

BANSAC: A dynamic BAyesian Network for adaptive SAmple Consensus

Valter Piedade

Instituto Superior Técnico, Lisboa

valter.piedade@tecnico.ulisboa.pt

Pedro Miraldo

Mitsubishi Electric Research Labs

miraldo@merl.com

Abstract

RANSAC-based algorithms are the standard techniques for robust estimation in computer vision. These algorithms are iterative and computationally expensive; they alternate between random sampling of data, computing hypotheses, and running inlier counting. Many authors tried different approaches to improve efficiency. One of the major improvements is having a guided sampling, letting the RANSAC cycle stop sooner. This paper presents a new adaptive sampling process for RANSAC. Previous methods either assume no prior information about the inlier/outlier classification of data points or use some previously computed scores in the sampling. In this paper, we derive a dynamic Bayesian network that updates individual data points' inlier scores while iterating RANSAC. At each iteration, we apply weighted sampling using the updated scores. Our method works with or without prior data point scorings. In addition, we use the updated inlier/outlier scoring for deriving a new stopping criterion for the RANSAC loop. We test our method in multiple real-world datasets for several applications and obtain state-of-the-art results. Our method outperforms the baselines in accuracy while needing less computational time. The code is available at <https://github.com/merlresearch/bansac>.

1 Introduction

Outliers are one of the primary causes of poor performance in computer vision. Robust estimators are essential since imaging sensors suffer from several types of noise and distortions. Removing outliers is one of the initial and more relevant steps in many computer vision tasks, such as relative pose estimation [12, 28, 32, 35, 39, 50], camera localization [10, 42, 43, 51], and mapping [22, 26, 29, 36, 44, 45]. The gold-standard robust estimator is RANSAC (RANdom SAmple Consensus), introduced in [23]. RANSAC-based algorithms are iterative methods that, at each iteration: sample minimal sets, estimate a model, and run inlier counting. The output is the solution with the largest consensus.

The original RANSAC dates back to 1981. Over the

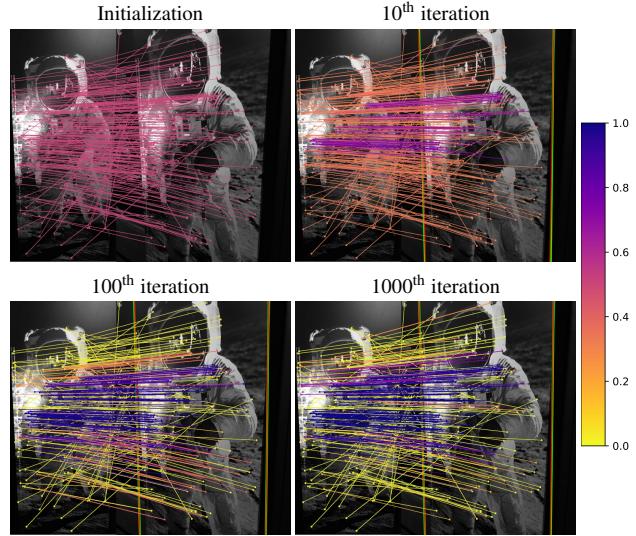


Figure 1. We run BANSAC in a homography estimation problem. We took a pair from the HPatches dataset [2] and show the updated inlier probabilities of data points (feature matches with color code at the right) over iterations. In the first row, from left to right, we show the probabilities at the start and iteration 10. The second row shows iterations 100 and 1000. For visualization purposes, we show only 250 randomly chosen matches.

years, many authors changed the original loop to alleviate some of its limitations. All these alternatives focus on improving the sampling process, getting a better hypothesis, improving the stopping criteria, or changing the inlier counting. Most modifications add significant gains in computational efficiency. This paper focuses on improving the sampling efficiency even further. The main question we want to tackle in this paper is: *Will changing the scoring weights over iterations help in sampling and defining the stopping criteria?* To answer this question, we propose BANSAC, a new sampling strategy for the RANSAC loop. Figure 1 illustrates our approach.

Previous methods such as [4, 11, 16, 37, 47, 48] focus on exploiting scoring priors or considering some geometric relationships. However, the best-performing methods

keep these scores fixed while running RANSAC. We focus on updating the scores online and using them for sampling minimal data. By modeling the problem with probabilities, we propose a new sampling strategy that uses a dynamic Bayesian network for updating the scores. These are the paper’s main contributions:

- A novel adaptive sampling strategy that uses a dynamic Bayesian network to update data points’ inlier scores. Our method does not need any prior information about the quality of the data, although it can use it;
- A new simple and intuitive stopping criterion using the updated scores; and
- Several experiments with multiple datasets show that our approach outperforms the best baselines in accuracy, being also more efficient.

We implemented BANSAC using C++, within the OpenCV USAC framework [1].

2 Related Work

Over the years, RANSAC has been improving in a variety of areas. [40] offers a single universal framework (USAC) that unites several improvements. Below we summarize some RANSAC improvements split into sampling and non-sampling strategies.

2.1 Sampling strategies

The original RANSAC [23] assumes that every data point has the same likelihood of being an inlier. Several new sampling improvements have been proposed. We split these methods into: heuristic [4, 14, 16, 27, 37, 48], probabilistic [8, 24, 33, 47, 49], and learning approaches [9, 11, 15].

Heuristic-based strategies: Heuristic-based strategies take advantage of problem-specific characteristics to guide sampling. NAPSAC [48] assumes that points in high-density areas are more likely to be inliers. The algorithm chooses the first point randomly and completes the sample within a certain distance from the first. NAPSAC often leads to local or degenerate models for more complex problems. PNAPSAC [4] improves some of NAPSAC issues by iteratively increasing the search space. One of the most used sampling strategies is PROSAC [16]. Using, *e.g.*, similarity scores between point matches, PROSAC prioritizes the sampling of points with better scores. A drawback of this method is that it cannot be applied in general since it requires some previously computed score. CS-RANSAC [27] argues that the matched features should neither be collinear nor adjacent to avoid degeneracies. CS-RANSAC defined the problem using a Constraint Satisfaction Problem (CSP) for homography matrix estimation. GroupSAC [37] assumes that data can be split into groups according to their

coordinates or based on the number of images observing the points.

