GroupSAC: Efficient Consensus in the Presence of Groupings

Kai Ni
Georgia Institute of Technology
nikai@cc.gatech.edu

Hailin Jin
Adobe Systems Inc.
hljin@adobe.com

Frank Dellaert
Georgia Institute of Technology
dellaert@cc.gatech.edu

Figure 1. **RANSAC can be significantly sped up if modified to take advantage of additional grouping information between features, e.g., as provided by optical flow based clustering.** *Left: one of the two images for which we seek epipolar geometry, with inlier and outlier correspondences overlaid in blue and red respectively. Middle and Right: Two largest groups of tentative correspondences after clustering based on feature locations and optical flows. Notice their inlier ratios are much higher than the average ratio in the entire image.*

Abstract

We present a novel variant of the RANSAC algorithm that is much more efficient, in particular when dealing with problems with low inlier ratios. Our algorithm assumes that there exists some grouping in the data, based on which we introduce a new binomial mixture model rather than the simple binomial model as used in RANSAC. We prove that in the new model it is more efficient to sample data from a smaller numbers of groups and groups with more tentative correspondences, which leads to a new sampling procedure that uses progressive numbers of groups. We demonstrate our algorithm on two classical geometric vision problems: wide-baseline matching and camera resectioning. The experiments show that the algorithm serves as a general framework that works well with three possible grouping strategies investigated in this paper, including a novel optical flow based clustering approach. The results show that our algorithm is able to achieve a significant performance gain compared to the standard RANSAC and PROSAC.

1. Introduction

A common problem in computer vision and many other domains is to infer the parameters of a given model from data contaminated with noise and outliers. RANSAC [7] and its many variants [21, 16, 13, 14, 1, 3, 17, 22] have been a popular choice for solving such problems, owing to their ability to operate in the presence of outlier data points. The original RANSAC algorithm works in a hypothesis-testing framework: it randomly samples a minimum set of data points from which the model parameters are computed (*hypothesis*) and verified against all the data points to determine a fitting score (*testing*). This process (a trial) is repeated until a termination criterion is satisfied.

As the inlier ratio drops, RANSAC becomes exponentially slower. For problems with moderate inlier ratios, RANSAC is fairly efficient. For instance, for computing camera poses from points with known 3D positions (camera resectioning) using the 6-point DLT algorithm [10], it only takes about 60 trials to find the correct parameters with a 95% confidence from data points contaminated with 40% outliers. However, with a 20% inlier ratio, it requires over $45,000$ trials to achieve the same confidence. Such low inlier ratios exist in many practical problems. For instance, for the very same resectioning problem when dealing with internet images [19], one image of interest may observe tens or hundreds of known 3D points, and many of those point correspondences are spurious. For these problems, the standard RANSAC would simply fail because of the extraordinarily large number of trials needed.

In this paper, we make an assumption that there exists a *grouping* of the data in which some of the groups have a high inlier ratio while the others contain mostly outliers. Most problems of interest have such natural groupings. For example, in the case of wide-baseline matching, one can group the tentative correspondences according to optical flow. The motivation of optical flow based clustering comes from the fact that it is fairly easy for a human to distinguish inliers from outliers in an image such as Figure 1, without calculating the underlying epipolar geometry: the optical flow supplies a strong clue and can be used to group the tentative correspondences. Another natural grouping is derived from image segmentation, as in Figure 3.2: only segments visible in both images will contain inlier correspondences. The same grouping assumption exists in camera resectioning. Typically, the geometries of known 3D points in the target image were recovered from their 2D measurements in previously visited images. Hence those 3D points can be grouped according to which visited images they come from. All these examples have one thing in common that they reveal a correlation between the inlier probabilities of individual data points.

If such a grouping can be identified we propose a novel algorithm to exploit it, *GroupSAC*, improving on RANSAC and its variants by using a hierarchical sampling paradigm. In GroupSAC, we first randomly select some groups and then draw tentative correspondences from those groups. We will show that doing so is beneficial in that minimum sample sets drawn from fewer groups have a substantially higher probability of containing only inliers. Because its theoretical justification relies on several modeling assumptions, we adopt the design used in PROSAC [2] and only change the order in which minimum samples are considered. Hence the method gracefully degrades to the classical RANSAC if the assumed model for the data is wrong.

In addition, we show that GroupSAC can be improved upon if the chosen grouping has the property that larger groups tend to have higher inlier ratios. For instance, if we cluster according to optical flow as in Figure 1, most outliers are classified into smaller clusters because their flow is more random, while the the inliers tend to be consistent with each other and hence clustered into larger groups. As a result, the larger size of a group often indicates a higher inlier ratio. The ranking on groups so obtained can be integrated into GroupSAC to further improve its efficiency.

