asymmetry of the data around the sample mean μ

$$\gamma = \frac{E(x-\mu)^3}{\sigma^3}. \quad (5)$$

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 captures the amount of peakedness of a distribution and is defined as

$$\beta = \frac{E(x-\mu)^4}{\sigma^4}. \quad (6)$$

For the two histograms shown in Figure 2, 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 the inlier's histogram of residuals has a much stronger peak than that of an outlier and can be used for further classification. Figure 4 shows the plots of the values of skewness and kurtosis for each data point in 2D for outlier ratios 0.3 and 0.5. As we can see the kurtosis and skewness are correlated. In our experiments, only kurtosis is used for identifying the inliers. Inliers can be easily separated, either by a k-means clustering algorithm or simply by ranking the points in the order of decreasing kurtosis and considering the top k to be inliers. Notice that the true inliers have kurtosis with a much larger variance than the outliers. Consequently, some inliers can be missclassified 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 to this new inlier set. The computational demands of this second stage are very low since the outlier percentage is very small in this case with no more than 10% outliers as our experiments show.



(a) (b)

Figure 4. Kurtosis and skewness of all data points, 'o' denotes inliers and 'x' outliers with 500 samples. (a) The outlier ratio is 0.3 (b) the outlier ration is 0.5. Outlier errors were uniformly distributed between $[-50, -20] \cup [20, 50]$;

Notice that in Figure 2 the outliers' histogram of residuals can also have a 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. Considering that 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. In the above experiments the number of samples used to generate the hypotheses is set to be $N = 500$. We examine this choice more closely in the following section. Algorithm 1 summarizes the procedure for inlier identification for the case of fundamental matrix estimation.

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 r_i^j with respect 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 a k-means clustering algorithm to cluster the data into two groups, which are identified as inliers and outliers, or simply rank the points by their kurtosis value.

Note that the proposed scheme does not need a predefined threshold for inlier identification, nor does it require prior knowledge of the outlier ratio. In the next section we will demonstrate, in simulation, the sensitivity of the method with respect to different choices of parameters.

4. Performance and sensitivity

Number of samples. Since the method uses the entire ensamble of hypotheses, it is difficult to establish a bound on the number of samples related to a desired confidence as done in standard RANSAC. We study this relationship in simulation by measuring the separability of the inliers and outliers as a function of the number of samples. The Jeffries-Matusita distance measure of separability is used to assess how well two classes may be separated. Assuming that the two classes can be represented by normal distributions with $N(\mu_i, C_i)$ and $N(\mu_j, C_j)$ the Jeffries-Matusita distance is computed as

$$JM_{ij} = \sqrt{2(1 - e^{-\alpha})} \tag{7}$$

where

$$\alpha = \frac{1}{8}(\mu_i - \mu_j)^T \left(\frac{2}{C_i + C_j}\right)(\mu_i - \mu_j) + \frac{1}{2} ln \left(\frac{\frac{1}{2}|C_i + C_j|}{\sqrt{|C_i||C_j|}}\right)$$

The Jeffries-Matusita distance has an upper bound of $1.414(\sqrt{2})$, and a lower bound of 0. When the calculated distance has a value of the upper bound, the signatures can be said to be totally separable, while the signatures are inseparable when the calculated distance is 0. We experimented to see how the number of samples affects the performance characterized using JM distance. If outliers and inliers are totally separable, their JM distance would be 1.414. On the other hand, if they are tangled together and inseparable, the distance would be 0. For a given outlier ratio, we ran 100 trials of the proposed method and estimated the two clusters based on their kurtosis measure. Mean and standard deviation for both outliers and inliers were obtained in each run. Then we used the average of those means and standard deviations to compute JM distance for the particular number of samples. Figure 5 shows the simulation for data with different outlier ratios. As shown, the separability improves when the number of samples increases. However, after 500 samples, the separability wouldn't improve much as the number of samples increases.

In the case of lower outlier ratios smaller number of samples is sufficient. Figure 6 shows the separation of inliers and outliers using 200 samples. Lowering the number of samples further will affect the quality of the skewness and kurtosis estimates. Since we would like to have a method which works for a range of outlier ratios we choose the number of samples which has been shown in simulation (Figure 5) to provide a good separation of outliers and inliers for high outlier ratios.

