

## Inverting RANSAC: Global Model Detection via Inlier Rate Estimation

Roee Litman\*, Simon Korman\*, Alex Bronstein and Shai Avidan  
Department of Electrical Engineering, Tel-Aviv university

The problem of image correspondence is a fundamental problem in computer vision, as it arises as a primitive in many tasks such as image retrieval, 3D reconstruction and panorama stitching. While some works approach this problem using direct methods [2, 4], the vast majority of recent methods use large sets of matching (pairs of) points as their entry point, discarding the content of the images. This is largely due to the tremendous improvement in algorithms for detecting stable image feature points and representing them by descriptors that are designed for the task of matching [3, 5, 6].

The desired outcome of such a point matching process is that a large portion of the matches is accurate, while only a few of them (preferably none) can have arbitrarily bad errors. These two groups of matches are called *inliers* and *outliers*, respectively. The final step of image matching is therefore to robustly detect the “true” transformation underlying the inliers while ignoring the outliers. In practice, this is most commonly formulated as the *consensus set maximization* (CSM) problem, where the goal is to find a maximal set of matches that agree on a model, up to some tolerance.

The standard approach to CSM is the RANSAC framework [1], in which the space of parameters is explored by repeatedly selecting random subsets of matches for which a model hypothesis is fitted and then verified. A recent comprehensive survey and evaluation of RANSAC techniques proposes USAC [7], a uniform pipeline that combines several of the known RANSAC extensions in addition to many practical and computational considerations.

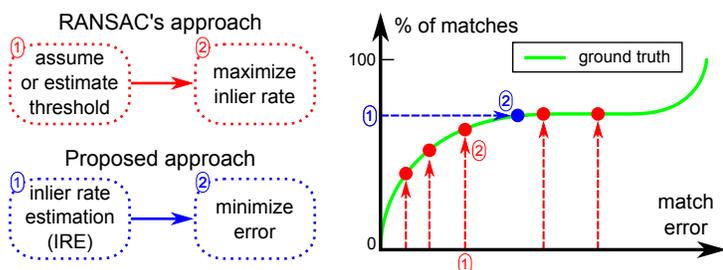


Figure 1: **Our approach is “orthogonal” to RANSAC**, which assumes a fixed error-threshold for inliers and then searches for a model that *maximizes the inlier rate*. Our method works in an opposite order: the inlier rate of matches is first estimated from the data and then, a model that *minimizes the error* of such a portion of inliers is searched for.

This work presents a different approach, as is illustrated in Figure 1. The green curve is a cumulative error distribution of matches between a pair of images, under a ground-truth transformation. In RANSAC (and other CSM approaches), a fixed error threshold is chosen and a model with a maximal number of inliers within the threshold is searched for (depicted by vertical red lines, for different thresholds). Our approach, on the other hand, first estimates the true inlier rate of the matches (about 70% in this example) and then searches for a model with lowest possible match errors over the detected portion of matches (depicted by the blue horizontal arrow).

The portion of inliers and the noise level of inlier matches are generally unknown, and any a-priori choice of error threshold is rather arbitrary. Our inlier rate estimation (IRE) method makes a principled prediction based on minimizing an indicative quantity, denoted  $v(p)$ , over any possible inlier rate  $p$ . The measure  $v$  ‘counts’ the number of transformations (or portion of transformation space) that have a  $p$ -tile error ‘similar’ to the best one possible. It turns out that  $v$  has a very particular behavior around the true inlier rate, where it attains a surprisingly clear minimum. This phenomenon can be seen in the middle plot of Figure 2 for a real case of 2D homography estimation. The resulting inlier-rate prediction  $\hat{p}$  is a critical input to our BnB search algorithm (see on right) which approximates the optimal transformation w.r.t. the inlier rate  $\hat{p}$ .

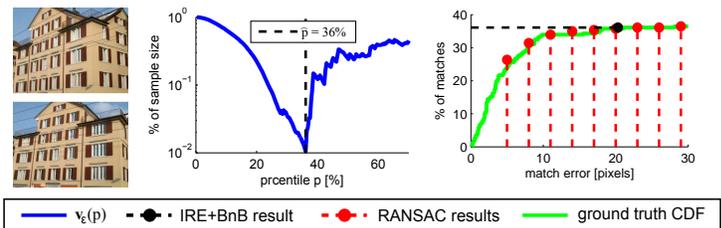


Figure 2: **2D homography estimation**. **Left**: an image pair from [7]; **Middle**: our IRE measure (blue curve) and our inlier-rate prediction (black dashed line) at the minimum (found at 36% of matches); **Right**: The result of our Algorithm (black circle) and the result of multiple USAC [7] runs for different thresholds (red circles), shown against the CDF (green curve) of match-errors w.r.t. the ground-truth transformation.

**Contributions** This paper has three main contributions. First, a scheme for efficiently sampling the space of homographies. Second, an algorithm (BnB) for finding the best transformation for a set of matches, given the rate of inliers, with global guarantees. This algorithm has low practical applicability without our third, main, contribution - an algorithm (IRE) for estimating the rate of inliers in a given set of matches, without explicitly detecting them. In addition, we present a rigorous analysis of the IRE algorithm and validate our analysis in several settings. We also show that our complete framework, termed GMD, improves on the state-of-the-art on several challenging data-sets. Refer to Figure 3 for example image pairs from the Mikolajczyk [6] data-set.

sequence: ‘Graffiti’  
image pair: 1-5  
inlier rate: 12.6%  
inlier noise:  $8.5 \pm 14$   
Sampson error: 6.55



sequence: ‘Graffiti-5’  
image pair: 1-6  
inlier rate: 9.3%  
inlier noise:  $6.3 \pm 11$   
Sampson error: 6.63



— ground truth map — GMD map ● outliers ● inliers

Figure 3: **Challenging image pairs from the Mikolajczyk [6] dataset**. GMD finds correct homographies for these instances of image pairs with matching SIFTs, which are extreme due to the combination of low inlier rate and a high inlier noise. The state-of-the-art USAC [7], fails to find a model, even when tested on multiple inlier match-error thresholds.

- [1] M. A. Fischler and R. C. Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 1981.
- [2] Y. HaCohen, E. Shechtman, D. B. Goldman, and D. Lischinski. Non-rigid dense correspondence with applications for image enhancement. *SIGGRAPH*, 2011.
- [3] D. Lowe. Distinctive image features from scale-invariant keypoints. *IJCV*, 2004.
- [4] B. D. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. In *IJCAI*, 1981.
- [5] K. Mikolajczyk and C. Schmid. Scale & affine invariant interest point detectors. *IJCV*, 2004.
- [6] K. Mikolajczyk and C. Schmid. A performance evaluation of local descriptors. *PAMI*, 2005.
- [7] R. Raguram, O. Chum, M. Pollefeys, J. Matas, and J. Frahm. Usac: A universal framework for random sample consensus. *PAMI*, 2013.