There is a lot of previous work along the lines of modifying the sampling strategy of standard RANSAC. It is usually achieved by making use of various types of prior information. NAPSAC [15] modifies the sampling strategy to select nearby points to deal with high-dimensional problems. Lo-RANSAC [4] conducts additional sampling among potential inliers. GASAC [18] adopts ideas from genetic algorithms to form samples with more likelihood of being all inliers. [20] suggests the use of additional prior information regarding absolute accuracy of each data point to guide the sampling procedure. PROSAC [2] uses relative accuracy information among the data points. Our algorithm differs from existing methods in the sense that we consider correlations between the inlier probabilities of the individual data points. While using additional prior information, existing algorithms have always assumed that data points are statistically independent, *i.e.*, no correlation among the probabilities of the data points being inliers has been exploited.

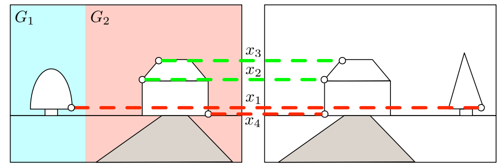


Figure 2. **An example of the wide-baseline matching problem**. *The inlier correspondences are colored in green, and the outliers are in red. The region G_2 in the left image contains part of the scene also visible in the right image. Hence this region not only contains more tentative correspondences (3 v.s. 1), but also has a higher inlier ratio (66.7% v.s. 0%).*

2. Group Sampling

In this section, we improve on the simple binomial model used in RANSAC by introducing a binomial mixture model for the case when a grouping in the data can be identified. This model suggests a novel hierarchical sampling scheme, which we term *group sampling*. The properties of group sampling are investigated, which establish the foundation of GroupSAC to be introduced in Section 3.

2.1. Simple Binomial Model of RANSAC

In the standard RANSAC algorithm and many of its variants, the probability of a data point x being an inlier is assumed to obey an i.i.d. Bernoulli distribution, i.e., the inlier probability of one point is assumed *independent* of any other data point. In RANSAC we consider the problem of estimating the parameters of a given model from N data points $\{x_j\}, j = 1 \ldots N$, contaminated with noise and outliers. We assume that the minimum number of points needed for computing the model parameters is m. For any minimum sample set S with m points, we have

$$\mathcal{I}_S \sim B(m, \epsilon) \tag{1}$$

where $\mathcal{I}_S$ is the number of inliers in S, and $B(m, \epsilon)$ is the binomial distribution in which ϵ is the parameter of a Bernoulli trial, i.e., the inlier probability of the set S in the RANSAC context. Therefore, the probability of all the points in S being inliers, $P(\mathcal{I}_S = m)$, is given by:

$$P(\mathcal{I}_S = m) = \prod_{j=1|x_j \in S}^{m} P(\mathcal{I}_j) = \epsilon^m \tag{2}$$

where $\mathcal{I}_j$ is an indicator variable signifying x_j is an inlier.

Although a lot of previous work has recognized the fact that ϵ is not necessarily the same for different data points, and has exploited this non-uniform property to speed up the sampling procedure [20, 2], in most if not all cases the inlier probabilities of different data points in those approaches are still considered independent of each other.

2.2. Binomial Mixture Model of GroupSAC

In this paper we make use of the fact that, for many problems, there exist groupings among the data points with the property that the groups tend to have either a high or a low fraction of inliers. One such example is the wide-baseline matching problem as shown schematically in Figure 2, in which the data points are the tentative correspondences between feature points in two images. Some regions in the first image will overlap with the scene seen from the second image, hence the corresponding point groups (G_2 in Figure 2) have a high chance to contain inliers, while the non-overlapping regions (G_1 in Figure 2) contribute no inliers. More examples from real image data will be discussed in Section 5.

More precisely, we assume that the probability of the inlier ratios in those groups can be modeled by a two-class mixture: the *high inlier class* (e.g., the overlapping region G_2 in Figure 2) and the *lower inlier class* (e.g., the non-overlapping region G_1). An obvious candidate to model each class is the beta distribution [6], the conjugate prior to the binomial distribution. However, the inlier ratio of the second class is usually close to zero, and we found that a mixture of two delta distributions suffices to model most problems, with the advantage of being simpler. In particular, we assume that the inlier ratio ϵ_i in any given group G_i is drawn from the following mixture

$$\epsilon_i \sim \pi_h \delta(\epsilon_0) + \pi_z \delta(0) \tag{3}$$

where π_h and π_z are the mixture weights for the high inlier class and the zero inlier class, with inlier ratios ϵ_0 and 0 respectively. Hence the probability of having $\mathcal{I}_{G_i}$ inliers in group G_i can be derived as

$$\begin{aligned} P(\mathcal{I}_{G_i}) &= \int_{\epsilon_i} P(\mathcal{I}_{G_i}|\epsilon_i) P(\epsilon_i) \\ &= \pi_h P(\mathcal{I}_{G_i}|\epsilon_i = \epsilon_0) + \pi_z P(\mathcal{I}_{G_i}|\epsilon_i = 0) \end{aligned} \tag{4}$$

In other words, the distribution on the number of inliers $\mathcal{I}_{G_i}$ in any given group is now a mixture of bimonials:

$$\mathcal{I}_{G_i} \sim \pi_h B(|G_i|, \epsilon_0) + \pi_z B(|G_i|, 0) = \pi_h B(|G_i|, \epsilon_0) \tag{5}$$

where $|G_i|$ is the number of data points in group G_i. Note that inliers are yielded by only a fraction π_h of the groups, referred to as *inlier groups*. So far we assume that each data point is associated with one group, whereas the multiple association case is introduced in Section 6.