Probability-based strategies: Probabilistic-based sampling strategies such as [24, 34, 47] focus on estimating prior probabilities for the data. These probabilities guide data selection during the sampling step. MLESAC [49] improves the hypotheses verification process in fundamental matrix estimation. Guided-MLESAC [47] further develops MLESAC by introducing a guided sampling modeled by two distinct distributions, one for matches and the other for mismatches. EVSAC [24] uses a Gamma and a Generalized Extreme Value distribution to model inliers and outliers, respectively. Although the problem differs from ours, [33] derives an approach that uses a probability for modeling inlier/outlier classification over iterations utilizing multiple match hypotheses (a single feature on an image matches more than one feature on the second image).

The closest work to ours is BAYSAC [8], which updates the inlier probabilities iteratively, using it to guide the sampling. At each iteration, after choosing a minimal set of data points and computing the respective model hypothesis, the method updates the probability of the data points in the minimal set based on how good the hypothesis was. Although this method updates the inlier probability at each iteration, these updates are limited to the sampled points, which do not perform well without a good prior. In this paper, we propose a new approach in which all the data points’ inlier probabilities are updated every iteration based on the inlier/outlier classifications.

Learning-based strategies: There has been widespread use of neural networks in many areas of computer vision, RANSAC sampling being no exception. NG-RANSAC [11] focuses on sampling by learning to estimate matching scores for the input correspondences for relative pose problems. Instead of scoring data points, DSAC [9] learns to score a set of previously computed hypotheses. NeFSAC [15] predicts the probability that a minimal sample leads to an accurate solution, thus preventing the estimation of models using bad minimal samples.

Our method does not require training. However, pre-computed matching scores from learning-based solutions can be given as input to BANSAC.

2.2 Non-sampling strategies

Below we list some key works on improvements to RANSAC concerning the inlier threshold, inlier counting, local optimization, and stopping criteria.

Inlier threshold: RANSAC uses the inlier ratio to select the best model (largest consensus). To compute the inlier ratio, a problem-dependent threshold is required. To avoid setting this parameter, MINPRAN [46] proposes to model

it using the model parameters. Alternatively, MAGSAC [6] and MAGSAC++ [7] reformulate the problem to use a weighted least squares fitting for model evaluation, using point scores as weights.

Inlier counting: Inlier/outlier classification is computationally heavy since, in each iteration, all the data needs to be checked. Some authors have developed strategies to avoid scoring all the points every iteration, [3, 13, 17, 30, 31]. Others check if the estimated models are valid, avoiding the scoring process for invalid models, such as [3, 19–21, 25]. To avoid scoring unnecessary data points, a bail-out test is proposed in [13]. The scoring stops when the current model fails to have a higher inlier count than the best model. SPRT [17, 31] estimates a likelihood ratio to decide if a model is good using the minimum possible amount of data.

Local optimization: To improve accuracy, some authors added a new step called local optimization. LO-RANSAC in [18] recomputes the model parameters when a new best model is found, using only the inliers. In [5], a method called Graph-Cut RANSAC is proposed. It takes advantage of the spatial coherence of the data to refine the estimated model by assuming that close neighbor points should have an equal classification in inlier/outlier. The data is arranged in a graph with edges between nearby points and is minimized by an energy cost function that penalizes neighbor points with different classifications.

Stopping criteria: The stopping criterion checks if RANSAC found a good enough solution and can exit the loop. The vanilla RANSAC in [23] estimates how many iterations are needed until one all-inlier model hypothesis is selected based on the inlier ratio of the so-far best model. It stops when the current number of iterations is higher than the one needed for getting one all-inlier model. Instead of attempting to guarantee the best solution, [38] sets a real-time limit to get an estimate. PROSAC [16] adds to the RANSAC criterion a condition to end when the probability of having a certain number of outliers in the current best set is lower than a predefined threshold. Finally, SPRT [31] terminates when the likelihood of missing a solution with a higher inlier set than the best solution found so far is below a certain threshold.

3 Background and Notations

RANSAC is an iterative method for solving a generic problem of type $f(\mathbf{x}, \theta) = 0$, where \mathbf{x} is some data satisfying the model θ . For simplicity, with a small abuse of notation, we call \mathbf{x} a data point, because it can represent other types of features such as matches. The method iterates for a maximum of K iterations while alternating between, 1) sampling $\mathcal{S}^k \subset \mathcal{Q}$ data points, where $\mathcal{Q} \triangleq \{\mathbf{x}_1, \dots, \mathbf{x}_N\}$; 2) computing model hypothesis θ^k ; and 3) doing inlier

Table 1. Summary of some important notations used in this paper.

Notation	Description
$\mathcal{X}^k \triangleq \{x_1^k, \dots, x_N^k\}$	Inlier/outlier guesses per iteration.
$\mathcal{X}_n \triangleq \{x_n^0, \dots, x_n^k\}$	All inlier/outlier guesses for \mathbf{x}_n .
$\mathcal{X} \triangleq \{\mathcal{X}^k, \mathcal{X}_n\}$	All data point inlier/outlier guesses.
$\mathcal{C}_n \triangleq \{c_n^1, \dots, c_n^k\}$	Set of inlier evidences per iteration.
$\mathcal{C} \triangleq \{\mathcal{C}^k, \mathcal{C}_n\}$	All data point inlier/outlier evidences.
$A^{0:m} \triangleq \{A^0, \dots, A^m\}$	Order subset. A can be a variable or set.
$\mathcal{P}^k \triangleq \{P(x_1^k C_1^{1:k}), \dots,$ $P(x_N^k C_N^{1:k})\}$	Set of guesses for data points inlier/outlier probabilities at k , given the evidence $c_n^{1:k}$.

counting, i.e., get $\mathcal{C}^k \triangleq \{c_1^k, \dots, c_N^k\}$, where c_n^k is the inlier/outlier classification of \mathbf{x}_n at iteration k . The output is the best-scored model and best inlier/outlier classification set, here denoted as $\{\theta^*, \mathcal{C}^*\}$.

4 BANSAC Method

This paper focuses on deriving an efficient sampling of data points, i.e., getting \mathcal{S}^k . We note that getting θ^k from \mathcal{S}^k is problem-dependent; BANSAC is independent of the problem. We take two simple assumptions:

1. We assume that sampling data points with higher inlier scores gives a better model hypothesis; and
2. As we iterate through RANSAC, we get a better sense of whether \mathbf{x}_n is an inlier or outlier.

