Open access • Book Chapter • DOI:10.1007/978-3-642-19282-1_44

Efficient multi-structure robust fitting with incremental top-k lists comparison

Hoi Sim Wong, Tat-Jun Chin, Jin Yu, David Suter

Institutions: University of Adelaide

Published on: 08 Nov 2010 - Asian Conference on Computer Vision

Topics: Sampling (statistics), RANSAC, Model selection and Sorting

Efficient Multi-Structure Robust Fitting with Incremental Top-k Lists Comparison

Hoi Sim Wong, Tat-Jun Chin, Jin Yu and David Suter

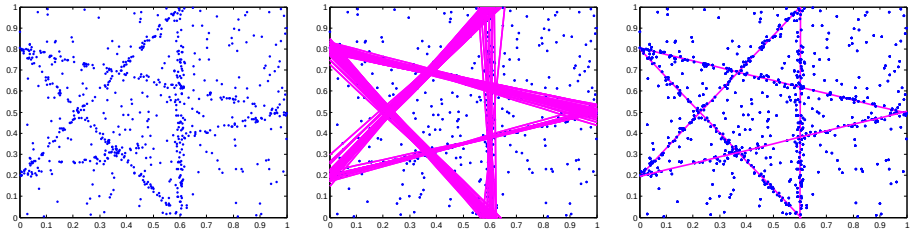
School of Computer Science,
The University of Adelaide, South Australia
{hoi.wong,tjchin,jin.yu,david.suter}@adelaide.edu.au

Abstract. Random hypothesis sampling lies at the core of many popular robust fitting techniques such as RANSAC. In this paper, we propose a novel hypothesis sampling scheme based on incremental computation of distances between partial rankings (top-k lists) derived from residual sorting information. Our method simultaneously (1) guides the sampling such that hypotheses corresponding to all true structures can be quickly retrieved and (2) filters the hypotheses such that only a small but very promising subset remain. This permits the usage of simple agglomerative clustering on the surviving hypotheses for accurate model selection. The outcome is a highly efficient multi-structure robust estimation technique. Experiments on synthetic and real data show the superior performance of our approach over previous methods.

1 Introduction

Robust model fitting techniques play an integral role in computer vision since the observations or measurements are frequently contaminated with outliers. Major applications include the estimation of various projective entities from multi-view data [1] which often contain false correspondences. At the core of many robust techniques is random hypothesis generation, i.e., iteratively generate many hypotheses of the geometric model from randomly sampled minimal subsets of the data. The hypotheses are then scored according to a robust criterion (e.g., RANSAC [2]) or clustered (e.g., Mean Shift [3]) to find the most promising model(s). Success rests upon retrieving an adequate number of *all-inlier* minimal subsets which may require a large enough number of sampling steps.

This paper addresses two major issues affecting the current paradigm of robust estimation. The first is that hypothesis generation tends to be time consuming for heavily contaminated data. Previous methods attempted to improve sampling efficiency by guiding the sampling such that the probability of selecting all-inlier minimal subsets is increased. These methods often depend on assumptions or domain knowledge of the data, e.g., inliers have higher keypoint matching scores [4, 5] or are correspondences that respect local geometry patterns [6]. Most methods, however, are not optimized for data with *multiple instances* (or *structures* [7]) of the geometric model. This is because they sample based on



(a) (b) (c)

Fig. 1. (a) Input data with 5 structures (lines) with 100 points per structure and 250 gross outliers. The inlier scale is 0.01. (b) 500 hypotheses are generated with the proposed *multi-structure* guided sampling scheme and *simultaneous* hypothesis filtering, producing 146 good hypotheses as shown in the figure. (c) Simple agglomerative clustering of the remaining 146 hypotheses gives the final fitting result.

estimated inlier probabilities alone while ignoring the fact that only inliers from the *same* structure should be included in the same minimal subset. Such methods may inefficiently generate a large number of samples before obtaining an all-inlier minimal subset for each genuine structure in the data.