2.3. Group Sampling

To deal with data points that can be grouped as described in Section 2.2, we introduce a novel hierarchical sampling scheme, namely *group sampling*, which is a two-step process: we first choose k groups yielding a *configuration* $\mathcal{G} = \{G_i\}$, $i = 1 \cdots k$, and then draw data points from the union of the groups G_i in $\mathcal{G}$. It is worth noting that the sampling strategy in RANSAC can be considered as a special case of group sampling in which all data points belong to the same group.

In group sampling, the probability of all the points in S being inliers can be computed as well. First, consider the

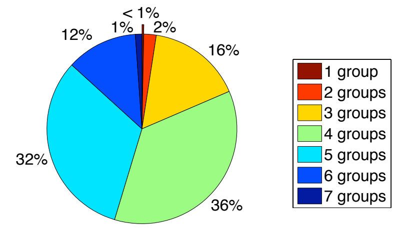


Figure 3. **The proportions of combinatorial numbers from 1,2,3...,7 groups, relevant for $m = 7$ when estimating a fundamental matrix**. *The number is less than 3% when sampling from one or two groups (the offset piece), and 20% from no more than three groups. In practice, GroupSAC finds an inlier sample from at most two or three groups, which results in large savings in computation time.*

easier case in which a set S_i contains $|S_i|$ samples (at least one) from a single group G_i. From Eq. (5), the probability of all points in S_i being inliers is given by

$$P(\mathcal{I}_{S_i} = |S_i|) = \pi_h \epsilon_0^{|S_i|} \tag{6}$$

Note that π_z does not appear above, as $B(|S_i|, 0) = 0$.

More generally, if a minimum sample set $S = S_1 \cup \cdots \cup S_k$ is drawn from configuration $\mathcal{G} = \{G_i\}_{i=1\cdots L}$, i.e., $S_i \subset G_i$, the probability of S containing only inliers is given by:

$$P(\mathcal{I}_S = m) = \prod_{i=1}^{|\mathcal{G}|} P(\mathcal{I}_{S_i} = |S_i|) = \pi_h^{|\mathcal{G}|} \epsilon_0^m \tag{7}$$

in which $|\mathcal{G}| = k$ is the number of the groups in configuration $\mathcal{G}$. Note that the difference between Eq. (7) and Eq. (2) is the factor $\pi_h^{|\mathcal{G}|}$ and the use of the inlier probability ϵ_0 for high inlier groups rather than the global ϵ.

2.4. Fewer Groups are Better

Under the assumptions made above, it is obvious from Eq. (7) that sampling data points from *fewer* groups yields a higher probability of obtaining an all-inlier sample S, as shown in the following corollary:

Corollary 2.1. *Given two minimum sample sets S_1 and S_2 drawn from configurations $\mathcal{G}_1$ and $\mathcal{G}_2$ respectively, we have*

$$|\mathcal{G}_1| \geq |\mathcal{G}_2| \Rightarrow P(\mathcal{I}_{S_1} = m) \leq P(\mathcal{I}_{S_2} = m). \tag{8}$$

Proof. Based on Eq. (7), we may write the inlier probabilities of S_1 and S_2 as

$$P(\mathcal{I}_{S_1} = m) = \pi_h^{|\mathcal{G}_1|} \epsilon_0^m \quad \text{and} \quad P(\mathcal{I}_{S_2} = m) = \pi_h^{|\mathcal{G}_2|} \epsilon_0^m. \tag{9}$$

We conclude the proof by noting that $0 < \pi_h \leq 1$. □

3. GroupSAC

In this section we present GroupSAC, a variant of RANSAC that gains additional efficiency by exploiting the properties of a grouping of the data points. GroupSAC is based on the same hypothesis-testing framework as that in the standard RANSAC, except for a new sampling scheme based on groupings.

3.1. GroupSAC Algorithm

To enable sampling from fewer groups, GroupSAC draws samples from increasing numbers of groups but is still able to achieve the same random sampling as in standard RANSAC, *i.e.*, each minimum sample set has the same chance of being sampled, when the computation budget is exhausted. Inspired by PROSAC [2], we use a heuristic ordering of group configurations without estimating the actual values of the mixture model parameters in Eq. (5).

More specifically, GroupSAC goes through all possible configurations in the order of their cardinalities. Let K be the total number of groups among all the data points, and all the configurations can be divided into $R = \min(m, K)$ subsets $\{\mathcal{C}_k\}_{k=1,\ldots,R}$ such that

$$\mathcal{C}_k = \{\mathcal{G}_u | \, |\mathcal{G}_u| = k\} \tag{10}$$

Accordingly, GroupSAC starts sampling from $\mathcal{C}_1$ until it finally reaches $\mathcal{C}_R$. For each subset $\mathcal{C}_k$, GroupSAC goes through each configuration $\mathcal{G}_u$ by drawing minimum sample sets from it. By the end of the R-th stage, all the configurations will have had their opportunity to be selected. The process can be summarized as below:

Algorithm 1 The GroupSAC Algorithm

Sort all the configurations according to Eq. (10) and (17).
Repeat the following

- Check the maximum rounds for the current configuration $\mathcal{G}_u$ Eq. (11). If reached, move to $\mathcal{G}_{u+1}$.
- Draw samples from all the groups in $\mathcal{G}_u$ such that each group at least contributes one data point.
- Estimate the parameter of the underlying model.
- Find the inliers for the new model and check the termination criteria in Section 3.2.