An intuitive way of deriving a sampling technique with the above assumptions is to have changeable inlier/outlier scores of data points per iteration. Our method updates the scores based on inlier/outlier classifications from previous iterations. For modeling changeable scores over iterations, we define unknown variables, namely x_n^k representing the best guess for inlier/outlier classification for a data point \mathbf{x}_n at iteration k . For modeling the scores for each data point at each iteration, we use probabilities, i.e., $P(x_n^k)$, where x_n^k can take the values of inlier and outlier. Unfortunately, we do not have a direct way of measuring $P(x_n^k = \text{inlier})$. Instead, in this paper, we use inliers/outliers classifications obtained from previous iterations as evidence, i.e., for sampling we use $P(x_n^k = \text{inlier} | C_n^{1:k})$, where $C_n^k \triangleq c_n^k$ = inlier/outlier. Table 1 lists some important notations we use in our derivations.

In the following subsection, we give a method overview.

4.1 Algorithm outline

Algorithm 1 outlines the proposed method. The RANSAC-based loop starts with an optionally given \mathcal{P}^0 , which can be obtained from matching scores (or other initial guesses), or set to a predefined value (i.e., not using previously pre-computed scores). At each iteration k , we generate a minimal set $\mathcal{S}^k \in \mathcal{Q}$ via weighted sampling,

Algorithm 1: BANSAC algorithm outline

Input – Data \mathcal{Q} , and (optional) pre-computed scores \mathcal{P}^0
Output – Best model θ^* , and \mathcal{C}^*

```

1  $k \leftarrow 1$ ;
2 while  $k < K$  do
3    $\mathcal{S}^k \leftarrow \text{weighted\_sampling}(\mathcal{Q}, \mathcal{P}^{k-1})$ ;     $\triangleright$  Sec. 4.3
4    $\theta^k \leftarrow \text{hypothesis}(\mathcal{S}^k)$ ;
5    $\mathcal{C}^k \leftarrow \text{model\_evaluation}(\mathcal{Q}, \theta^k)$ ;
6    $\theta^*, \mathcal{C}^* \leftarrow \text{best\_model}(\mathcal{C}^k, \theta^k)$ ;
7    $\mathcal{P}^k \leftarrow \text{update\_probabilities}(\mathcal{C}^k, \mathcal{X}^{0:k-1})$ ;  $\triangleright$  Sec. 4.2
8   if stopping_criteria( $\mathcal{P}^k$ ) break;       $\triangleright$  Sec. 4.4
9    $k \leftarrow k + 1$ ;

```

using \mathcal{P}^{k-1} as weights (see notations in Tab. 1), which is shown in Line 3. Next, we compute the hypothesis model θ^k , run inlier counting, and update the best model if needed, as described in Lines 4 to 6. Line 7 updates the probabilities for the next iteration, i.e., computes \mathcal{P}^k . In addition to the new sampling and updating the probabilities, using \mathcal{P}^k , we derive a new stopping criterion for our RANSAC-based loop, which we use at Line 8.

Section 4.2 derives the probabilistic model, Sec. 4.3 describes the weighted sampling strategy, and Sec. 4.4 introduces the new stopping criterion.

4.2 Probabilistic model and inference

We describe our method by modeling data points' inlier/outlier probability.

Probabilistic model: Since we want the data point probabilities to change over iterations, we use a dynamic Bayesian network (DBN) as our probabilistic model. A DBN is a probabilistic graphical model that uses variables (nodes representing states and observations) and their conditional dependencies (edges) in a directed acyclic graph (see [41]). Starting with the variables, at each iteration k , we have the data points state and the inlier/outlier classifications (evidence) obtained so far (i.e., from 1 to k). So, for iteration k , we have nodes \mathcal{X} and \mathcal{C} (see notations in Tab. 1).

For the graph edges, we have the following constraints:

1. We want our sampling to be general. Then, for a certain iteration k , we assume that the inlier/outlier probabilities of different data points and the classifications are independent of each other. At each iteration k , the probability of x_n^k is updated based on the x_n 's previous probabilities and the previous inlier/outlier classifications (our evidence needs to constrain only the next probability estimate);
2. The inlier/outlier evidence at each iteration, c_n^k , depends only on the model θ^k , which depends on \mathcal{S}^k and, by consequence, on x_n^{k-1} .

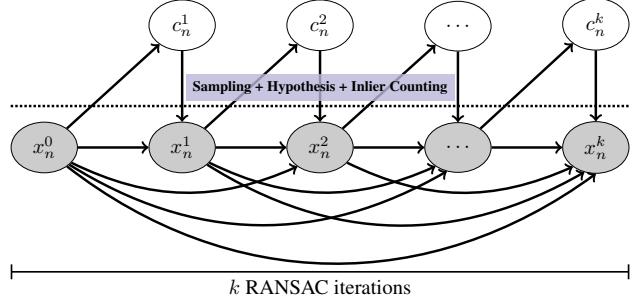


Figure 2. Dynamic Bayesian network proposed to model the data inlier/outlier probabilities for data point x_n .

Formally, for iteration k , we have the following constraints:

$$x_n^k \perp \mathcal{X} \setminus \mathcal{X}_n, \mathcal{C} \setminus c_n^k \mid \mathcal{X}_n, c_n^k \quad (1)$$

$$c_n^k \perp \mathcal{X} \setminus x_n^{k-1}, \mathcal{C} \mid x_n^{k-1}. \quad (2)$$

The first important consequence of these conditionally independent constraints is that we have an independent DBN for each data point, x_n , each with its own weights. Then, for iteration k , we define the DBN per data point as shown in Fig. 2.

Markov assumptions: The DBN derived above has the unbounded problem of increasing exponentially with the number of iterations¹. To solve this problem, we follow the typically used Markov assumptions. For simplicity, here we derive the first-order Markov assumption for our problem. The second and third-order assumptions are tested in Sec. 5.1, and the derivations are provided in the supplementary materials; they are similar but slightly more intricate. In addition to the conditionally independent constraints derived above, we have

$$x_n^j \perp x_n^{0:j-2} \mid x_n^{j-1}, c_n^j, \forall j, \quad (3)$$

which means

$$P(x_n^j \mid x_n^{0:j-1}, c_n^j) = P(x_n^j \mid x_n^{j-1}, c_n^j), \quad \forall j. \quad (4)$$

Now, by applying the chain rule of probabilities, we write the joint probability at iteration k as

$$P(x_n^{0:k}, c_n^{1:k}) = P(x_n^0) \prod_{j=1}^k \phi_j(x_n^{0:k}, c_n^{1:k}), \quad (5)$$

where

$$\phi_j(x_n^{0:k}, c_n^{1:k}) = P(x_n^j \mid x_n^{j-1}, c_n^j)P(c_n^j \mid x_n^{j-1}). \quad (6)$$

¹Would need huge conditional probability tables (CPT) and the increase in computational cost.