The second crucial issue is the lack of a principled approach to fit the multiple structures in the data. Many previous works [8, 9] simply apply RANSAC sequentially, i.e., fit one structure, remove corresponding inliers, then repeat. This is risky because inaccuracies in the initial fits will be amplified in the subsequent fits [10]. Moreover, finding a stopping criterion for sequential fitting that accurately reflects the true number of structures is non-trivial. Methods based on clustering [11] or mode detection [3, 12] given the generated hypotheses are not affected by the dangers of sequential fitting. However, if there are insufficient hypotheses corresponding to the true structures, the genuine clusters will easily be overwhelmed by the irrelevant hypotheses. Consequently, these methods often miss the true structures or find spurious structures.

The inability to retrieve "good" hypotheses at sufficiently large quantities represents the fundamental obstacle to the satisfactory performance of previous methods. To address this limitation, we propose a novel hypothesis sampling scheme based on incremental computation of distances between *partial rankings* or *top-k lists* [13] derived from residual sorting information. Our approach enhances hypothesis generation in two ways: (1) The computed distances guide the sampling such that inliers from a *single* coherent structure are more likely to be simultaneously selected. This dramatically improves the chances of hitting all-inlier minimal subsets for *each* structure in the data. (2) The qualities of the generated hypotheses are evaluated based on the computed distances. This permits an on-the-fly filtering scheme to reject "bad" hypotheses. The outcome is a set of only the most promising hypotheses which facilitate a simple agglomerative clustering step to fit all the genuine structures in the data. Fig. 1 summarizes the proposed approach.

The rest of the paper is organized as follows: Sec. 2 describes how to derive data similarities from residual sorting information by comparing top-k lists. Sec. 3 describes our guided sampling scheme with simultaneous hypothesis filtering and incremental computations of distances between top-k lists. Sec. 4 describes how multi-structure fitting can be done by a simple agglomerative clustering on the promising hypotheses returned by our method. Sec. 5 presents results on synthetic and real data which validate our approach. Finally, we draw conclusions in Sec. 6.

2 Data Similarity by Comparing Top-k Lists

A key ingredient of our guided sampling scheme is a data similarity measure. This section describes how to derive such a measure from residual sorting information.

2.1 Top-k Lists from Residual Sorting Information

We measure the similarity between two input data based on the idea that if they are inliers from the same structure, then their preferences to the hypotheses as measured by residuals will be similar. Such preferences can be effectively captured by lists of ranked residuals.

Let $X = \{x_i\}_{i=1}^N$ be a set of N input data and $\theta = \{\theta_j\}_{j=1}^M$ a set of M hypotheses, where each hypothesis θ_j is fitted from a minimal subset of p points (e.g., p=2 for line fitting). For each datum x_i, we compute its absolute residual $r_i = \{r_1^{(i)}, r_2^{(i)}, \cdots, r_M^{(i)}\}$ as measured to M hypotheses. We sort the elements in r_i to obtain the list of sorted residual $\tilde{r}_i = \{r_{\lambda_1^{(i)}}^{(i)}, \cdots, r_{\lambda_M^{(i)}}^{(i)}\}$ such that $r_{\lambda_1^{(i)}}^{(i)} \leq \cdots \leq r_{\lambda_M^{(i)}}^{(i)}$. The top-$k$ list of data x_i is defined as the first k elements in the permutation $\{\lambda_1^{(i)}, \cdots, \lambda_M^{(i)}\}$, i.e.,

$$\tau_i = \{\lambda_1^{(i)}, \cdots, \lambda_k^{(i)}\}. \tag{1}$$

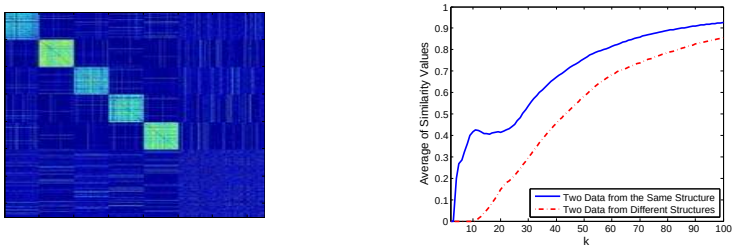
The top-k list τ_i essentially gives the top-k hypotheses preferred by x_i, i.e., x_i is more likely to be an inlier to the hypotheses which have higher rank.

