Hybrid and Non-minimal Planar Motion Estimation from Point Correspondences (Supplementary)

Juan C. Dibene and Enrique Dunn

Stevens Institute of Technology

1 Relevance of our proposed solvers

We target calibrated egomotion estimation for visual odometry modules in mobile platforms, focusing on the under-explored prior-free planar motion scenario (e.g. robot navigation, autonomous driving). Introducing hybrid solvers enables previously unavailable trade-offs between solver complexity, runtime, and solution quality (critical for online deployment). Given the ubiquity of stereo and RGBD cameras, we analyze how varying degrees of 3D data integration into the estimation process streamlines the resulting geometric formulation and affects estimation quality due to now introducing 3D noise.
Egomotion estimation inputs frame-to-frame correspondences w/o an available 3D map. Hence, 3D-3D matches are obtained by feature tracking on sequences of triangulated stereo pairs (e.g., KITTI [2]) or RGBD image pairs (e.g., TUM [5]). Alternatively, 2D-3D matches may come from RGBD feature matching combined with data decimation or arise naturally in mobile stereo capture with partial FOV overlap. For situations where 3D data is unavailable, considering pure 2D-2D matches renders our approach flexible. Our non-minimal solvers elucidate how integrating additional 2D-2D matches can yield less complex and faster planar solvers, with unique solutions in all cases.

2 Performance of our solvers vs SOTA

We illustrate the overall performance (error vs runtime) of all solvers in Figure 1. The data for Figure 1 is sourced directly from Table 3, Table 4, and Table 5 of the original submission (i.e., it is not from new experiments). On both KITTI [2] and TUM [5], our 5-pt and 7-pt solvers achieve comparable results to 4pst0 [3] but much faster, while our 6-pt solver is both more accurate and faster than 4pst0 [3]. On KITTI, our solvers (except hpc1) are dominated by the 6-DOF Lambda Twist solver [4]. We attribute this to higher non-planarities present in the dataset, which was also observed in [1, 3]. However, on the more planar TUM sequences, our hpc2, nm5, nm6, and nm7 solvers cluster near the bottom left (low error and low runtime) and dominate most of the solvers, including full DOF solvers, demonstrating that our solvers can outperform, both in accuracy and runtime, SOTA solvers in situations where the planar motion assumption is better supported. Then, solver selection is contingent on the availability

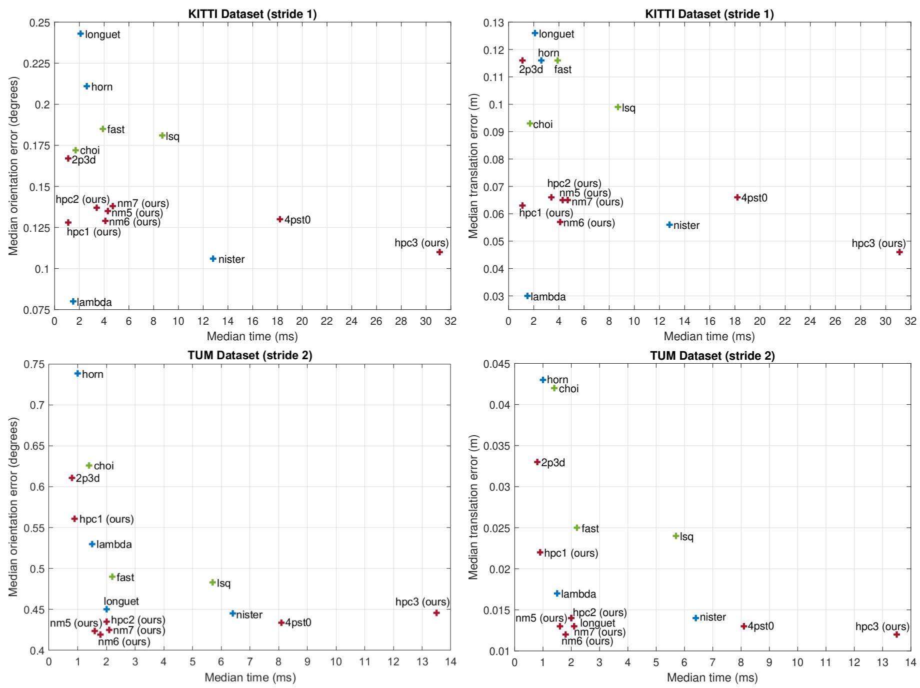

Fig. 1: Median errors vs RANSAC time. Data sourced from Tables 3, 4, and 5 of the original submission. +: no plane prior, +: plane prior, +: full DOF.

of 3D data and user preference of accuracy vs speed. Acknowledging potential implementation-specific optimizations, our MATLAB-based implementation timings are intended for relative comparison of solver execution times. RANSAC-based estimation on empirical data (subject to measurement noise and motion model non-compliance) clearly exemplifies our solvers' empirical performance trade-offs.

3 Variations of Solver Formulations

In the original submission, we briefly discuss the engineering choices made to implement our solvers. Here, we elaborate on the reasons of these choices and their experimental justification.

Zero Trace Constraint Parameterization. As mentioned in the original submission, there are two possible ways to embed the zero trace constraint for planar motion into existing essential matrix (E) estimators: 1) Stack the coefficients of the zero trace equation into the matrix of epipolar constraints (9 parameters), 2) Replace one the diagonal elements of E with the negated sum of the other two diagonal elements, e.g. $e_{11} = -(e_{22} + e_{33})$ (8 parameters). To implement our solvers, we chose option 2 based on the results of an experiment on synthetic data with 2D pixel noise, performed as described in the original submission. The results are shown in Figure 2. The 7-point solver is marginally affected by either parameterization, while the 6-point solver has similar performance, with 8 parameters performing slightly better. However, the 8 parameters parameterization is significantly better for the 5-point point solver. Therefore, we use 8 parameters for all of our three non-minimal 2D-2D solvers.

