Supplementary Material:
NID-SLAM: Robust Monocular SLAM using Normalised Information Distance

Geoffrey Pascoe, Will Maddern, Michael Tanner, Pedro Piniés and Paul Newman
Oxford Robotics Institute
University of Oxford, UK
{gmp,wm,mtanner,ppinies,pnewman}@robots.ox.ac.uk

Abstract

In this ***supplementary document*** *we present additional details on the experimental datasets used for evaluation and the parameters for each of the three algorithms (NID-SLAM, LSD-SLAM and ORB-SLAM) used to generate results for the primary submission (Paper ID 536). We also present additional qualitative results for the indoor and outdoor datasets, illustrating the tracking performance of NID-SLAM under changing lighting and weather conditions. This document accompanies a* ***supplementary video*** *demonstrating the tracking and mapping performance of NID-SLAM in comparison to LSD-SLAM and ORB-SLAM for both indoor and outdoor environments.*

1. Experimental Datasets

Here we list the exact images and ground truth data used for both indoor and outdoor evaluation.

1.1. New Tsukuba Dataset

The New Tsukuba Dataset [3] was obtained from the original authors at http://www.cvlab.cs.tsukuba.ac.jp/dataset/tsukubastereo.php. We used all four illumination conditions (Daylight, Flourescent, Flashlight and Lamps) for evaluation, using only the 640×480 left image of the stereo pair for monocular SLAM. The provided ground truth poses were used for the tracking evaluation in Section 3.1 in the primary submission, and the metric ground truth depth maps were used for the mapping evaluation in Section 3.2 in the primary submission.

1.2. Oxford RobotCar Dataset

The sections of the Oxford RobotCar Dataset [2] were obtained from the original authors at http://robotcar-dataset.robots.ox.ac.uk/. We selected six subsets of the dataset in six different conditions detailed in Table 1. The start and end frame of each subset were manually aligned to the same physical location as closely as possible, and hence all subsets cover the same 500m distance. The images were obtained from the left camera of the forward-facing BumblebeeXB3 stereo camera. Bayer demosiacing and undistortion was performed at full resolution using the provided tools: http://robotcar-dataset.robots.ox.ac.uk/downloads/. The original 1280×960 images were then downscaled to 640×480 before use.

1.3. Outdoor Ground Truth

As metric ground truth is not available between traversals for the Oxford RobotCar Dataset, we generated a nearest-frame look-up-table using the provided visual odometry (VO) data. For each traversal we integrated the total distance travelled for each frame (between 0 and 500m). Ground truth correspondances to the other five traversals are determined by minimum absolute different in accumulated distance. Fig. 1 presents the accumulated distances travelled by all six traversals, and Fig. 2 illustrates sample correspondances for one of the traversals. The frame-based ground truth was used for the outdoor tracking reliability evaluation in Section 3.3 in the primary submission.

2. Algorithm Parameters

For repeatability we list the parameters used for each of the three algorithms (NID-SLAM, LSD-SLAM [1], ORB-SLAM [4]). Where parameters for LSD-SLAM or ORB-SLAM are not listed, we use the default values provided with the software. NID-SLAM parameters are listed in Table 2, LSD-SLAM in Table 3 and ORB-SLAM in Table 4.

3. Qualitative Results

In this section we present additional qualitative results illustrating the localisation performance of NID-SLAM in

Condition	Frames	Start Timestamp	End Timestamp	Filename
Sun	924	1403622333595478	1403622393158652	2014-06-24-15-03-07_stereo_left_01.tar
Rain	802	1437747791877082	1437747841995156	2015-07-24-14-17-50_stereo_left_01.tar
Overcast	908	1417176772138658	1417176830255804	2014-11-28-12-07-13_stereo_left_01.tar
Night	874	1441138754508641	1441138809134112	2015-09-01-21-16-05_stereo_left_01.tar
Dusk	1081	1424450359745880	1424450427299121	2015-02-20-16-34-06_stereo_left_01.tar
Snow	1491	1422953346020135	1422953439195435	2015-02-03-08-45-10_stereo_left_01.tar

Table 1. Images used for each traversal of the Oxford RobotCar Dataset. Each condition label links to a summary page describing the traversal, and each filename links directly to the file containing the images. Timestamps correspond directly to an image file in the dataset (e.g. 1403622333595478.png).