2.2 The Spearman Footrule Distance

Given the top-k lists, we measure their similarity using the Spearman Footrule (SF) distance [13]. Let τ be a top-k list and D_τ a set of elements contained in τ. Denote the position of the element $m \in D_\tau$ in τ by $\tau(m)$. The SF distance between two top-k lists τ_i and τ_j is defined as

$$F^{(\ell)}(\tau_i, \tau_j) \;=\; \sum_{m \in D_{\tau_i} \cup D_{\tau_j}} \left|\tau_i'(m) - \tau_j'(m)\right|, \tag{2}$$

where $\ell > 0$ is the so-called location parameter (often set to k+1), $\tau_i'(m) = \tau_i(m)$ if $m \in D_{\tau_i}$; otherwise $\tau_i'(m) = \ell$, and τ_j' is similarly obtained from τ_j.



(a) (b)

Fig. 2. (a) Similarity matrix K for data shown in Fig. 1(a) (data is arranged according to structure membership for representation only) (b) The average of similarity values between two data from the same structure and two data from different structures under various k.

2.3 Measuring Similarity between Data

To measure the similarity between two data, we use the SF distance (Eq. 2) between their corresponding top-k lists. The similarity value between two data x_i and x_j is defined as

$$d(\tau_i, \tau_j) = 1 - \frac{1}{k \times \ell} F^{(\ell)}(\tau_i, \tau_j). \tag{3}$$

Note that we normalize $F^{(\ell)}(\tau_i, \tau_j)$ such that $d(\tau_i, \tau_j)$ is between 0 (dissimilar) and 1(identical). By comparing the top-k lists between all data, we obtain a $N \times N$ similarity matrix K with

$$K(i, j) = d(\tau_i, \tau_j), \tag{4}$$

where $K(i, j)$ denotes the element at its i-th row and j-th column. Fig. 2(a) shows an example of K, which is generated from the input data shown in Fig. 1(a). The evident block structures in K correspond to the 5 lines in Fig. 1(a). As shown in Fig. 2(b), across a wide range of k, the similarity value between two data from the same structure (solid) is higher than that from different structures (dotted).

3 Guided Sampling with Hypothesis Filtering

This section describes our guided sampling scheme which involves a simultaneous hypothesis filtering scheme. We also provide an efficient incremental update for computing the sampling weights.

3.1 Guided Sampling

We use the similarity matrix K (Eq. 4) to sample data in a guided fashion. Let $Q = \{s_u\}_{u=1}^{p}$ be the indices of data in a minimal subset of size p, where s_u are

indexed by the order in which they are sampled. The first element s_1 in Q is randomly selected from X. To sample the next element s_2, we use $K(s_1,:)$ as the weight to guide the sampling, i.e., the similarity values of all input data with respect to s_1. We set $K(s_1,s_1)$ to 0 to avoid sampling the same data again.

Suppose data $s_1,\cdots,s_u$ have been selected, then the next datum s_{u+1} is chosen conditionally on the selected data. Its sampling weight is defined as

$$K'(s_1,:)\ \cdot\ K'(s_2,:)\ \cdot\ \ldots\ \cdot\ K'(s_u,:), \tag{5}$$

where $\cdot$ is the element-wise multiplication and $K'(s_u,:)$ is just $K(s_u,:)$ with $K(s_u,s_u)=0$. Eq. 5 means that in order to have higher probabilities of being sampled, a datum need to be similar (measured by Eq. 3) to all the data that have been selected into the minimal subset.

3.2 Incremental Top-k Lists Comparison

Our sampling method computes an update to the similarity matrix K (Eq. 4) once a block (of size b) of new hypotheses are generated. This involves comparing top-k lists of ranked residuals. The computation of top-k lists can be done efficiently via merge sort. However, comparing top-k lists between all data, i.e., constructing K, can be computationally expensive. Here we provide efficient incremental updates for K that can substantially accelerate the computation.