Enforcement of Eigenvalues. As shown in the original submission we have $\mathrm{tr}(E) = 0$ for planar motion, and it is known that for all essential matrices $\det(E) = 0$. This implies that the eigenvalues of E for planar motion are of the form $[\lambda, -\lambda, 0]$. However, for imperfect input observations (e.g., corrupted by noise), the estimated E (by non-minimal solvers) does not necessarily conform to this structure. We considered enforcing this structure via eigendecomposition. Let $E = Q\,\mathrm{diag}(\lambda_1, \lambda_2, \lambda_3)Q^{-1}$, $\|\lambda_1\| > \|\lambda_2\| > \|\lambda_3\|$, be the eigendecomposition of E. Then, a conformant E can be reconstructed as $E = Q\,\mathrm{diag}(\lambda, -\lambda, 0)Q^{-1}$. We use the same setup as for the noisy synthetic experiments in the original submission, considering our three non-minimal 5-point, 6-point, and 7-point solvers, with and without eigenvalue enforcement. The results are shown in Figure 3. We observe that the solvers enforcing the eigenvalue form (denoted with an E) are consistently worse than those without it. Therefore, we do not enforce the eigenvalue form in our solvers. Consequently, the estimated E may not have zero trace (i.e., E is not exactly planar) in non-ideal conditions. Nevertheless, this is not a problem as shown by our experimental results in the original submission. Note that the singular value structure $[\sigma, \sigma, 0]$ of proper essential matrices is automatically enforced when decomposing E into R and t.

4 Planar Scene Structure

We analyze the effect of coplanar input observations on our three non-minimal 2D-2D solvers by means of a noiseless synthetic data experiment. We consider seven 3D points on a plane at $Z = 10$. Then, we move zero, one, two, and three of these points a short distance outside of this plane, so the input observations are comprised of seven, six, five and four coplanar points respectively. We test for the 10,000 random planar poses in the synthetic dataset and the results are shown in Table 1. Our three non-minimal 2D-2D solvers require that at least one of the input points is not coplanar with the others.

References

1. Dibene, J.C., Min, Z., Dunn, E.: General planar motion from a pair of 3d correspondences. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). pp. 8060–8070 (October 2023) 1
2. Geiger, A., Lenz, P., Urtasun, R.: Are we ready for autonomous driving? the kitti vision benchmark suite. In: Conference on Computer Vision and Pattern Recognition (CVPR) (2012) 1

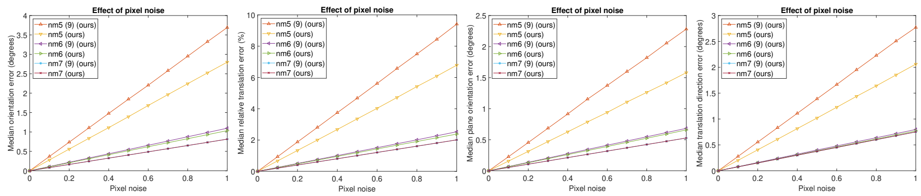

Fig. 2: Effect of 2D pixel noise considering 9 vs. 8 E parameters for our three non-minimal 2D-2D solvers.

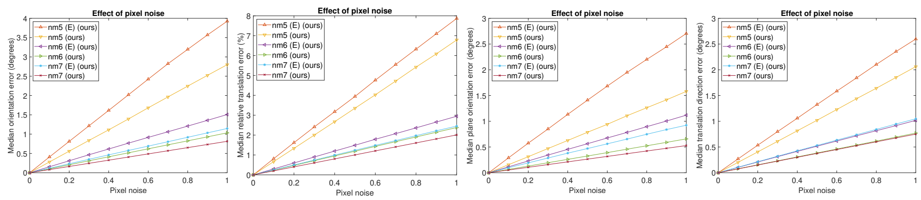

Fig. 3: Effect of 2D pixel noise for our three non-minimal 2D-2D solvers with eigenvalue enforcement (denoted by E) and without.

3. Li, B., Martyushev, E., Lee, G.H.: Relative pose estimation of calibrated cameras with known se(3) invariants. In: Computer Vision – ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part IX. p. 215–231. Springer-Verlag, Berlin, Heidelberg (2020) 1
4. Persson, M., Nordberg, K.: Lambda twist: An accurate fast robust perspective three point (p3p) solver. In: Proceedings of the European Conference on Computer Vision (ECCV) (September 2018) 1
5. Sturm, J., Engelhard, N., Endres, F., Burgard, W., Cremers, D.: A benchmark for the evaluation of rgb-d slam systems. In: Proc. of the International Conference on Intelligent Robot Systems (IROS) (Oct 2012) 1

Table 1: Mean orientation error (degrees) and relative translation error (%) for coplanar input observations to our non-minimal solvers. Failure cases highlighted in red.

	4 coplanar		5 coplanar		6 coplanar		7 coplanar	
	deg	%	deg	%	deg	%	deg	%
nm5 (ours)	4.31e-07	1.92e-07	16.816	53.32	-	-	-	-
nm6 (ours)	3.74e-07	2.39e-10	3.89e-07	1.04e-09	15.06	47.61	-	-
nm7 (ours)	3.89e-07	4.86e-12	3.88e-07	5.44e-12	4.28e-07	7.31e-11	16.10	67.67