Practical Robust Two-View Translation Estimation

Johan Fredriksson, Viktor Larsson, Carl Olsson
Centre for Mathematical Sciences
Lund University

Outliers pose a problem in all real structure from motion systems. Due to the use of automatic matching methods one has to expect that a (sometimes very large) portion of the detected correspondences can be incorrect. In this paper we propose a method that estimates the relative translation between two cameras and simultaneously maximizes the number of inlier correspondences.

Traditionally, outlier removal tasks have been addressed using RANSAC approaches. However, these are random in nature and offer no guarantees of finding a good solution. If the amount of mismatches is large, the approach becomes costly because of the need to evaluate a large number of random samples. In contrast, our approach is based on the branch and bound methodology which guarantees that an optimal solution will be found. While most optimal methods trade speed for optimality, the proposed algorithm has competitive running times on problem sizes well beyond what is common in practice. Experiments on both real and synthetic data show that the method outperforms state-of-the-art alternatives, including RANSAC, in terms of solution quality. In addition, the approach is shown to be faster than RANSAC in settings with a large amount of outliers.

The geometry of the problem has been investigated in prior works such as [1, 2]. If the motion is restricted to translation the epipolar constraint reduces to the image points and the translation being coplanar. If we allow for some reprojection error ε the translation must instead lie inside one of two wedges formed by taking the tangent great circles to the ε-cones around the image points. See Figure 1.

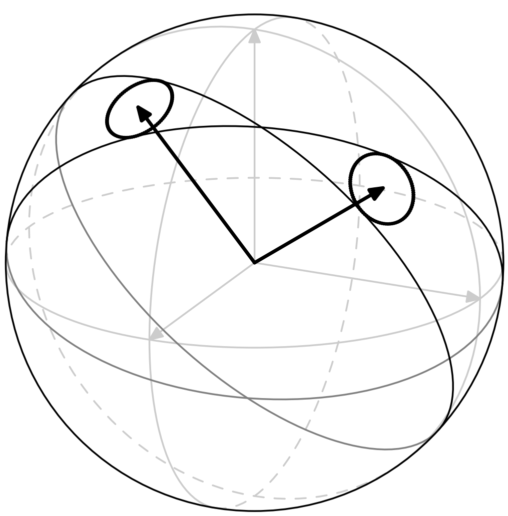

Figure 1: Epipolar geometry under translation.

In [1], Fredriksson et al. gave an optimal algorithm for the two-view translation problem in the presence of outliers. For a given error threshold the algorithm finds a translation with the maximal amount of inlier correspondences, with a theoretical time complexity of $\mathcal{O}(n^2 \log n)$. It was demonstrated that in settings with large amounts of outliers the approach not only produces better results but is also faster than RANSAC.

In this paper we present a branch and bound approach tailored for the two-view translation problem that outperforms Fredriksson et al. [1]. The key to designing an effective formulation is to find strong bounding functions that can be evaluated efficiently. We achieve this by a division of the parameter space $\mathcal{S}^2$ using spherical triangles. Checking feasibility of a correspondence within a parameter triangle reduces to computing a few intersections between great circles, which can be done extremely efficiently using a simple cross-product. See Figure 2.

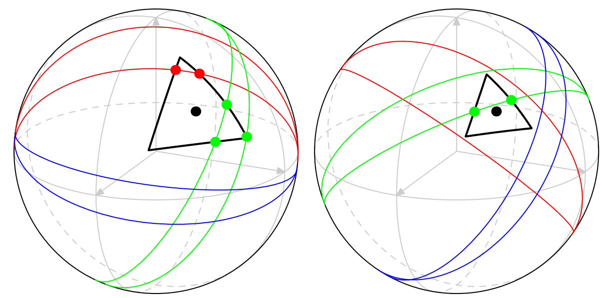

Figure 2: The bounds for a triangle is found by counting the number of wedges which either intersect it or completely contains it.

In experiments we show that the method outperforms RANSAC in terms of quality of the solution and can in some cases be faster, in particular if the portion of outliers is high. In addition, the approach outperforms [1] with orders of magnitude while providing the same optimality guarantees. Figure 3 shows a comparison with [1] on synthetic data. Since both methods solve the problem optimally we only show the running time.

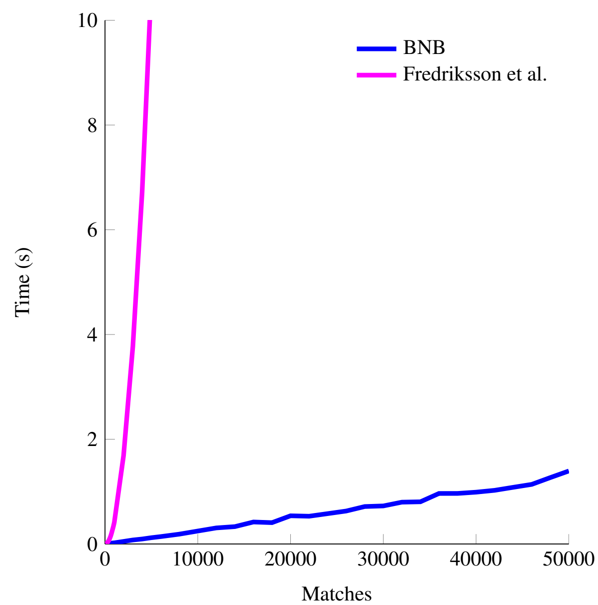

Figure 3: Comparisons of the running times on synthetic data for Fredriksson et al.[1] (purple) and the proposed method (blue). The number of inliers are kept at 10% and the number of point pairs vary between 50 and 50000.

[1] Johan Fredriksson, Olof Enqvist, and Fredrik Kahl. Fast and reliable two-view translation estimation. In *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014.

[2] Richard I Hartley and Fredrik Kahl. Global optimization through rotation space search. *International Journal of Computer Vision*, 82(1): 64–79, 2009.

This is an extended abstract. The full paper is available at the Computer Vision Foundation webpage.