As proved in [13], the SF distance (Eq. 2) can be equivalently computed as

$$F^{(\ell)}(\tau_i,\tau_j)=2(k-|Z|)\ell+\sum_{m\in Z}|\tau_i(m)-\tau_j(m)|-\sum_{m\in S}\tau_i(m)-\sum_{m\in T}\tau_j(m), \tag{6}$$

where $Z=D_{\tau_i}\cap D_{\tau_j}$, $S=D_{\tau_i}\backslash D_{\tau_j}$ and $T=D_{\tau_j}\backslash D_{\tau_i}$. In fact, S is simply the elements in D_{τ_i} but not in Z, i.e., $S=D_{\tau_i}\backslash Z$, similarly for T. Hence, we have

$$\sum_{m\in S}\tau_i(m)=\sum_{m=1}^{k}m-\sum_{m\in Z}\tau_i(m)=\frac{1}{2}k(k+1)-\sum_{m\in Z}\tau_i(m), \tag{7}$$

similarly for $\sum_{m\in T}\tau_j(m)$. By setting $\ell=k+1$ and using Eq. 7, we can rewrite Eq. 6 to be in terms of Z only,

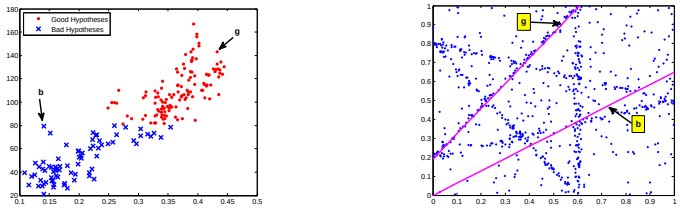
$$F^{(k+1)}(\tau_i,\tau_j)=(k+1)(k-2\,|Z|)+\sum_{m\in Z}\left(|\tau_i(m)-\tau_j(m)|+\tau_i(m)+\tau_j(m)\right). \tag{8}$$

Let A and B be two $N\times N$ symmetric matrices with zero on diagonal, and set the elements at the i-th row and the j-th column of A and B to

$$A(i,j)=|Z|\ \text{ and }\ B(i,j)=\sum_{m\in Z}\left(|\tau_i(m)-\tau_j(m)|+\tau_i(m)+\tau_j(m)\right). \tag{9}$$

From Equations 3, 4, and 8, the similarity matrix K can be constructed by

$$K=1-\frac{1}{k}(kI_N-2A)-\frac{1}{k(k+1)}B, \tag{10}$$



(a) (b)

Fig. 3. (a) Feature space where x axis is $f_m^{(1)}$ and y axis is $f_m^{(2)}$ (Best view in color). (b) An example of "good" (denoted "g") and "bad" (denoted "b") hypotheses.

where I_N is an $N \times N$ identity matrix. Observe from Eq. 9 that the matrices A and B can be efficiently updated by keeping track of the elements that move into or out of Z. This information is readily available from the merge sort. Once A and B are updated, K can be updated via Eq. 10.

3.3 Simultaneous Hypothesis Filtering

During sampling, we want to simultaneously filter hypotheses such that only a small but very promising subset remains. Let $R_m = \{R_m^{(1)}, \cdots, R_m^{(N)}\}$ be the absolute residual of N input data as measured to a hypothesis m. For this hypothesis, we construct a feature vector

$$f_m = \left[f_m^{(1)},\ f_m^{(2)}\right] = \left[\frac{\sum_{(i,j)\in E} K(i,j)}{|E|},\ \frac{|\Omega_m|}{\sum_{\{i|x_i\in\Omega_m\}} R_m^{(i)}}\right], \tag{11}$$

where $E = \{(i,j)|i \neq j \text{ and } x_i, x_j \in \Omega_m\}$ with $\Omega_m = \{x_i \in X \mid m \in D_{\tau_i}\}$, and K is the similarity matrix computed by Eq. 10. The set Ω_m contains all data that include the hypothesis m in their top-k lists. If the hypothesis m is "good", then Ω_m should contain many inliers from a structure. Hence, $f_m^{(1)}$, the average of similarity values between all data in Ω_m should be high. Moreover, the average residual of data in Ω_m should be low, i.e., high $f_m^{(2)}$. Therefore, we want to find a set of hypotheses which have high value in both $f_m^{(1)}$ and $f_m^{(2)}$. To this end, we apply k-means on the feature vectors (Eq. 11) to separate "good" and "bad" hypotheses. As illustrated in Fig. 3(a), the cluster whose center has larger norm (dots) contains good hypotheses. We incrementally maintain a set of "good" hypotheses as the guided sampling proceeds.

4 Multi-Structure Fitting

By leveraging the simultaneous hypothesis filtering, a set of "good" hypotheses is immediately available once the sampling is done. The minimal subsets of these "good" hypotheses should mainly contain inliers from different structures.