Separation of error distributions. Since the method relies on the different properties of residuals of outliers and inliers, it is affected by the assumptions about the separation of the two error distributions. Figure 7 shows the simulation for the case when the error distribution of outliers is uniform on the interval $[-20, -3] \cup [3, 20]$, as well as uniform on the interval $[-50, 50]$. The inliers are distributed normally with $N(0, 2)$. Since the error ranges are less separated, some of the data points can be missclassified compared to

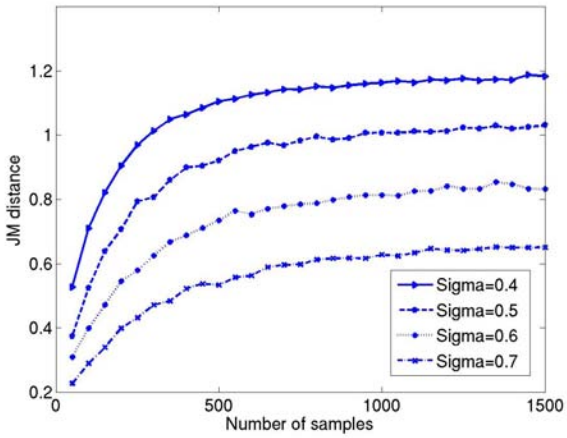
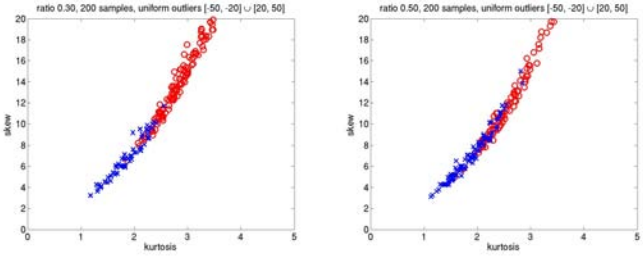
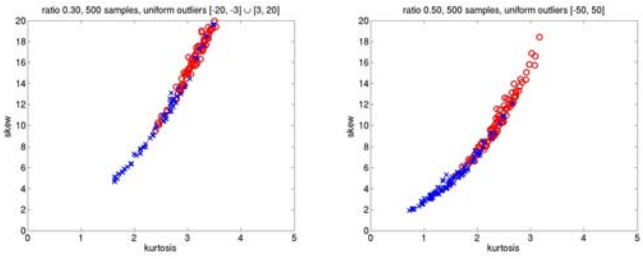


Figure 5. The JM distance changes with the number of samples. Note that it increases little after 500 samples.



(a) (b)

Figure 6. Kurtosis and skewness of all data points, 'o' denotes inliers and 'x' outliers computed based on 200 samples. (a) The outlier ratio is 0.3 and (b) the outlier ratio is 0.5. Outlier errors were uniformly distributed between $[-50, -20] \cup [20, 50]$.



(a) (b)

Figure 7. Kurtosis and skewness of all data points, 'o' denotes inliers and 'x' outliers. (a) The outlier ratio is 0.3, with outlier errors uniformly distributed in $[-20, -3] \cup [3, 20]$ and (b) the outlier ratio is 0.5 and the outlier errors uniformly distributed in $[-50, 50]$. 500 samples were used in both experiments.

examples in Figure 4. This however is consistent with the observation that the missclasified outliers with low errors are in fact consistent with the epipolar geometry. Hence, the proposed method can still effectively identify a majority of the outliers and can be used as a preprocessing step to the standard RANSAC, which can then be employed with a much lower number of samples. Changing the variance of outliers does not affect the separation capability.

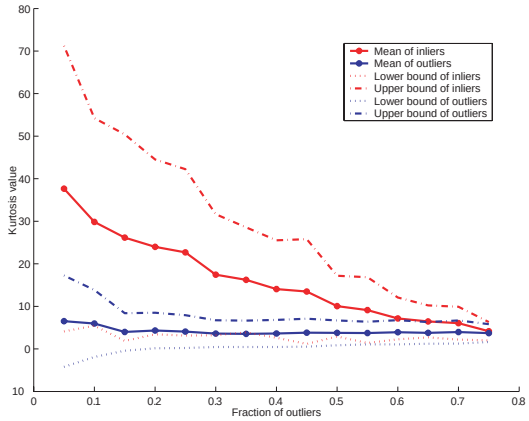
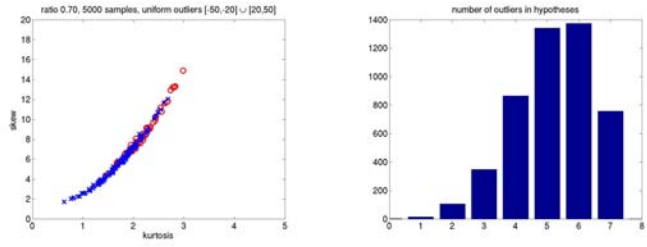


Figure 8. Mean and 95% confidence interval of inliers' kurtosis are shown in red, mean and 95% confidence interval of outliers' kurtosis are shown in blue.