Exact inference: In our sampling strategy, we use $P(x_n^k = \text{inlier} | C_n^{1:k})$, for all n . This means that we are doing inference of $x_n^k = \text{inlier}$, with evidences C_n , and hidden variables $\mathcal{X}_n \setminus x_n^k$. Given Eq. 6, after some derivations, the exact inference is given by (see [41, Sec.14.4]):

$$P(x_n^k = \text{inlier} | C_n^{1:k}) = \alpha \Phi(x_n^k = \text{inlier}, x_n^{0:k-1}, C_n^{1:k}), \quad (7)$$

where α is a normalization factor², and

$$\begin{aligned} \Phi(x_n^k, x_n^{0:k-1}, C_n^{1:k}) &= \\ &\sum_{x_n^{k-1}} \phi_k(x_n^{0:k}, c_n^{1:k}) \sum_{x_n^{k-2}} \phi_{k-1}(x_n^{0:k}, c_n^{1:k}) \\ &\cdots \sum_{x_n^1} \phi_2(x_n^{0:k}, c_n^{1:k}) \sum_{x_n^0} \phi_1(x_n^{0:k}, c_n^{1:k}). \end{aligned} \quad (8)$$

Notice x_n^k can take two values; it can be either inlier or outlier. This means that $\Phi(\cdot)$ is a 2-dimension tuple; in Eq. 7 we pick the case $x_n^k = \text{inlier}$. In addition, the summations for x_n^j have two terms for all j .

A convenient result of Eq. 8 is that $\Phi(\cdot)$ can be computed recursively as follows:

$$\begin{aligned} \Phi(x_n^k, x_n^{0:k-1}, C_n^{1:k}) &= \\ &\sum_{x_n^{k-1}} \phi_k(x_n^{0:k}, c_n^{1:k}) \Phi(x_n^{k-1}, x_n^{0:k-2}, C_n^{1:k-1}) \end{aligned} \quad (9)$$

for $k \geq 1$, and $\Phi(x_n^0, -, -) = P(x_n^0)$. This means that at each iteration k , we can use the $\Phi(\cdot)$ computed from the previous iteration, with no additional computational cost when increasing the number of iterations. The pseudo-code for the probability update is in the supplementary materials.

We tried several alternatives for the Conditional Probability Tables (CPT) in Eq. 6. In our experiments, for $P(c_n^k | x_n^{k-1})$, we use a Leaky ReLU, weighted using the inlier counting. For the CPT of $P(x_n^k | x_n^{k-1}, c_n^k)$, we get the probabilities empirically. Due to space limitations, we present more details about the CPTs in the supplementary material, including some experiments.

4.3 Weighted sampling

Weighted sampling aims at getting the minimal set $\mathcal{S}^k \subset \mathcal{Q}$ for computing hypothesis θ^k . To increase the chances of only selecting inliers, we take the estimated probabilities \mathcal{P}^{k-1} and create a weighted discrete distribution. Data points with a higher probability of being an inlier will have a higher chance of being sampled.

In addition to directly using \mathcal{P}^{k-1} for weighting the discrete distribution, we tested using various activation functions such as leaky ReLU, sigmoid, and tanh functions. In

²From complementary rule, α is such that $P(x_n^k = \text{inlier}, x_n^{0:k-1}, C_n^{1:k}) + P(x_n^k = \text{outlier}, x_n^{0:k-1}, C_n^{1:k}) = 1$

the experiments, we directly use \mathcal{P}^{k-1} . Due to space constraints, we show results with other activation functions in the supplementary materials.

4.4 Stopping criterion

Besides using the probabilities \mathcal{P}^k in sampling, we also exploit them in defining a stopping criterion. We know that \mathcal{P}^k will influence the sampling, meaning that, after reaching a low probability threshold (which we denoted as τ), we can say that a data point will not be considered for sampling. Based on this idea, we derive a new, simple stopping criterion. At each iteration k , we add the following steps to Algorithm 1:

1. In Line 6, we keep the smallest number of outliers (best case scenario so far), denoted as O^* ; and
2. After updating the probabilities in Line 7, we compute the number of data points with $P(x_n^k = \text{inlier} | C_n^{1:k})$ lower than a predefined threshold τ , which we denote here as \tilde{O}^k .

Our stopping criterion is triggered when $\tilde{O}^k \geq O^*$, which we check at each iteration in Line 8. This means that the current best model has a bigger or equal number of inliers than the set of $P(x_n^k = \text{inlier} | C_n^{1:k})$ being selected for sampling; it can be assumed we have only inliers in the sampling set. We exit the loop because there is a low chance of getting a better solution.

Our stopping criterion can be used alone or added to existing ones, such as the RANSAC [23], SPRT [31] or PROSAC [16]. We do experiments on the combinations of these criteria in the supplementary materials.

4.5 Degenerative configurations

Degenerate configurations in minimal solvers can lead to poor RANSAC estimates. BANSAC is susceptible to such settings as other RANSAC-based methods. In a worst-case scenario, since the probability updates depend on the inlier ratio, after sampling degenerate configurations, BANSAC can create a bias towards degenerative sampling. However, it may take several iterations of consecutive degenerate configurations before that bias has some impact on the sampling. Moreover, since each data point always has a chance of being selected, even if it is minimal, BANSAC can reverse that bias as soon as new non-degenerative solutions are obtained.

Although possible, we highlight that we did not experience any issues with degenerative configurations during our experiments. In addition, we are using the OpenCV USAC framework that has built-in methods to handle many of those cases. When those degenerative configurations are detected, the probabilities are not updated.

5 Experiments

We evaluate BANSAC in three computer vision problems: calibrated relative pose, uncalibrated relative pose, and homography estimation. We start our experiments with an ablation study to access the need for different Markov assumption orders in Sec. 5.1. Next, we evaluate the efficiency of our method concerning a varying number of fixed iterations (no stopping criterion is used) and for a varying inlier ratio, in Secs. 5.2 and 5.3, respectively. Section 5.4 offers results for 1) calibrated relative pose, 2) uncalibrated relative pose, 3) homography estimation, and 4) same as 1), 2) and 3) with the addition of local-optimization.

Evaluation metrics: We use the mean Average Accuracy (mAA) with thresholds at 5 and 10 degrees for the calibrated and uncalibrated relative pose problems, and at 5 and 10 pixels for homography estimation. We borrow the evaluation scripts for rotation, translation, and homography error metrics from the “RANSAC in 2020” tutorial package³. Additionally, we show the average execution time.