Hence, we can perform the final fitting by first, clustering the minimal subsets of "good" hypotheses, and then fitting geometric models to each cluster of data.

We use the agglomerative clustering (See [14] for a detailed description) to cluster the minimal subsets. First, we need to define a distance measure between two minimal subsets. From Fig. 2(a), we can see that if two data x_i and x_j are from the same structure, the rows $K(i,:)$ and $K(j,:)$ of the similarity matrix K must have higher values on the same dimension, implying that the sum of the element-wise multiplication of these two rows must have higher value. Based on this observation, we assign to each minimal subset Q_u a $1 \times N$ feature vector

$$\alpha_u = K(s_1^u,:) \ \cdot \ K(s_2^u,:) \ \cdot \ \ldots \ \cdot \ K(s_p^u,:), \tag{12}$$

where $\cdot$ is the element-wise multiplication and $s_1^u, \ \cdots, s_p^u$ are indices of data in Q_u. The distance between two minimal subsets Q_u and Q_v is given by

$$d(Q_u, Q_v) = \frac{1}{\|\alpha_u \cdot \alpha_v\|_1}, \tag{13}$$

where $\|\cdot\|_1$ denotes the L_1 norm. Using this distance measure, the clustering is then performed through the standard agglomerative clustering mechanism.

5 Experiments

We test the proposed method (ITKSF) on several synthetic and real datasets. To evaluate the efficiency of the proposed guided sampling scheme, we compare our method against 6 sampling techniques: Uniform random sampling in RANSAC (Random) [2], proximity sampling (Proximity) [9, 11], LO-RANSAC [15], Guided-MLESAC [4] and PROSAC [5].

In all experiments, the scale parameter of Proximity (σ^2 as in Equation 1 in [11]) is set to twice the squared average nearest neighbor distance. For LO-RANSAC, the inlier threshold is set to the average residual of inliers as measured to their corresponding structures. For PROSAC, T_N is set to 5×10^4. For our method, we fix $b = 10$ and $k = \lceil 0.1 \times t \rceil$ throughout, b being the block size (cf Sec. 3.2) and t the number of hypotheses generated so far. All experiments are run on a machine with 2.53GHz Intel Core 2 Duo processor and 4GB RAM.

5.1 Multiple 2D Line and Circle Fitting

We test the performance of various sampling methods on multiple 2D line and circle fitting under various numbers of gross outliers. Fig. 4 (Left) shows the test data. The inliers scale is set to 0.01, and the number of inliers per structure is 50 for lines and 80 for circles. We simulate the quality score required by PROSAC and Guided-MLESAC by probabilistically assigning higher scores to inliers than gross outliers. Each method is given 50 random runs, each for 2 CPU seconds.

As can be seen in Fig. 4 (Second column), in all cases the average percentage of all-inlier samples found by ITKSF within the given time budget is significantly

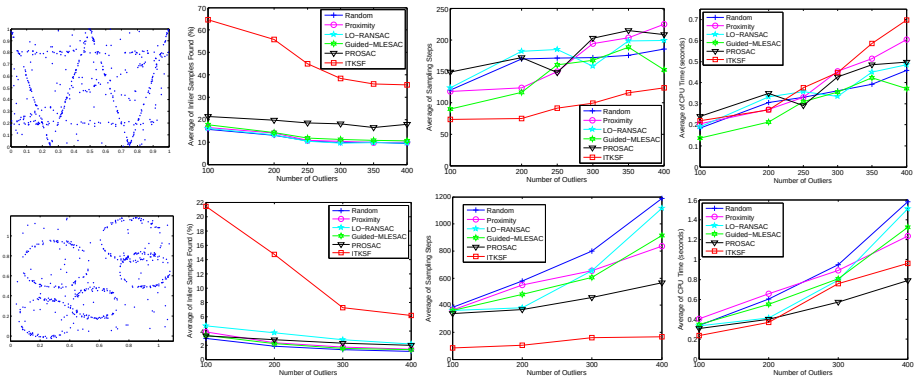


Fig. 4. Performance of various sampling methods on 2D lines and circles data under various numbers of gross outliers. First column: input data. Second column: the average percentage of all-inlier samples found within 2 CPU seconds. Third column: the average sampling steps (respectively CPU time, last column) needed to hit at least one all-inlier minimal subset for each structure.

higher than its competing methods. It also consistently requires less sampling steps than other methods to hit at least one all-inlier minimal subset for each structure (Third column). In terms of CPU time, ITKSF exhibits comparable performance to other methods (Fig. 4, Last column). Note that ITKSF simultaneously picks "good" hypotheses during sampling (Sec. 3.3). The CPU time spent on this operation is counted toward the reported CPU time for ITKSF.