Similar to PROSAC, it is crucial to make the correct number of trials in each configuration $\mathcal{G}_u$. Assume our computation budget is T_0 trials in total for all possible $M_0 = \binom{N}{m}$ minimum sample sets, and let T_u be the number of trials for configuration $\mathcal{G}_u = \{G_i\}_{\mathcal{F}}$, in which $\mathcal{F}$ is the set of its group indices. T_u is chosen such that the sampling ratio of $\mathcal{G}_u$ is equal to that of the entire M_0 sample sets:

$$T_u / \left\| \bigcap_{i\in\mathcal{F}} G_i \right\| = T_0/M_0 \tag{11}$$

where $\left\| \bigcap_{i\in\mathcal{F}} G_i \right\|$ refers to the binomial coefficient of picking points from *all* the groups in $\mathcal{G}_u$, *i.e.*, each group contributes at least one point.

In order to compute $\left\| \bigcap_{i\in\mathcal{F}} G_i \right\|$ in Eq. (11), we invoke the inclusion-exclusion principle in combinatorics [9]:

$$\begin{aligned}\left\|\bigcap_{i\in\mathcal{F}} G_i\right\| &= \left\|\bigcup_{i\in\mathcal{F}} G_i\right\| - \sum_{i_1\in\mathcal{F}} \left\| \bigcup_{i_2\in\mathcal{F}\setminus i_1} G_{i_2}\right\| + \cdots \\ &+ (-1)^{n-2} \sum_{i_1,i_2\in\mathcal{F}} \left\|G_{i_1} \bigcup G_{i_2}\right\| + (-1)^{n-1}\sum_{i\in\mathcal{F}}\left\|G_i\right\|\end{aligned} \tag{12}$$

where $\left\| \bigcup_{i\in\mathcal{F}} G_i \right\|$ represents the binomial coefficient of picking points from *any* one or more groups in $\{G_i\}_{\mathcal{F}}$:

$$\left\| \bigcup_{i\in\mathcal{F}} G_i \right\| = \binom{\sum_{i\in\mathcal{F}} |G_i|}{m} \tag{13}$$

The proportion of the combinatorial numbers for $\{\mathcal{C}_k\}$ in a wide-baseline matching example is shown in Figure 3.

3.2. Termination Criteria

Since GroupSAC computes a model from a subset of the data points in a fashion similar to PROSAC, we adapt the termination scheme in the latter to our algorithm. First, non-randomness is checked to prevent the case that an incorrect model is selected as the final solution because of an accidental support from outliers. To be more precise, we want the probability that the model to be validated is supported by j random points is smaller than a certain threshold ψ: the minimal number of inliers n^* required to maintain the non-randomness can be computed as

$$\min\{n^* : \sum_{i=n^*}^{n} \beta^{i-m}(1-\beta)^{n-i+m}\binom{n-m}{i-m} < \psi\}, \tag{14}$$

where β is the probability of an incorrect model from a minimum sample set happening to be supported by a point that is not from the sample set, and n is the number of points in the current configuration. In our implementation, we use $\psi = 5\%$ for all the experiments.

The second termination condition is that the confidence that there does not exist another model which is consistent with more data points is smaller than a given threshold η:

$$(1 - \epsilon_{\mathcal{G}}^m)^p \le \eta \tag{15}$$

where $\epsilon_{\mathcal{G}}$ is the inlier ratio of the current configuration $\mathcal{G}$, and p is the number of random sampling trials. Note that Eq. (15) differs from the counterpart used in the standard RANSAC algorithm in the sense that only the points in $\mathcal{G}$ are considered. In this case, even if the entire data set has a low inlier ratio, GroupSAC is still able to terminate before reaching the maximum number of sampling trials.

4. Integrating Additional Orderings

Above we have shown that fewer groups should be preferred. Now we argue that in many cases groups with larger sizes should also be preferred. More specifically, in many applications there exists a correlation between the number of points in a group and its probability being an inlier group. For instance, in camera resectioning, the visited image that has more tentative correspondences to the known 3D points in the target image is more likely to be an image with true correspondences. The same holds for wide-baseline matching problems, in which one can expect more tentative correspondences from the true overlapping regions.

