NeFSAC: Neurally Filtered Minimal Samples

Luca Cavalli[1], Marc Pollefeys[1,2], and Daniel Barath[2]

[1] Department of Computer Science, ETH Zurich, Zurich, Switzerland
[2] Microsoft Mixed Reality and AI Zurich Lab

1 Implementation details

All MLPs in our network use leaky ReLU activation with slope 0.01, except for the final prediction branch which uses the sigmoid activation function. Labels l_i for pose and epipolar errors e_i are assigned the value zero if $e_i > e_{\max}$, the value of one if $e_i < e_{\min}$ and are linearly interpolated for error values in $[e_{\min}, e_{\max}]$. We set $e_{\max} = 5pix$ and $e_{\min} = 2pix$ for epipolar error using the Sampson distance and $e_{\max} = 30°$ and $e_{\min} = 5°$ for pose errors using the maximum between translation and rotation errors. When the minimal solver finds multiple solutions, we score the sample according to the best one, and when it fails to find any solution we consider it equivalent to the maximum possible error. The Sampson error assigned to a minimal sample is the maximum Sampson error of all of its correspondences with respect to the ground truth epipolar geometry. All correspondences both for training and for the experimental validation are obtained by matching SIFT [5] with ratio-test filtering of 0.8.

The error of the expert branch for autonomous driving takes inspiration from several existing works on planar constrained motion [3,4,7]. We assume that the correct motion is a rotation purely around the vertical axis and a translation without any vertical component, and penalize the maximum rotation or translation component that deviates from such assumption. Finally errors are mapped to classification labels with the same thresholds as with the previously defined pose errors. We further define a tentative error for an expert branch to partially capture the weaker regularities in a scenario of PhotoTourism [8] image collections. We consider a motion plausible when the relative rotation is either almost horizontal or almost vertical, and penalize the deviation from such model.

Since our technique proposed in Section 3.2 of the main paper does not suffice to produce a balanced dataset, we further address its residual imbalance with class weighting. For each branch, we keep a running count of positive and negative samples, and weigh each sample according to the inverse of the frequency of its class [9].

2 Experiments with MAGSAC++

In this section, we provide additional experiments of integrating NeFSAC on a substantially different RANSAC variant than USAC [6]. We combined NeFSAC with MAGSAC++ [2,1] and ran essential matrix estimation on scene British

Museum from the PhotoTourism dataset, KITTI with the frame difference of 4, and the Malaga dataset. The AUC@10 scores of the maximum rotation and translation errors, the average number of models tested, and the run-time in *ms* are shown in Table 1.

	PhotoTourism		**KITTI**		**Malaga**	
NeFSAC	w/o	w/	w/o	w/	w/o	w/
AUC@10 ↑	0.82	**0.83**	0.80	**0.82**	0.82	**0.86**
# models ↓	409	**153**	838	**204**	3888	**435**
time (ms) ↓	59	**44**	415	**127**	1456	**122**

Table 1: Augmenting MAGSAC++ with NeFSAC.

NeFSAC increases the accuracy, decreases the number of tested models and, thus, the run-time on all datasets

3 Testing NeFSAC with an inlier oracle

We argue that the contribution of NeFSAC is orthogonal to outlier filters. In this section, we use NeFSAC in combination to an idea inlier probability predictor on KITTI, which orders the available pool of correspondences according to their residuals with respect to the ground truth essential matrix. Due to using the PROSAC sampler, the probability of finding an all-inlier sample early is extremely high. We report results of USAC + Oracle, with or without NeFSAC, in Table 2. NeFSAC is still able to improve accuracy even in this ideal case, showing that learning to avoid degeneracies can still contribute on top of an ideal inlier selection.

	USAC + Oracle	USAC + Oracle + NeFSAC
AUC@10 ↑	0.93	**0.95**
# models ↓	65	**47**
time (ms) ↓	9	**10**

Table 2: NeFSAC influence with Oracle inlier probability predictor.

References

1. Barath, D., Noskova, J., Ivashechkin, M., Matas, J.: MAGSAC++, a fast, reliable and accurate robust estimator. In: CVPR (2020)
2. Barath, D., Noskova, J., Matas, J.: MAGSAC: marginalizing sample consensus. In: CVPR (2019), https://github.com/danini/magsac
3. Choi, S., Kim, J.: Fast and reliable minimal relative pose estimation under planar motion. Image Vis. Comput. **69**, 103–112 (2018)
4. Hajder, L., Barath, D.: Relative planar motion for vehicle-mounted cameras from a single affine correspondence. IEEE International Conference on Robotics and Automation (2020)
5. Lowe, D.G.: Object recognition from local scale-invariant features. In: ICCV. IEEE (1999)
6. Raguram, R., Chum, O., Pollefeys, M., Matas, J., Frahm, J.M.: USAC: a universal framework for random sample consensus. TPAMI (2013), `https://www.cs.unc.edu/~rraguram/usac`
7. Scaramuzza, D.: 1-point-ransac structure from motion for vehicle-mounted cameras by exploiting non-holonomic constraints. IJCV **95**(1), 74–85 (2011)
8. Snavely, N., Seitz, S.M., Szeliski, R.: Photo tourism: exploring photo collections in 3d. In: ACM siggraph 2006 papers, pp. 835–846 (2006)
9. Zadrozny, B., Langford, J., Abe, N.: Cost-sensitive learning by cost-proportionate example weighting. In: International Conference on Data Mining. pp. 435–442. IEEE (2003)