Fig. 5 (Left) shows the "good" hypotheses returned by ITKSF after sampling. It is evident that they are concentrated on the genuine structures present in data. Fig. 5 (Center) shows that the minimal subsets of these promising hypotheses are correctly clustered by the agglomerative clustering (Sec. 4). Model fitting on

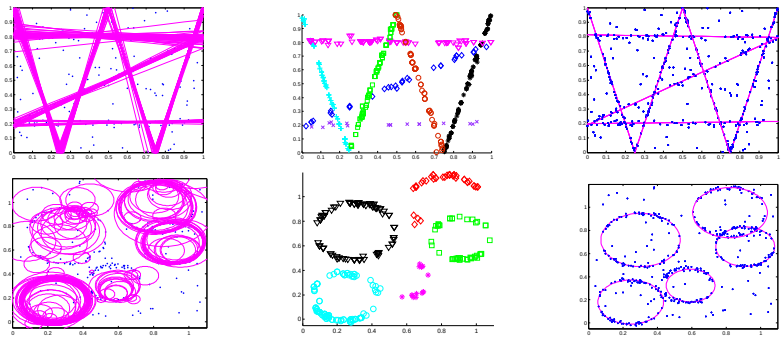


Fig. 5. Left: "good" hypotheses returned by ITKSF after sampling. Center: clusters found by clustering the minimal subsets of "good" hypotheses. Right: final fitting result.

(a) CollegeII (b) CollegeIII (c) CollegeI

Fig. 6. Clusters found by clustering the minimal subsets of the "good" hypotheses returned by our method (Best view in color).

each individual cluster of data then leads to the final fitting results shown in Fig. 5 (Right).

5.2 Homography Estimation

Our second set of experiments involves estimating multiple planar homographies on real images data.[1] For each image pair, keypoint correspondences (including false correspondences) and their matching scores are generated by SIFT matching[2] [16]. The image data with marked keypoint correspondences can be found in Table 1, the false correspondences are marked as yellow crosses. We use 4 correspondences to estimate a homography using Direct Linear Transformation [1]. Each method is given 50 random runs, each for 15 CPU seconds.

Table 1 summarizes our experimental results. We can see that ITKSF outperforms other methods across all performance measures in almost all cases; only on the College II data, it is slightly slower than PROSAC in terms of CPU time. Noticeably, our method finds much more all-inlier samples within the given time budget (Structures) than its competing methods. The average percentage of all-inlier samples (IS) found by ITKSF is up to an order of magnitude higher than other methods. Moreover, it is one of the only two methods that succeed in finding at least an all-inlier sample for each structure in all 50 runs.

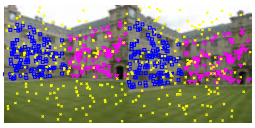
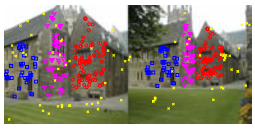
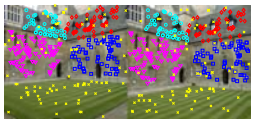
Fig. 6 shows that the minimal subsets of the hypotheses provided by our guided sampling procedure are indeed inliers, and they are correctly clustered according to their membership to a planar structure. Given these clusters, we can obtain multiple homographies by model fitting on each cluster of data.

5.3 Fundamental Matrix Estimation

We now evaluate the performance of various sampling methods for the task of fundamental matrix estimation on the Hopkins data.[3] The image data with marked keypoint correspondences (obtained from SIFT matching) are shown in Table 2. We use the standard 7-point algorithm [1] to estimate the fundamental matrix.[4] Each method is given 50 random runs, each for 30 CPU seconds.