With this heuristic information, it is reasonable to order the groups in each class $\mathcal{C}_k$ according to the number of

Figure 4. **RANSAC can be significantly sped up if modified to take advantage of additional grouping information between features, e.g., as provided by image segmentation.** *Left and Right: two images for which we seek epipolar geometry. Middle: inliers (green) and outliers (red) overlaid on the image segments: notice the clear correlation between segmentation and inlier/outlier classification.*

points they contain and apply a PROSAC-like [2] algorithm to select groups. However, we find in experiments that this strategy does not always work well. The reason is that often outlier groups with many points are over-favored and this may lead to poor performance. Instead, we propose a different weighting scheme which works more reliably. Let S_1 and S_2 be any two size-m minimum sample sets from two configurations $\mathcal{G}_1$ and $\mathcal{G}_2$ respectively. We assume that a configuration with more points *in total* has a higher inlier probability than a configuration with fewer points:

$$\sum_{G_{i_1} \in \mathcal{G}_1} |G_{i_1}| \geq \sum_{G_{i_2} \in \mathcal{G}_2} |G_{i_2}| \Rightarrow P(\mathcal{I}_{S_1} = m) \geq P(\mathcal{I}_{S_2} = m). \tag{16}$$

The change from before is that we do not favor any individual group but instead we favor a configuration which consists a larger number of data points. This significantly reduces the risks of over-sampling from a small set of bad groups with more points.

The GroupSAC algorithm can be slighted modified to integrate the additional ordering based on the group sizes. Eq. (10) does not prefer any configurations inside $\mathcal{C}_k$, but with Eq. (16), we can order the configurations in $\mathcal{C}_k$ as $\{\mathcal{G}_u | \, |\mathcal{G}_u| = k\}$, such that for any two arbitrary $\mathcal{G}_{u_1}$ and $\mathcal{G}_{u_2}$ with $u_1 < u_2$, we have:

$$\sum_{G_i \in \mathcal{G}_{u_1}} |G_i| \geq \sum_{G_i \in \mathcal{G}_{u_2}} |G_i| \,. \tag{17}$$

In addition to the ordering of different configurations, another possible improvement is to order the points inside a configuration according to their inlier probabilities if such information is available. This basically amounts to applying PROSAC [2] in the inner loop. Since GroupSAC works in the same way as the standard RANSAC for a given configuration, it is straightforward to apply PROSAC, and the details are omitted here.

5. Application I: Wide-Baseline Matching

We first evaluate GroupSAC on the wide-baseline matching problem, in which there are two main challenges: small overlapping regions between the interested images and repetitive texture in the scene. Both challenges lead to low inlier ratios in tentative correspondences, and we will show how GroupSAC is able to alleviate this significantly.

First tentative correspondences are established by matching the SIFT features [11] detected in both input images. To apply our algorithm, we employ two strategies to group those tentative correspondences: one based on optical flow clustering and the other based on image segmentation. In both strategies, one image is arbitrarily chosen to serve as the reference image, and each tentative correspondence in that image is associated with *one* of the groups. The outliers in the tentative correspondences are then filtered by fitting a fundamental matrix [10].

To avoid the degeneracy when estimating fundamental matrix, we always start sampling from two groups and make a sanity check as in QDEGSAC [8]. Note that as GroupSAC only modifies the sampling stage of the standard RANSAC, it is rather easy to integrate tweaks on the other stages, e.g., QDEGSAC.

5.1. Grouping using Optical Flow Based Clustering

The first grouping strategy we tried for wide-baseline matching problems is to cluster the optical flow of the detected SIFT features. In Figure 1, it is easy to see that the inliers in a certain region tend to have consistent optical flow vectors, while those from outliers are much more random.

Specifically, we model each tentative correspondence as a 4-dimension vector $(u, v, \delta u, \delta v)$, in which u and v are the 2D feature position in the reference image, and δu and δv are the offset between the corresponding features in both images. We then define the distance between two tentative correspondences as the sum of squared distances, and the four dimensions are weighted by $(1, 1, 10, 10)$ respectively. The reason to use higher weights for the offset is that we favor making the optical flows consistent with each other while still allowing features to spread out somewhat. Based on the weighted distance, all the tentative correspondences can be clustered using mean shift [5] in which the bandwidth threshold is fixed to be proportional to the image size. The resulted two largest groups are illustrated in the 2nd and 3rd images of Figure 1.

5.2. Grouping using Image Segmentation

The second grouping strategy is based on image segmentation. The intuition behind this is that correspondences

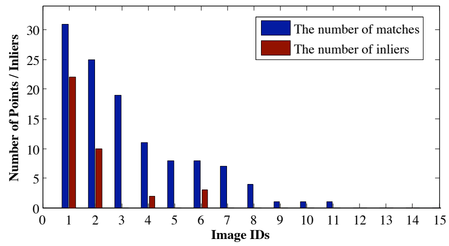


Figure 5. **The number of points and inliers associated with different groups for the hotel images shown in Figure 3.2.**

within one image segmentation are more likely to share the same inlier ratio. In particular, we generate a thumbnail for the reference image with the longer edge about 200 pixels, which is then segmented using Berkeley Segmentation Engine (BSE) as in [12]. The segmentation is efficient due to the small size of thumbnail image. The features in the reference image are then associated to the image segments with respect to their 2D position, and so are the tentative correspondences. One such example is shown in Figure 3.2. Note that GroupSAC does not heavily depend on the quality of the image segmentation, as it only yields a *rough* grouping of the correspondences.