Outlier ratio. Once the outlier percentage exceeds 65%, the proposed method is not as effective. This constitutes the limiting case of our method. Increasing the number of samples would not help, since with the increasing number of samples the number of hypotheses with a significant fraction of outliers would also increase. Figure 8 depicts the separation of inliers and outliers in the skewness/kurtosis space as the outlier ratio increases. he settings for the experiment were the same as in Figure 2. 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. It then decreases as more outliers are added because outliers would disperse the peaks. When the 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 an error histogram associated with inliers and outliers are statistically different, and an 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 a model parameters with

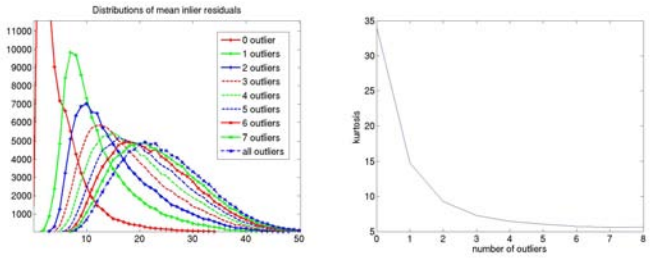
a small number of samples. When the outlier percentage goes further to 0.75, inliers and outliers become indistinguishable. Figure 9 shows the effect of increasing the number of samples for the outlier ratio 0.7. At this high outlier ratio, the probability of obtaining a sample with a small number of inliers is very low. The increase in the number of samples which contain one, two or three outliers (and contribute to the kurtotic shape of inliers' distribution) is almost negligable. In the majority of the samples more than half of the points are outliers. The fact that only the hypothesis



(a) (b)

Figure 9. Kurtosis and skewness of all data points, 'o' denotes inliers and 'x' outliers. The outlier ratio is 0.7 and the outlier errors were uniformly distributed between $[-50,-20] \cup [20,50]$. (a) 5000 samples; (b) histogram of hypotheses with different outlier numbers. Increasing the number of samples does not improve the separability of inliers and outliers.

with one, two and three outliers significantly contribute to the kurtotic behavior of inliers' residual distribution is demonstrated in Figure 10. The plot 10(a) shows different distributions of the mean inlier residuals, for different number of outliers present in each sample and (b) their corresponding kurtosis. The number of outliers in each sample was fixed for each distribution.



(a) (b)

Figure 10. (a) Distribution of mean residuals of inlier points as a function of number of outliers in the generated hypotheses; (b) kurtosis values of the same distributions. Note that for hypotheses with more then 3 outliers the curve flattens out.

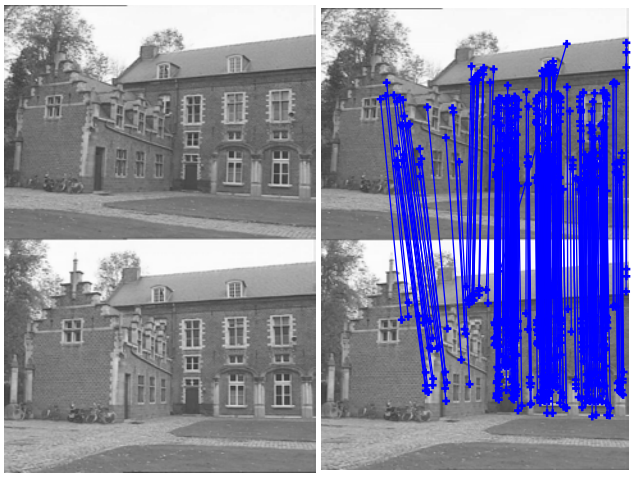
Asymptotic running time analysis. The steps 3, 4, and 5 of Algorithm 1 require extra computation compared to standard RANSAC. Given N samples and C correspondences, constructing the histograms takes $O(NC)$ and computing the value of kurtosis takes $O(NC)$ multiplications; the classification in 1-D kurtosis space is very efficient. Overall computation time is less than the second hypothesis evaluation stage of standard RANSAC which requires $O(NC)$ matrix multiplication. For the scenarios examined by simulations and in real experiments, the number of samples used was 500 with outlier ratios greater then 0.6. This is an improvement of an order of magnitude compared to the standard RANSAC method.