[1] http://www.robots.ox.ac.uk/ vgg/data
[2] Code from http://www.vlfeat.org/ vedaldi/code/sift.html
[3] http://www.vision.jhu.edu/downloads/data/hopkins155/
[4] http://www.robots.ox.ac.uk/ vgg/hzbook/code/

Table 1. Performance of various sampling methods in multiple homographies estimation. We record the average CPU time (Time) (respectively sampling steps (Steps)) required and the number of random runs a method fails (Fail) to hit at least one all-inlier minimal subset for each structure. We also report the average percentage of all-inlier samples found within the given time budget (IS). The average number of all-inlier samples found for each structure (Structures) is separately listed in square bracket. The reported result is taken over successful runs only with the best result boldfaced.

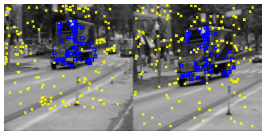
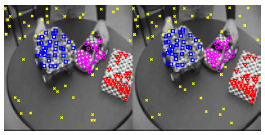
Data	Sampling Method	Time (seconds)	Steps	Fail	Structures	IS (%)
(a) CollegeII 	Random	1.03	298	0	[109,25]	1.84
	Proximity	0.58	163	0	[137,33]	2.48
	LO-RANSAC	0.46	146	0	[162,37]	2.56
	Guided-MLESAC	0.5	179	0	[131,31]	2.6
	PROSAC	**0.23**	86	0	[**354**,86]	7.46
	ITKSF	0.25	**45**	0	[337,**162**]	**33.74**
(b) CollegeIII 	Random	2.27	945	0	[6,20,125]	2.24
	Proximity	0.69	253	0	[18,21,151]	2.99
	LO-RANSAC	1.66	641	0	[6,22,148]	2.66
	Guided-MLESAC	1.74	685	0	[7,20,247]	4.21
	PROSAC	1.71	677	0	[6,12,189]	3.19
	ITKSF	**0.31**	**75**	**0**	[**228**,**98**,**305**]	**28.81**
(c) CollegeI 	Random	2.82	2080	15	[5,12,3,2]	0.37
	Proximity	8.72	953	**0**	[20,57,13,6]	1.69
	LO-RANSAC	2.82	1750	3	[7,21,5,3]	0.61
	Guided-MLESAC	4.91	1855	3	[4,17,8,7]	0.60
	PROSAC	5.17	1819	2	[6,18,6,4]	0.57
	ITKSF	**1.36**	**198**	**0**	[**74**,**93**,**30**,**8**]	**14.4**

In Table 2, we can see that in the case of single fundamental matrix estimation (on the Truck data) PROSAC is the most effective sampling method in terms of CPU time required to find an all-inlier minimal subset for the single structure, while ITKSF performs best in terms of the average number of all-inlier samples found within the given time budget. Overall, all sampling enhancement methods are effective on this simple single structure recovery task.

We now move to the more challenging case where more than one structure is present. Table 2 shows that previous methods fail disastrously on the Cars[5] and the Toy Cars data, which contain 2 and 3 structures, respectively. All previous methods fail to hit an all-inlier sample for each structure in 24%-100% of the given 50 runs, while ITKSF succeeds in every run. The average CPU time required by ITKSF to find at least one all-inlier sample for each structure is about 80% less than the best-performing competing method. Within the given time budget, the overall number of all-inlier samples found by ITKSF is again substantially larger than other methods. For instance, on the Toy Cars data, the

[5] The Cars data originally contains 3 structures. We use two in our experiment in order to create different levels of difficulties in the three datasets used in our experiments.

Table 2. Performance of various sampling methods in fundamental matrix estimation. The same notations as used in Table 1 are used here.

Data	Sampling Method	Time (seconds)	Steps	Fail	Structures	IS (%)
(a) Truck 	Random	5.94	1156	0	[5]	0.08
	Proximity	1.73	318	0	[16]	0.26
	LO-RANSAC	0.35	78	0	[10]	0.15
	Guided-MLESAC	0.23	62	0	[165]	3.23
	PROSAC	**0.008**	**1**	0	[58]	0.93
	ITKSF	0.2	27	0	[**471**]	**16.24**
(b) Cars 	Random	×	×	50	[0,0]	0
	Proximity	13.2	3997	41	[67,1]	0.69
	LO-RANSAC	13.63	4227	49	[25,1]	0.25
	Guided-MLESAC	6.94	3796	47	[334,1]	3.08
	PROSAC	×	×	50	[0,0]	0
	ITKSF	**2.19**	**450**	**0**	[**813**,**14**]	**16.77**
(c) Toy Cars 	Random	9.92	3015	49	[1,1,1]	0.02
	Proximity	15.09	4448	12	[5,2,3]	0.07
	LO-RANSAC	20.63	6092	48	[2,2,3]	0.04
	Guided-MLESAC	17.85	6662	36	[3,2,1]	0.04
	PROSAC	6.16	2436	36	[2,2,1]	0.03
	ITKSF	**1.93**	**305**	**0**	[**244**,**28**,**11**]	**8.11**