Since the overlapping part between two images is typically small in challenging problems, most of the image segments only contain outlier groups as shown in Figure 5, In addition, the repetitive textures such as building facades in Figure 3.2 yields a lot of random false correspondences. In fact, the problems we are dealing with often have inlier ratio as low as 20%. However, the image segments corresponding to the overlapping view usually contain a considerable proportion of inliers, which become exactly the inlier groups that GroupSAC tries to locate during the sampling process. Also note that the inlier groups usually contain more tentative correspondences, which is consistent with our group ordering assumption.

5.3. Efficiency Comparisons

We compare the performance of the standard RANSAC, PROSAC [2] and the proposed GroupSAC on some real image data, which includes some well-known images (Table 5.3 A-D) as well as some (Table 5.3 E) that we collected by ourselves. We set the maximum allowed trials to 5000, T_0 to $250,000$ and η to 0.99, and all the results are averaged over 100 runs.

Due to low inlier ratios, the minimum number of trials to satisfy the termination criteria in RANSAC is exponentially high. As a result, RANSAC always takes the longest time and sometimes even exits without enough certainty when it reaches the maximal trial limit. PROSAC improves on RANSAC in all the tests but is still much slower than GroupSAC most of time. The main reason is that the residual used by PROSAC, the residuals in feature descriptor matching as proposed in [2], is indeed a point-wise evidence, while the grouping structure used by GroupSAC reveals the correlation between inlier points and is shown to provide stronger evidence. For example, the point-wise residual of outlier correspondences can be very small for scenes with repetitive textures, but the global grouping remains same in this case.

On the other hand, GroupSAC gains efficiency by fully exploiting the grouping between data points. For instance, in the hotel data set shown in Figure 5, if GroupSAC starts sampling from two groups, *i.e.*, $\mathcal{C}_2$, the first configuration in the sorted list is $\{G_1, G_2\}$, because it contains the most points according to Eq. (17). Since $\{G_1, G_2\}$ has a relatively high inlier ratio about 50%, GroupSAC has a very high chance to finish sampling earlier and hence is able to avoid outlier groups and inlier groups with low inlier ratios. In contrast, the standard RANSAC samples from all the data points with a much lower inlier ratio.

Another important insight is that the inlier ratios of the two largest groups in GroupSAC are much higher than those of the entire data set. This again proves that the underlying grouping structure can be recovered and used to speed up RANSAC dramatically. The high local inlier ratio also makes the proposed GroupSAC exit properly, because Eq. (15) can be satisfied as long as a certain configuration has a relatively high inlier ratio, while both RANSAC and PROSAC fails to exit with enough certainty on image pair F.

6. Multiple Associations

So far we have introduced how to do group sampling on data points in which each point has a single parent group, but in some problems a point may be associated with multiple groups. For instance, in the camera resectioning, we are given a set of points with 3D coordinates and their projections in the target image. These points can be naturally grouped according to the visited images that they are visible in, and the points from an image that overlaps with the target image have a higher chance to be inliers than those from an image that does not overlap with the target image.

More generally, let S be a minimum sample set associated with multiple configurations. Those configurations may yield different values when evaluating Eq. (7), and the one with the minimal cardinality gives the largest inlier probability. Since it is computationally intractable to evaluate the exact inlier probabilities in this case, we define $\mathcal{G}^*(S)$ as the one with the minimal cardinality, called the *minimum spanning configuration*, and assume the inlier probability of S is dominated by $\mathcal{G}^*(S)$:

$$P(\mathcal{I}_S = m) = \pi_h^{|\mathcal{G}^*|}\epsilon_0^m. \tag{18}$$

We then have a result similar to Corollary 2.1 that the probabilities of sample sets can be compared using the size of the corresponding minimum spanning configurations.

Corollary 6.1. *Let the probability of a group being an inlier group be π_h and the inlier ratio be $\hat{\epsilon}$. For two sample sets*