5. Experiments

The proposed scheme was tested with real correspondence sets obtained from wide baseline matching. The putative correspondences were obtained based on the matching of SIFT keypoints [4]. We tested the method in the domain of wide baseline matching between two views of urban scenes and/or buildings. In addition to a large change of viewpoint between the views, these scenes contain many repetitive structures, making the problem of finding correspondences by means of matching local feature descriptors highly ambiguous. Our focus here is on the inlier identification capability of the proposed scheme. In this stage, the identified inliers are not refined with additional RANSAC round, so they might still contain few true outliers for severely contaminated data sets. Two examples are shown in Figures 12 and 11. The identified inlier sets include most of the true inliers and very few outliers. Since we do not rely, per se, on the presence of an outlier free hypothesis, our method for inlier identification is not affected by the dominant plane problem. Even in cases when the linear algorithm for fundamental matrix estimation is poorly conditioned, due to the fact that the selected subset of points comes from the plane, all the inliers will still have small residuals with respect to the closest fundamental matrix estimate. In this case, however, once a subset of inliers is identified a proper model selection has to be carried out, but now with a substantially lower number of outliers. We also carried out some preliminary experiments, which indicate that the presented method is applicable in the context of general systems of linear equations contaminated by outliers.

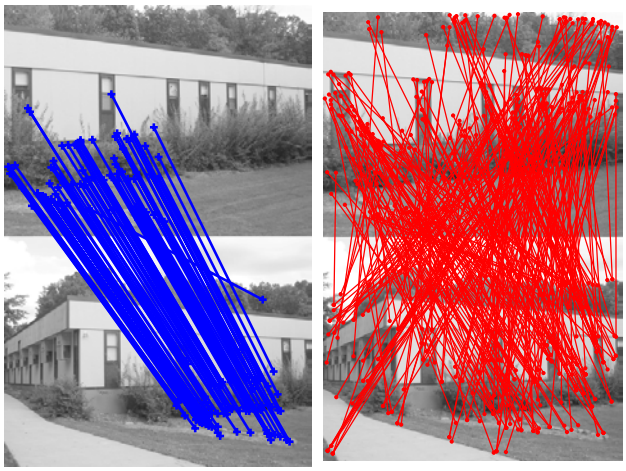
6. Conclusion and future work

When the percentage of outliers is low, our approach can identify inliers and outliers directly almost without



(a) (b)

Figure 11. Leuven image pair with identified inliers.



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

Figure 12. 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.

mistake. The low percentage of outliers can also be handled by RANSAC without excessive computational overhead. We emphasize that our approach is most advantageous on correspondences sets with significant portion of outliers of more than 40% as often encountered in wide baseline matching. In this paper, we proposed a new inlier identification scheme for robust motion estimation, which can efficiently handle data sets containing significant levels of outliers. Inliers can be identified directly without relying on the existence of an outilier free hypothesis, thus avoiding the need for a 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 the need for prior knowledge about the percentage of outliers. The proposed scheme has been tested extensively with both synthetic and real data. We are also in the process of carrying out more extensive experiments and analysis in order to asses the applicability of the presented method for other estimation problems.

References

[1] O. Chum, T. Werner, and J. Matas. Two view geometry estimation unaffected by dominant plane. *CVPR*, 2005.

[2] B. Efron and G. Gong. A leisurely look at the bootstrap, the jackknife and cross validation. *The American Statistitian*, 37(1):36–48, 1983.

[3] 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, 1999.

[4] D. Lowe. Distinctive image features from scale-invariant keypoints. *IJCV*, 37(1):36–48, 2004.

[5] J. Matas and O. Chum. Randomized ransac with T(d,d) test. In *BMVC*, pages 458–467, 2002.

[6] D. Nister. Preemptive RANSAC for live structure and motion estimation. In *ICCV*, pages 199–206, 2003.

[7] C. Tomasi and J. Zhang. A resampling methods for computer vision. In *Proceedings of Nith International Symposium on Robotics Research*, pages 89–96, 1999.

[8] B. Tordoff and D. Murray. Guided sampling and consensus for motion estimation. *ECCV*, 2002.

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

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

[11] W. Zhang and J. Košecká. A new inlier identification scheme for robust estimation problems. In *submitted to Rotics Science and Systems*, page (to appear), 2006.