Methods: We utilize two variations of our method: with and without pre-computed scores. Without pre-computed scores, we refer to our method as BANSAC, and we use a combination of BANSAC (Sec. 4.4) and SPRT [31] stopping criteria. When using pre-computed scores, we refer to our method as P-BANSAC, and we use a combination of BANSAC and PROSAC’s stopping criteria [16]. The stopping criteria for BANSAC and P-BANSAC are chosen for better accuracy (keeping a reasonable speed) and computational efficiency, respectively. In addition, BANSAC and P-BANSAC have some parameters that need to be set, namely the CPT values, which we kept fixed for all experiments.

As baselines, we use RANSAC [23] and NAPSAC [48] when not using pre-computed scores, and P-NAPSAC [4] and PROSAC [16] otherwise. Only the sampling and the stopping criterion vary between all methods. All the remaining RANSAC components are the same.

For all methods, at the end of the RANSAC cycle, the inliers of the best model are used to refine the final solution using a non-minimal solver, following a typical robust estimation pipeline. By default, no local optimization is run inside the RANSAC loop, unless we explicitly say so.

Experiments evaluating BANSAC using different stopping criteria and CPT parameters are present in the supplementary materials. BaySAC [8] is not shown in the paper since its results do not differ much from RANSAC. In the supplementary materials, we evaluate BANSAC against RANSAC and BaySAC using synthetic data.

Settings: For the calibrated relative pose problem, we use an error threshold of $1e-3$ (normalized points), 1000 maximum iterations, and a confidence of 0.999. For the uncali-

Table 2. *Ablation study on the Markov assumption.* We run BANSAC using the 1st, 2nd, and 3rd Markov assumption orders on a calibrated relative pose problem.

Metrics	Markov assumptions		
	1st Order	2nd Order	3rd Order
Rotation mAA (5°) \uparrow	0.836	0.822	0.808
Rotation mAA (10°) \uparrow	0.864	0.853	0.843
Translation mAA (5°) \uparrow	0.775	0.755	0.739
Translation mAA (10°) \uparrow	0.825	0.811	0.798
Avg. execution time [ms] \downarrow	13.9	12.2	11.8

brated relative pose problem, we use an error threshold of 0.5 pixel, 10000 maximum iterations, and a confidence of 0.999. For the homography estimation problem, we use an error threshold value of 1 pixel, 1000 maximum iterations, and a confidence of 0.999. In all these problems, we set our proposed stopping criteria threshold τ to 0.01 in BANSAC and 0.1 in P-BANSAC (see Sec. 4.4).

Datasets: For the relative pose problems (essential and fundamental matrices estimation), we use the dataset “CVPR IMW 2020 PhotoTourism challenge”⁴, which has 12 scenes with around 100K pairs and 2 sequences with around 5K pairs (inlier rates vary between 30 and 60%, approximately). For the homography estimation problem, we use the EVD⁵ and HPatches [2], with 7 and 145 pairs of images, respectively (we used the validation set since the test set does not provide ground-truth). Matches and pre-computed weights were obtained with RootSIFT features and nearest-neighbor matching. Dataloaders were borrowed from the “RANSAC in 2020” tutorial webpage.

For both relative pose problems, results were obtained by taking, for each scene, the first 4K pairs and repeating each trial 5 times to ensure we have replicable accuracy and computational time readings for different runs. In the homography matrix results, we use all the available pairs and repeat each trial 10 times. For the experiments with different Markov order assumptions, varying number of fixed iterations, and varying inlier ratio, we use the `sacre_coeur` sequence entirely (around 5K pairs). We also use this sequence to tune our BANSAC parameters (stopping criterion and probability update model).

5.1 Different Markov assumption orders

We start the experiments by running an ablation study on the Markov assumption described in Sec. 4.2. We use the 1st, 2nd, and 3rd Markov assumption orders (the 2nd and 3rd are derived in the supplementary materials) and run BANSAC for a calibrated relative pose problem. Results

⁴<https://www.cs.ubc.ca/research/image-matching-challenge/current/> [August 18, 2023]

⁵<http://cmp.felk.cvut.cz/wbs/> [August 18, 2023]

³github.com/ducha-aiki/ransac-tutorial-2020-data [August 18, 2023]

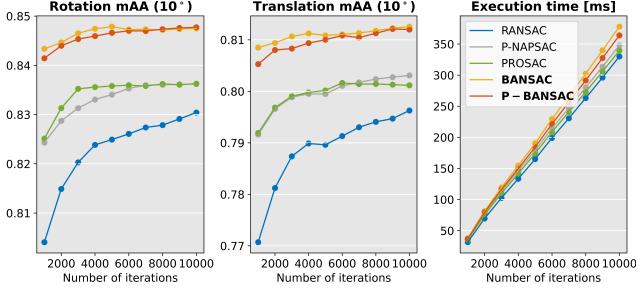


Figure 3. Results for a fixed number of iterations, i.e., without stopping criterion. We vary the number of iterations between 1000 and 10000.

are shown in Tab. 2.

We observe that there are minor differences between the different Markov assumption orders. Rotation and translation errors are lower on the 1st-order assumption, and execution time is lower on the 3rd-order assumption. This occurs because probabilities converge more rapidly for higher orders, activating the stopping criterion faster. Prioritizing accuracy, in the following experiments, both BANSAC and P-BANSAC use the 1st-order Markov assumption in the probability modeling.

5.2 Varying number of fixed iterations

In this experiment, we aim to evaluate the efficiency of the sampling process. We run experiments fixing the number of iterations for all the methods; no early stopping criterion is used. We vary the number of iterations between 1000 and 10000 on an uncalibrated relative pose problem and measure the rotation error, translation error, and execution time. The results obtained are shown in Fig. 3.

As expected, overall, we verify that with an increasing number of iterations, the error decreases for all the methods. We observe that BANSAC and P-BANSAC have the lowest rotation and translation errors. Concerning execution time, although both our methods require additional steps to update the probabilities every iteration, we notice that the results are marginally the same.

5.3 Varying inlier ratio

The inlier ratio has a strong impact on RANSAC-based methods performance. To evaluate how each method performs for different inlier ratios, we vary the confidence threshold for filtering matches from 0.82 to 0.92, which gives us inlier rates ranging between around 60% to 30%, respectively. Results for rotation and translation errors and execution time for an uncalibrated relative pose problem are shown in Fig. 4.