average percentage of all-inlier samples found by ITKSF is at least an order of magnitude higher than others.

Fig. 7 shows the clusters found by clustering the minimal subsets of the "good" hypotheses returned by ITKSF. It can be seen that each cluster is formed of inliers to one structure present in data. The fundamental matrix for each structure can therefore be effectively obtained by model fitting on each cluster of data.

(a) Truck (b) Cars (c) Toy Cars

Fig. 7. Clusters found by clustering the minimal subsets of the "good" hypotheses returned by our method (Best view in color).

6 Conclusions

We propose a novel guided sampling scheme based on the distances between top-k lists that are derived from residual sorting information. In contrast to

many existing sampling enhancement techniques, our method does not rely on any domain-specific knowledge, and is capable of handling multiple structures. Moreover, while performing sampling, our method simultaneously filters the hypotheses such that only a small but very promising subset remains. This permits the use of simple agglomerative clustering on the surviving hypotheses for accurate model selection. Experiments on synthetic and real data show the superior performance of our approach over previous methods.

References

1. Hartley, R.I., Zisserman, A.: Multiple View Geometry in Computer Vision. Second edn. Cambridge University Press (2004)
2. Fischler, M.A., Bolles, R.C.: RANSAC: A paradigm for model fitting with applications to image analysis and automated cartography. Comm. of the ACM **24** (1981) 381–395
3. Comaniciu, D., Meer, P.: Mean shift: A robust approach toward feature space analysis. IEEE Transactions on pattern analysis and machine intelligence **24** (2002) 603–619
4. Tordoff, B.J., Murray, D.W.: Guided-MLESAC: Faster image transform estimation by using matching priors. TPAMI **27** (2005) 1523–1535
5. Chum, O., Matas, J.: Matching with PROSAC- progressive sample consensus. In: CVPR. (2005)
6. Sattler, T., Leibe, B., Kobbelt, L.: SCRAMSAC: Improving RANSAC's efficiency with a spatial consistency filter. In: ICCV. (2009)
7. Stewart, C.V.: Robust parameter estimation in Computer Vision. SIAM Review **41** (1999) 513–537
8. Vincent, E., Laganiere, R.: Detecting planar homographies in an image pair. In: Image and Signal Processing and Analysis, 2001. ISPA 2001. Proceedings of the 2nd International Symposium on. (2001) 182–187
9. Kanazawa, Y., Kawakami, H.: Detection of planar regions with uncalibrated stereo using distributions of feature points. In: BMVC. (2004)
10. Zuliani, M., Kenney, C., Manjunath, B.: The multiransac algorithm and its application to detect planar homographies. In: IEEE International Conference on Image Processing, 2005. ICIP 2005. Volume 3. (2005)
11. Toldo, R., Fusiello, A.: Robust multiple structures estimation with j-linkage. In: European Conference on Computer Vision, Springer (2008) 537–547
12. Xu, L., Oja, E., Kultanen, P.: A new curve detection method: randomized Hough transform (RHT). Pattern Recognition Letters **11** (1990) 331–338
13. Fagin, R., Kumar, R., Sivakumar, D.: Comparing Top *k Lists*. In: Proceedings of the fourteenth annual ACM-SIAM symposium on Discrete algorithms, Society for Industrial and Applied Mathematics (2003) 36
14. Hastie, T., Tibshirani, R., Friedman, J.: The Elements of Statistical Learning. Springer (2009)
15. Chum, O., Matas, J., Kittler, J.: Locally optimized RANSAC. In: DAGM. (2003)
16. Lowe, D.: Distinctive image features from scale-invariant keypoints. IJCV **60** (2004) 91–110