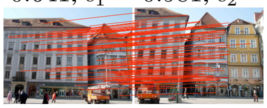
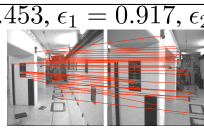
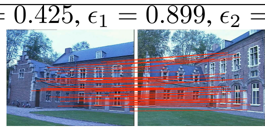
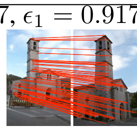
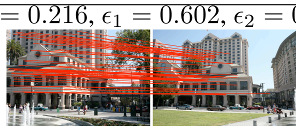
		k	$k_{\min}$	$k_{\max}$	vpm	time(sec)	speed-up
A: $\epsilon = 0.485$, $\epsilon_1 = 0.925$, $\epsilon_2 = 0.879$	RANSAC	898.6	667	1269	838.0	137.0	1
	PROSAC	250.1	136	512	127.8	12.9	10.6
	GroupSAC$_I$	52.1	42	92	353.8	3.74	36.6
	GroupSAC$_F$	10.25	7	17	451.0	1.13	**121**
B: $\epsilon = 0.641$, $\epsilon_1 = 0.981$, $\epsilon_2 = 0.934$	RANSAC	103.9	83	174	1025.5	21.3	1
	PROSAC	131.6	125	140	69.6	8.9	2.40
	GroupSAC$_I$	31.3	22	83	382.1	2.54	8.39
	GroupSAC$_F$	4.6	4	6	444.7	0.521	**40.8**
C: $\epsilon = 0.453$, $\epsilon_1 = 0.917$, $\epsilon_2 = 0.947$	RANSAC	772.6	396	1169	225.3	35.3	1
	PROSAC	15.9	8	39	15.9	0.174	202
	GroupSAC$_I$	44.1	27	59	39.9	0.363	97.2
	GroupSAC$_F$	6.2	4	11	78.8	0.132	**267**
D: $\epsilon = 0.425$, $\epsilon_1 = 0.899$, $\epsilon_2 = 0.973$	RANSAC	1823.1	1661	2316	501.3	168.6	1
	PROSAC	100.6	31	240	55.8	2.87	58.7
	GroupSAC$_I$	106.8	84	142	85.8	2.12	79.5
	GroupSAC$_F$	6.35	5	12	178.4	0.366	**460.7**
E: $\epsilon = 0.547$, $\epsilon_1 = 0.917$, $\epsilon_2 = 0.938$	RANSAC	702.7	588	886	190.3	25.1	1
	PROSAC	32.3	6	59	23.6	0.321	78.2
	GroupSAC$_I$	45.2	29	62	43.0	0.490	51.2
	GroupSAC$_F$	7.4	5	14	77.7	0.137	**183.2**
F: $\epsilon = 0.216$, $\epsilon_1 = 0.602$, $\epsilon_2 = 0.414$	RANSAC	5000	5000	5000	607.1	548.2	1
	PROSAC	5000	5000	5000	305.9	300.7	1.82
	GroupSAC$_I$	1383.6	942	2202	169.1	44.4	12.3
	GroupSAC$_F$	271.15	194	370	176.7	9.71	**31.0**
G: $\epsilon = 0.227$, $\epsilon_1 = 0.727$, $\epsilon_2 = 0.368$	RANSAC	5000	5000	5000	97.0	51.0	1
	PROSAC	5000	5000	5000	92.6	50.1	1.02
	GroupSAC	464.1	255	696	35.1	2.31	**22.1**

Table 1. **Comparisons between the standard RANSAC, PROSAC [2] and the proposed GroupSAC on wide baseline matching problems (A-F) and camera resectioning problems (F)**. *All the results are averaged over* 100 *runs.* k *is the number of models computed during the run, and vpm is the number of verifications per model. On top of the thumbnails, three inlier ratios are listed: the global* ϵ*, and the inlier ratio of the largest group* ϵ_1 *and the second largest group* ϵ_2 *after applying optical flow based clustering. A-F: GroupSAC*$_I$ *and GroupSAC*$_F$ *correspond to two grouping strategies: image segmentation and optical flow based clustering. Only part of the correspondences are drawn in the thumbnails for clear visualizations. G: The thumbnails are four representative images in a 23-frame house sequence. The first one is of interest in the camera resectioning problem.*

S_1 and S_2 whose minimum spanning configurations are $\mathcal{G}_1^*$ and $\mathcal{G}_2^*$ respectively, we have

$$|\mathcal{G}_1^*| \geq |\mathcal{G}_2^*| \Rightarrow P(\mathcal{I}_{S_1} = m) \leq P(\mathcal{I}_{S_2} = m). \quad (19)$$

Proof. The proof is similar to that of Corollary 2.1 and hence is omitted here. □

As a result, $\mathcal{G}$ in Eq. (10) is replaced with $\mathcal{G}^*$. The other part of the GroupSAC algorithm remains the same.

7. Application II: Camera Resectioning

Below we detail the experimental setup for the camera resectioning application and show that the assumptions underlying this paper indeed hold there. First we generate tentative correspondences by matching the SIFT descriptors [11] in the target image and those in the visited images that correspond to the known 3D points. All the tentative correspondences are passed to RANSAC, and the best inliers found are used to compute the underlying projection matrix using the 6-point DLT algorithm [10].

We demonstrate the performance of GroupSAC on an exemplar 23-frame image sequence taken around a house, whose correspondences and inliers along with their containing groups are plotted in Figure 6. The grouping in Figure 6 shows that only 11 out of 22 images contain the inliers, *i.e.*, half the groups are outlier groups. In fact, the correct matchings *only* exist in the images that overlap with the image of interest. On the other hand, a pure descriptor-based matching scheme tends to generate many false matchings because of the repetitive textures and the resulted ambiguities between their corresponding SIFT features. If we order the groups according to the number of their points as in Figure 6, the first half of images contains 7 inlier groups,

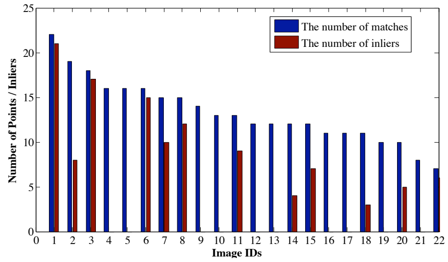


Figure 6. **The number of points and inliers associated with different groups for the house sequence in the camera resectioning problem**. *The blue bars indicate the numbers of tentative correspondences in each group/image, and the red bars indicate the numbers of inliers.*

while the second half only contains 4 inlier groups. This validates the assumption we made in Eq. (16) that orders different configurations. In terms of efficiency, GroupSAC again overwhelms RANSAC and PROSAC by a factor 22 as shown in Table 5.3.