In contrast to the baselines, we observe that the decrease in the inlier ratio (by filtering fewer matches) increases the accuracy for BANSAC and P-BANSAC. Concerning exe-

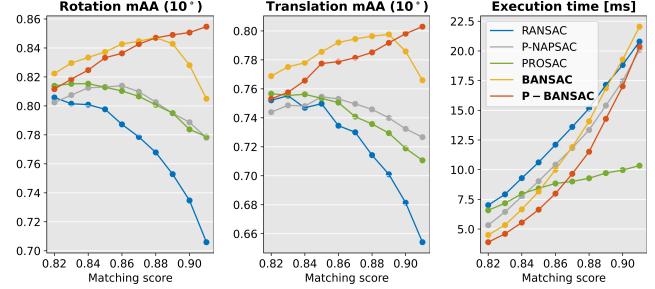


Figure 4. Results varying the quality of the input data. As the matching score increases from 0.82 to 0.92, the inlier ratio decreases from 60% to 30%.

cution time, it increases in all methods similarly, except in PROSAC where it grows less.

5.4 Results

Next, we present results for three computer vision problems with and without a local-optimization step.

Calibrated relative pose: Results for the calibrated relative pose (essential matrix estimation) are shown in Tab. 3. They show that BANSAC is the most accurate method, followed by P-BANSAC across all the scenes. In execution time, P-BANSAC is the fastest method, followed by BANSAC, indicating that the pre-computed scores help BANSAC on exiting the RANSAC loop earlier.

Uncalibrated relative pose: Results for the uncalibrated relative pose (fundamental matrix estimation) are shown in Tab. 3. Similarly to the previous results, we observe that BANSAC is consistently the most accurate method, and P-BANSAC is the second most accurate in most scenes. In runtime, P-BANSAC is the fastest and BANSAC the second fastest. Figure 5 shows the probabilities values after 10, 100, 1000, and 10000 iterations in an image pair from the *sacre-coeur* sequence, using BANSAC.

Homography estimation: Results for homography estimation are shown in Tab. 3. In this experiment, P-BANSAC is the best in accuracy, with BANSAC being second best. In runtime, RANSAC is the fastest. We note that BANSAC requires an additional loop over all data points per iteration for updating scores (see the discussion in Sec. 6) when compared with RANSAC. This extra computational effort is visible when BANSAC does not exit the loop sufficiently earlier than the baselines, as shown in Fig. 3. Figure 1 shows the initial probabilities of some randomly chosen matches (all started at 0.5) and the updated probabilities after 10, 100, and 1000 iterations in an image pair from the HPatches dataset using BANSAC, demonstrating the probability updates over iterations. The ground truth and the estimated homography are marked in green and red, respectively.

Table 3. Experimental results for the calibrated relative pose, uncalibrated relative pose, and homography estimation. We compare BANSAC and P-BANSAC with RANSAC [23], NAPSAC [48], P-NAPSAC [4], and PROSAC [16]. We show results with and without the local-optimization step of LO-RANSAC [18].

Metrics	Without Local Optimization					With Local Optimization						
	RANSAC	NAPSAC	P-NAPSAC	PROSAC	BANSAC	P-BANSAC	RANSAC	NAPSAC	P-NAPSAC	PROSAC	BANSAC	P-BANSAC
<i>Calibrated Relative Pose Estimation (essential matrix estimation)</i>												
Rotation mAA (5°) \uparrow	0.568	0.158	0.551	0.569	0.610	0.603	0.569	0.216	0.557	0.570	0.611	0.604
Rotation mAA (10°) \uparrow	0.645	0.226	0.641	0.653	0.680	0.675	0.645	0.292	0.645	0.655	0.680	0.675
Translation mAA (5°) \uparrow	0.422	0.0810	0.402	0.417	0.460	0.454	0.423	0.114	0.409	0.419	0.461	0.454
Translation mAA (10°) \uparrow	0.532	0.137	0.514	0.527	0.566	0.559	0.532	0.176	0.520	0.528	0.566	0.560
Avg. execution time [ms] \downarrow	25.5	40.1	20.9	21.5	15.6	15.2	27.6	42.9	26.6	22.6	18.0	17.4
<i>Uncalibrated Relative Pose Estimation (fundamental matrix estimation)</i>												
Rotation mAA (5°) \uparrow	0.467	0.206	0.460	0.464	0.500	0.478	0.514	0.499	0.517	0.511	0.526	0.501
Rotation mAA (10°) \uparrow	0.559	0.308	0.557	0.560	0.589	0.571	0.595	0.572	0.600	0.595	0.610	0.589
Translation mAA (5°) \uparrow	0.267	0.0780	0.260	0.264	0.292	0.274	0.308	0.300	0.309	0.307	0.317	0.294
Translation mAA (10°) \uparrow	0.353	0.129	0.345	0.349	0.380	0.360	0.394	0.381	0.396	0.392	0.405	0.381
Avg. execution time [ms] \downarrow	14.4	26.3	12.2	12.5	9.80	7.88	15.4	18.4	13.6	13.5	11.4	8.87
<i>Homography Estimation</i>												
Homography mAA (5 px) \uparrow	0.422	0.152	0.158	0.210	0.443	0.446	0.513	0.498	0.488	0.333	0.542	0.517
Homography mAA (10 px) \uparrow	0.552	0.208	0.235	0.291	0.569	0.573	0.647	0.617	0.617	0.426	0.672	0.650
Avg. execution time [ms] \downarrow	2.30	2.78	2.96	3.09	4.04	3.07	3.24	5.94	6.43	2.89	4.09	4.33

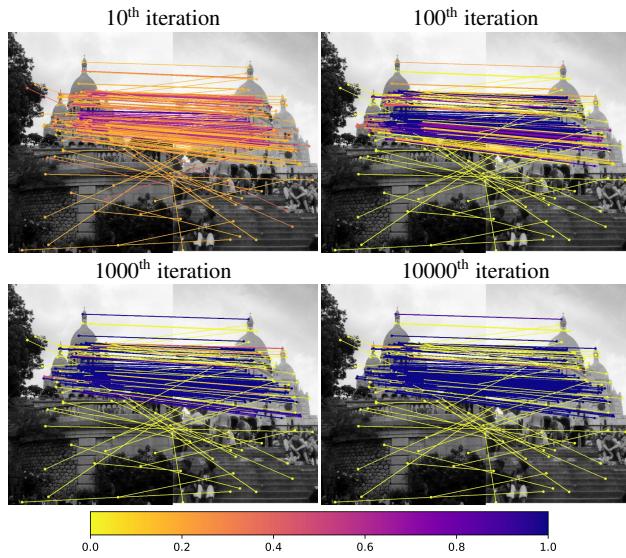


Figure 5. Data inlier probabilities over iterations for a fundamental matrix estimation problem using BANSAC (color code at the right). Image pair is from the PhotoTourism sacre_coeur scene. For visualization purposes, we show only 250 randomly chosen matches.

