Prior Guided Dropout for Robust Visual Localization in Dynamic Environments —Supplementary Document

Zhaoyang Huang[1,2] Yan Xu[2] Jianping Shi[2] Xiaowei Zhou[1] Hujun Bao[1*] Guofeng Zhang[1*]

[1]State Key Lab of CAD&CG, Zhejiang University[†]
[2]SenseTime Research

In this supplementary document, we provide more training details and additional experiments.

1. Training Details

All of the models are trained by Adam optimizer with learning rate of 1e-4, weight decay of 5e-4, batch size of 64 and the sampling threshold used in prior guided dropout is 0.2. On the Cambridge Landmarks, both of PoseNet and AD-PoseNet are trained 1000 epochs with initializations $\beta = 0.0$, $\gamma = -3.0$. On the Oxford RobotCar, we find that initializing loss weight with $\beta = -3.0$, $\gamma = -3.0$ helps PoseNet converge faster and achieve better accuracy. The color jitter data augmentation, which is assigned as 0.1, is applied to PoseNet and MapNet to reduce overfitting. The interval between the sampled image pairs is set to 10 frames when training MapNet to facilitate the learning of relative pose variations.

The *full* is comprised of 3 sequences (*2014-11-28-12-07-13*, *2014-12-02-15-30-08* and *2014-12-09-13-21-02*), which are captured from similar routes but at different time. We choose *2014-12-09-13-21-02* as test set and train the model on the other two sequences. Similarly, we use *2014-06-26-09-24-58* and *2014-06-23-15-41-25* in the *loop* when training, while *2014-06-26-08-53-56* and *2014-06-23-15-36-04*) when testing.

2. Additional Experiments

The uncertainty of pose prediction is obtained from multiple hypotheses, so the effectiveness of estimated variance is determined by sampling times and takes charge of the usefulness of predicted poses. To investigate the uncertainty generated by PGD (prior guided dropout), we carry out an experiment illustrated in Fig. 1 which explores the relation among sampling times, estimated variance and accuracy. The accuracy improvement (with the increasing sampling times and the decreasing variance threshold) suggests that the proposed two modules, prior guided dropout and composite self-attention, can compatibly cooperate with each other to esitmate the variance, which reliably improves the performance when utilized in uncertainty-aware PGO (pose graph optimization).

*Corresponding authors: {bao, zhangguofeng}@cad.zju.edu.cn

†The authors from State Key Lab of CAD&CG, Zhejiang University are also affiliated with ZJU-SenseTime Joint Lab of 3D Vision. This work was partially supported by NSF of China (Nos. 61822310 and 61672457), and the Fundamental Research Funds for the Central Universities (Nos. 2018FZA5011 and 2019XZZX004-09).

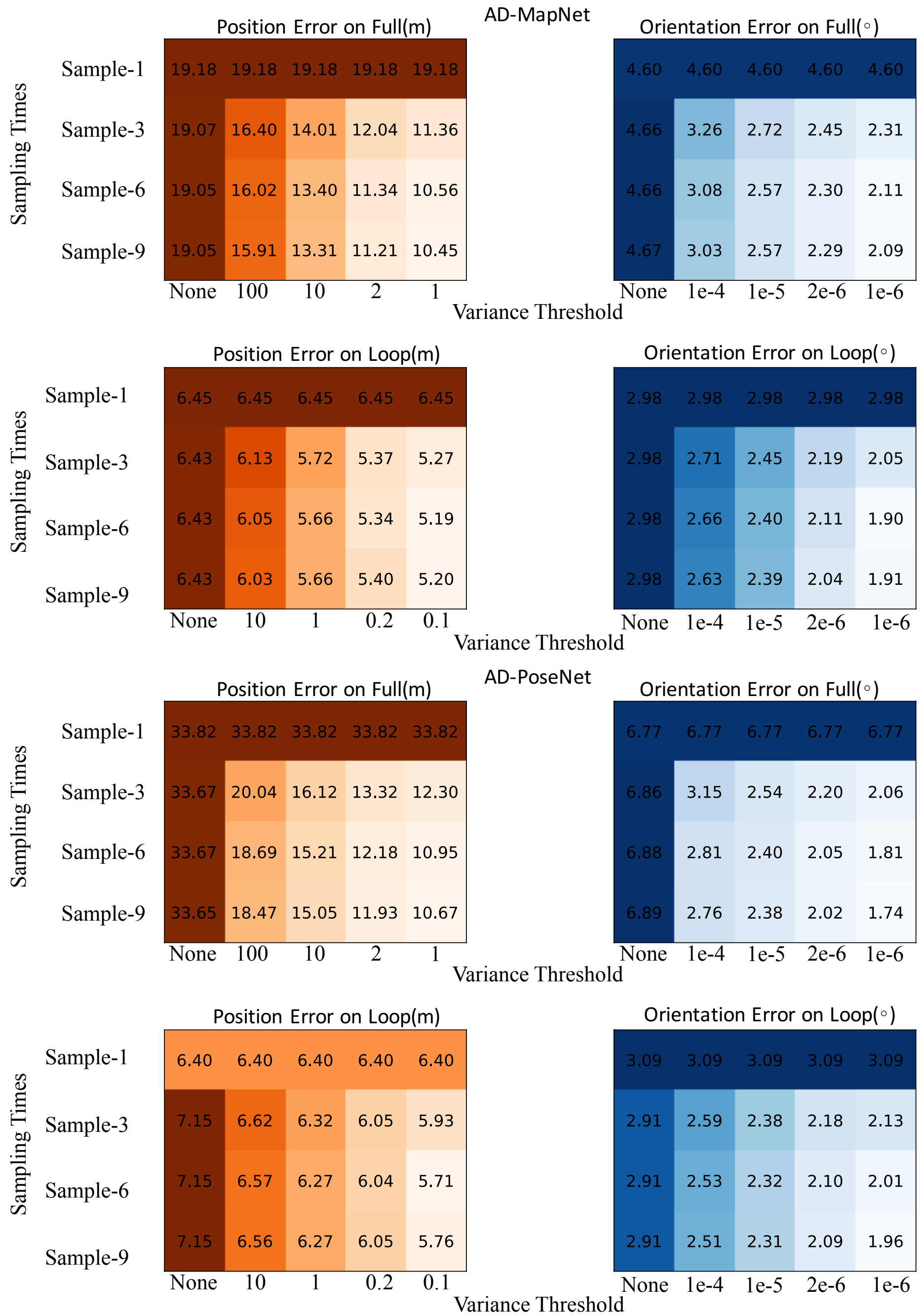

Figure 1: Relation analysis of sampling times, estimated variance and accuracy. We show the analysis of AD-MapNet on full (the top row) and loop (the second row), and AD-PoseNet on full (the third row) and loop (the bottom row) in the figure. In each matrix, the mean error is computed after distilling high-variance poses according to a decreasing threshold and the variance is generated by different sampling times. The accuracy does not change in the first row of each matrix because the variance is zero when we get one sample. Based on a fixed variance threshold, the error is lower when more sampling times are applied, and 3 samples are enough for variance estimation.