In practice, a common way to solve the resectioning problem is to draw samples from the image with most of tentative correspondences. In fact, this heuristic approach is just the first step of the case that GroupSAC starts sampling from subset $\mathcal{C}_\infty$, and the largest group is exactly the image with the most tentative correspondences. Moreover, GroupSAC is inherently more robust since it handles other combinations in a probabilistically sound way.

8. Conclusions

We present a novel RANSAC variant that exploits group structures in data points to guide the sampling process. Our algorithm is able to handle problems with low inlier ratios and is shown to work magnitudes faster than the standard RANSAC and PROSAC in two important geometric vision problems: camera resectioning and wide-baseline matching. Our future work includes exploring other possible grouping strategies and trying to employ a more sophisticated model rather than the binomial mixture model.

9. Acknowledgements

This work is supported in part by Adobe Systems Incorporated. Kai Ni and Frank Dellaert gratefully acknowledge support by NSF awards Nrs. 0713162 and 0448111.

References

[1] D. Capel. An effective bail-out test for RANSAC consensus scoring. In *British Machine Vision Conf. (BMVC)*, 2005.

[2] O. Chum and J. Matas. Matching with PROSAC - progressive sample consensus. In *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2005.

[3] O. Chum and J. Matas. Optimal randomized RANSAC. *IEEE Trans. Pattern Anal. Machine Intell.*, 30(8):1472–1482, August 2008.

[4] O. Chum, J. Matas, and J. Kittler. Locally optimized RANSAC. In *Ann. Symp. German Assoc. Pattern Recognition*, 2003.

[5] D. Comaniciu and P. Meer. Mean shift: A robust approach toward feature space analysis. *IEEE Trans. Pattern Anal. Machine Intell.*, 24(5):603–619, 2002.

[6] M. Evans, N. Hastings, and B. Peacock. *Statistical Distributions*. Wiley, 3 edition, 2000.

[7] M. Fischler and R. Bolles. Random sample consensus: a paradigm for model fitting with application to image analysis and automated cartography. *Commun. Assoc. Comp. Mach.*, 24:381–395, 1981.

[8] J. Frahm and M. Pollefeys. RANSAC for (quasi-)degenerate data (QDEGSAC). In *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2006.

[9] M. Hall. *Combinatorial Theory, 2nd Edition*. Wiley-Interscience, 1998.

[10] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2000.

[11] D. Lowe. Distinctive image features from scale-invariant keypoints. *Intl. J. of Computer Vision*, 60(2), 2004.

[12] D. R. Martin, C. C. Fowlkes, and J. Malik. Learning to detect natural image boundaries using local brightness, color, and texture cues. *IEEE Trans. Pattern Anal. Machine Intell.*, 26(5):530–549, 2004.

[13] J. Matas and O. Chum. Randomized RANSAC with $T_{d,d}$ test. *Image and Vision Computing*, 22(10), 2004.

[14] J. Matas and O. Chum. Randomized RANSAC with sequential probability ratio test. In *Intl. Conf. on Computer Vision (ICCV)*, pages 1727–1732, 2005.

[15] D. Myatt, P. Torr, S. Nasuto, J. Bishop, and R. Craddock. Napsac: High noise, high dimensional robust estimation - it's in the bag. In *British Machine Vision Conf. (BMVC)*, pages 458–467, 2002.

[16] D. Nistér. Preemptive RANSAC for live structure and motion estimation. In *Intl. Conf. on Computer Vision*, 2003.

[17] R. Raguram, J.-M. Frahm, and M. Pollefeys. A comparative analysis of RANSAC techniques leading to adaptive real-time random sample consensus. In *Eur. Conf. on Computer Vision (ECCV)*, 2008.

[18] V. Rodehorst and O. Hellwich. Genetic algorithm sample consensus (GASAC) - a parallel strategy for robust parameter estimation. In *Proc. Conf. Computer Vision and Pattern Recognition Workshop*, 2006.

[19] N. Snavely, S. M. Seitz, and R. Szeliski. Modeling the world from internet photo collections. *Intl. J. of Computer Vision*, 80(2):189–210, November 2008.

[20] B. Tordoff and D. Murray. Guided sampling and consensus for motion estimation. In *Eur. Conf. on Computer Vision (ECCV)*, 2002.

[21] P. Torr and A. Zisserman. MLESAC: A new robust estimator with application to estimating image geometry. *Computer Vision and Image Understanding*, 78(1):138–156, 2000.

[22] C. Zach, A. Irschara, and H. Bischof. What can missing correspondences tell us about 3D structure and motion. In *IEEE Conf. on Computer Vision and Pattern Recognition*, 2008.