Local optimization: We add the local-optimization (LO) step in [18] to all methods and repeat the previous experiments. Results are present in Tab. 3. We observe that LO improves the accuracy for every method, with an increase in execution time. We also note that the improvement in accuracy is significant for some of the baselines, e.g., NAPSAC and P-NAPSAC. Overall, BANSAC and P-BANSAC continue to outperform the baselines by some margin.

6 Conclusion

This paper proposes BANSAC, a new sampling strategy for RANSAC using dynamic Bayesian networks. The method performs weighted sampling using probabilities for scoring data points. These probabilities are adaptively updated every iteration based on the successive inlier/outlier classifications. Additionally, we propose a stopping criterion using the estimated probabilities.

We present results on challenging real-world datasets showing that the proposed algorithm can learn the data inlier probability and that these probabilities can guide the sampling efficiently; the updates to RANSAC bring improvements in accuracy and execution time.

We note that BANSAC updates the scoring weights depending on the quality of the hypothesis, which is obtained from inlier counting. This means that we need an extra loop cycle every iteration for updating the scores. Although we do not observe a significant increase in computational cost compared to the best baselines, there is room for improvement. In future work, we plan to include a more efficient hypothesis prediction for incorporating the probability update in the inlier counting loop.

Acknowledgments

Valter Piedade was supported by the National Centre for Research and Development under the Smart Growth Operational Programme as part of project POIR.01.01.01-00-0102/20 and by the LARSyS–FCT Project UIDB/50009/2020. We thank all the reviewers and ACs for their valuable feedback.

References

- [1] OpenCV: Open source computer vision library. <https://github.com/opencv/opencv>. 2
- [2] Vassileios Balntas, Karel Lenc, Andrea Vedaldi, and Krystian Mikolajczyk. Hpatches: A benchmark and evaluation of handcrafted and learned local descriptors. In *IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR)*, 2017. 1, 6
- [3] Daniel Barath, Luca Cavalli, and Marc Pollefeys. Learning to find good models in ransac. In *IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR)*, pages 15744–15753, 2022. 3
- [4] Daniel Barath, Maksym Ivashevkin, and Jiri Matas. Progressive napsac: sampling from gradually growing neighborhoods. *arXiv preprint arXiv:1906.02295*, 2019. 1, 2, 6, 8
- [5] Daniel Barath and Jiri Matas. Graph-cut ransac. In *IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, pages 6733–6741, 2018. 3
- [6] Daniel Barath, Jiri Matas, and Jana Noskova. Magsac: marginalizing sample consensus. In *IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR)*, pages 10197–10205, 2019. 3
- [7] Daniel Barath, Jana Noskova, Maksym Ivashevkin, and Jiri Matas. Magsac++, a fast, reliable and accurate robust estimator. In *IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR)*, pages 1304–1312, 2020. 3
- [8] Tom Botterill, Steven Mills, and Richard D Green. New conditional sampling strategies for speeded-up ransac. In *British Machine Vision Conference (BMVC)*, pages 1–11, 2009. 2, 6
- [9] Eric Brachmann, Alexander Krull, Sebastian Nowozin, Jamie Shotton, Frank Michel, Stefan Gumhold, and Carsten Rother. Dsac-differentiable ransac for camera localization. In *IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR)*, pages 6684–6692, 2017. 2
- [10] Eric Brachmann and Carsten Rother. Learning less is more: 6d camera localization via 3d surface regression. In *IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, pages 4654–4662, 2018. 1
- [11] Eric Brachmann and Carsten Rother. Neural-guided ransac: Learning where to sample model hypotheses. In *IEEE/CVF Int'l Conf. Computer Vision (ICCV)*, pages 4322–4331, 2019. 1, 2
- [12] Dingding Cai, Janne Heikkila, and Esa Rahtu. Ove6d: Object viewpoint encoding for depth-based 6d object pose estimation. In *IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR)*, pages 6803–6813, 2022. 1
- [13] David P Capel. An effective bail-out test for ransac consensus scoring. In *British Machine Vision Conference (BMVC)*, volume 1, page 2, 2005. 3
- [14] Luca Cavalli, Viktor Larsson, Martin Ralf Oswald, Torsten Sattler, and Marc Pollefeys. Handcrafted outlier detection revisited. In *European Conf. Computer Vision (ECCV)*, pages 770–787, 2020. 2
- [15] Luca Cavalli, Marc Pollefeys, and Daniel Barath. Nefsac: neurally filtered minimal samples. In *European Conf. Computer Vision (ECCV)*, pages 351–366, 2022. 2
- [16] Ondrej Chum and Jiri Matas. Matching with prosac—progressive sample consensus. In *IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, volume 1, pages 220–226, 2005. 1, 2, 3, 5, 6, 8
- [17] Ondrej Chum and Jiri Matas. Optimal randomized ransac. *IEEE Trans. Pattern Analysis and Machine Intelligence (TPAMI)*, 30(8):1472–1482, 2008. 3
- [18] Ondrej Chum, Jiri Matas, and Josef Kittler. Locally optimized ransac. In *Joint Pattern Recognition Symposium*, pages 236–243, 2003. 3, 8
- [19] Ondrej Chum, Tomas Werner, and Jiri Matas. Epipolar geometry estimation via ransac benefits from the oriented epipolar constraint. In *IEEE Int'l Conf. Pattern Recognition (ICPR)*, volume 1, pages 112–115, 2004. 3
- [20] Ondrej Chum, Tomas Werner, and Jiri Matas. Two-view geometry estimation unaffected by a dominant plane. In *IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, volume 1, pages 772–779, 2005. 3
- [21] Hongyi Fan, Joe Kileel, and Benjamin Kimia. On the instability of relative pose estimation and ransac's role. In *IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR)*, pages 8935–8943, 2022. 3
- [22] Maxime Ferrera, Alexandre Eudes, Julien Moras, Martial Sanfourche, and Guy Le Besnerais. Ov2slam: A fully online and versatile visual slam for real-time applications. *IEEE Robotis and Automation Letters (RA-L)*, 6(2):1399–1406, 2021. 1
- [23] Martin A Fischler and Robert C Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395, 1981. 1, 2, 3, 5, 6, 8
- [24] Victor Fragoso, Pradeep Sen, Sergio Rodriguez, and Matthew Turk. Evsac: accelerating hypotheses generation by modeling matching scores with extreme value theory. In *IEEE Int'l Conf. Computer Vision (ICCV)*, pages 2472–2479, 2013. 2
- [25] Maksym Ivashevkin, Daniel Barath, and Jiří Matas. Vsac: Efficient and accurate estimator for h and f. In *IEEE/CVF Int'l Conf. Computer Vision (ICCV)*, pages 15243–15252, 2021. 3
- [26] Michal Jancosek and Tomas Pajdla. Multi-view reconstruction preserving weakly-supported surfaces. In *IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, pages 3121–3128, 2011. 1
- [27] Geun Jo, Kee-Sung Lee, Devy Chandra, Chol-Hee Jang, and Myung-Hyun Ga. Ransac versus cs-ransac. In *AAAI Conference on Artificial Intelligence*, volume 29, 2015. 2
- [28] Viktor Larsson, Zuzana Kukelova, and Yingqiang Zheng. Making minimal solvers for absolute pose estimation compact and robust. In *IEEE Int'l Conf. Computer Vision (ICCV)*, pages 2316–2324, 2017. 1
- [29] Philipp Lindenberger, Paul-Edouard Sarlin, Viktor Larsson, and Marc Pollefeys. Pixel-perfect structure-from-motion with featuremetric refinement. In *IEEE/CVF Int'l Conf. Computer Vision (ICCV)*, pages 5987–5997, 2021. 1
- [30] Jiri Matas and Ondrej Chum. Randomized ransac with td, d test. *Image and Vision Computing (IVC)*, 22(10):837–842, 2004. 3

- [31] Jiri Matas and Ondrej Chum. Randomized ransac with sequential probability ratio test. In *IEEE Int'l Conf. Computer Vision (ICCV)*, volume 2, pages 1727–1732, 2005. 3, 5, 6
- [32] Andre Mateus, Srikumar Ramalingam, and Pedro Miraldo. Minimal solvers for 3d scan alignment with pairs of intersecting lines. In *IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR)*, pages 7234–7244, 2020. 1
- [33] Paul McIlroy, Edward Rosten, Simon Taylor, and Tom Drummond. Deterministic sample consensus with multiple match hypotheses. In *British Machine Vision Conference (BMVC)*, pages 1–11, 2010. 2
- [34] Antoine Melier, Marion Decrouez, and James L Crowley. Betasac: A new conditional sampling for ransac. In *British Machine Vision Conference (BMVC)*, 2010. 2
- [35] Pedro Miraldo, Tiago Dias, and Srikumar Ramalingam. A minimal closed-form solution for multi-perspective pose estimation using points and lines. In *European Conf. Computer Vision (ECCV)*, pages 474–490, 2018. 1
- [36] Raul Mur-Artal and Juan D. Tardos. Orb-slam2: An open-source slam system for monocular, stereo, and rgbd cameras. *IEEE Trans. Robotics (T-RO)*, 33(5):1255–1262, 2017. 1
- [37] Kai Ni, Hailin Jin, and Frank Dellaert. Groupsac: Efficient consensus in the presence of groupings. In *IEEE Int'l Conf. Computer Vision (ICCV)*, pages 2193–2200, 2009. 1, 2
- [38] David Nister. Preemptive ransac for live structure and motion estimation. *Machine Vision and Applications*, 16(5):321–329, 2005. 3
- [39] Linfei Pan, Marc Pollefeys, and Viktor Larsson. Camera pose estimation using implicit distortion models. In *IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR)*, pages 12819–12828, 2022. 1
- [40] Rahul Raguram, Ondrej Chum, Marc Pollefeys, Jiri Matas, and Jan-Michael Frahm. Usac: A universal framework for random sample consensus. *IEEE Trans. Pattern Analysis and Machine Intelligence (T-PAMI)*, 35(8):2022–2038, 2012. 2
- [41] Stuart J. Russell and Peter Norvig. *Artificial Intelligence: a modern approach*. Pearson, 3 edition, 2009. 4, 5
- [42] Paul-Edouard Sarlin, Ajaykumar Unagar, Mans Larsson, Hugo Germain, Carl Toft, Viktor Larsson, Marc Pollefeys, Vincent Lepetit, Lars Hammarstrand, Fredrik Kahl, et al. Back to the feature: Learning robust camera localization from pixels to pose. In *IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR)*, pages 3247–3257, 2021. 1
- [43] Torsten Sattler, Bastian Leibe, and Leif Kobbelt. Fast image-based localization using direct 2d-to-3d matching. In *IEEE Int'l Conf. Computer Vision (ICCV)*, pages 667–674, 2011. 1
- [44] Johannes L Schonberger and Jan-Michael Frahm. Structure-from-motion revisited. In *IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, pages 4104–4113, 2016. 1
- [45] Johannes L Schonberger, Enliang Zheng, Jan-Michael Frahm, and Marc Pollefeys. Pixelwise view selection for unstructured multi-view stereo. In *European Conf. Computer Vision (ECCV)*, pages 501–518, 2016. 1
- [46] Charles V. Stewart. Minpran: A new robust estimator for computer vision. *IEEE Trans. Pattern Analysis and Machine Intelligence (T-PAMI)*, 17(10):925–938, 1995. 2
- [47] Ben J Tordoff and David William Murray. Guided-mlesac: Faster image transform estimation by using matching priors. *IEEE Trans. Pattern Analysis and Machine Intelligence (T-PAMI)*, 27(10):1523–1535, 2005. 1, 2
- [48] Philip Hilaire Torr, Slawomir J Nasuto, and John Mark Bishop. Napsac: High noise, high dimensional robust estimation-it's in the bag. In *British Machine Vision Conference (BMVC)*, volume 2, page 3, 2002. 1, 2, 6, 8
- [49] Philip HS Torr and Andrew Zisserman. Mlesac: A new robust estimator with application to estimating image geometry. *Computer Vision and Image Understanding (CVIU)*, 78(1):138–156, 2000. 2
- [50] Alexander Vakhitov, Jan Funke, and Francesc Moreno-Noguer. Accurate and linear time pose estimation from points and lines. In *European Conf. Computer Vision (ECCV)*, pages 583–599, 2016. 1
- [51] Brian Williams, Georg Klein, and Ian Reid. Automatic re-localization and loop closing for real-time monocular slam. *IEEE Trans. Pattern Analysis and Machine Intelligence (T-PAMI)*, 33(9):1699–1712, 2011. 1