Parameter	Value
Num Histogram Bins (n)	16
Num Histogram Pyramid Levels (l)	3
Min Gradient for Depth	5
Min Frames per Key-frame	5
Min Key-frame Overlap	40%
BFGS Max Iterations Per Level	50
BFGS Max Line Search Iterations	20

Table 2. NID-SLAM Parameters

Parameter	Value
Num Pyramid Levels	5
Min Gradient for Depth	5
Min Frames per Key-frame	5
LM Min Step Size	10^{-8}
LM λ Success Factor	0.5
LM λ Fail Factor	2.0
Huber Width	3

Table 3. LSD-SLAM Parameters

Parameter	Value
Num Features	2000
Baseline Threshold	35
Num Levels	8
Level Scale Factor	1.2
Initial FAST Threshold	20
Min FAST Threshold	7

Table 4. ORB-SLAM Parameters

both indoor and outdoor environments. Fig. 3 presents key-frame appearances and depths for the map built with the indoor daylight traversal, along with localised frames with projected depth-maps for each of the four subsequent traversals. Despite significant appearance changes, NID-SLAM correctly registers the depth-map against the current image.

Fig. 4 presents the corresponding results for the outdoor dusk traversal. Again, despite changes in outdoor lighting and weather, NID-SLAM tracks the camera pose for the outdoor and snow traversal. However, it is not successful when attempting to localise at night against a map built at dusk.

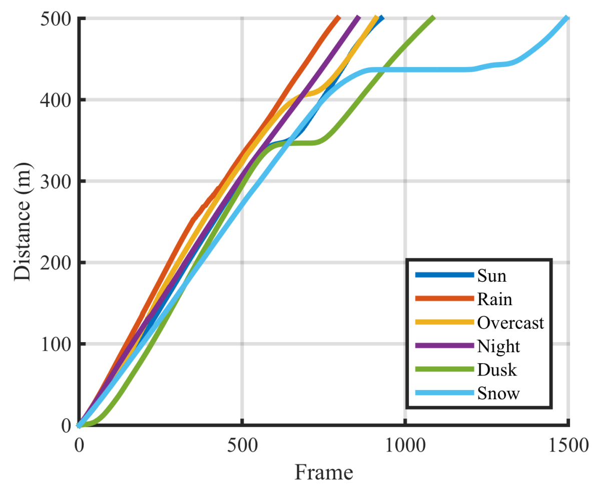

Figure 1. Accumulated ground truth VO distance for each of the six 500m traversals of the outdoor Oxford RobotCar Dataset.

References

[1] J. Engel, T. Schöps, and D. Cremers. LSD-SLAM: Large-scale direct monocular SLAM. In *European Conference on Computer Vision*, pages 834–849. Springer, 2014. 1

[2] W. Maddern, G. Pascoe, C. Linegar, and P. Newman. 1 Year, 1000km: The Oxford RobotCar Dataset. *The International Journal of Robotics Research (IJRR)*, to appear. 1

[3] S. Martull, M. Peris, and K. Fukui. Realistic CG stereo image dataset with ground truth disparity maps. In *ICPR Workshop: TrakMark2012*, volume 111, pages 117–118, 2012. 1

[4] R. Mur-Artal and J. D. Tardós. ORB-SLAM2: an open-source SLAM system for monocular, stereo and RGB-D cameras. *arXiv preprint arXiv:1610.06475*, 2016. 1

Figure 2. Sample images from nearest frame ground truth computed for six 500m traversals of the outdoor Oxford RobotCar Dataset. From left to right: sun, rain, overcast, night, dusk, and snow conditions. Correspondances are determined using accumulated VO distance.

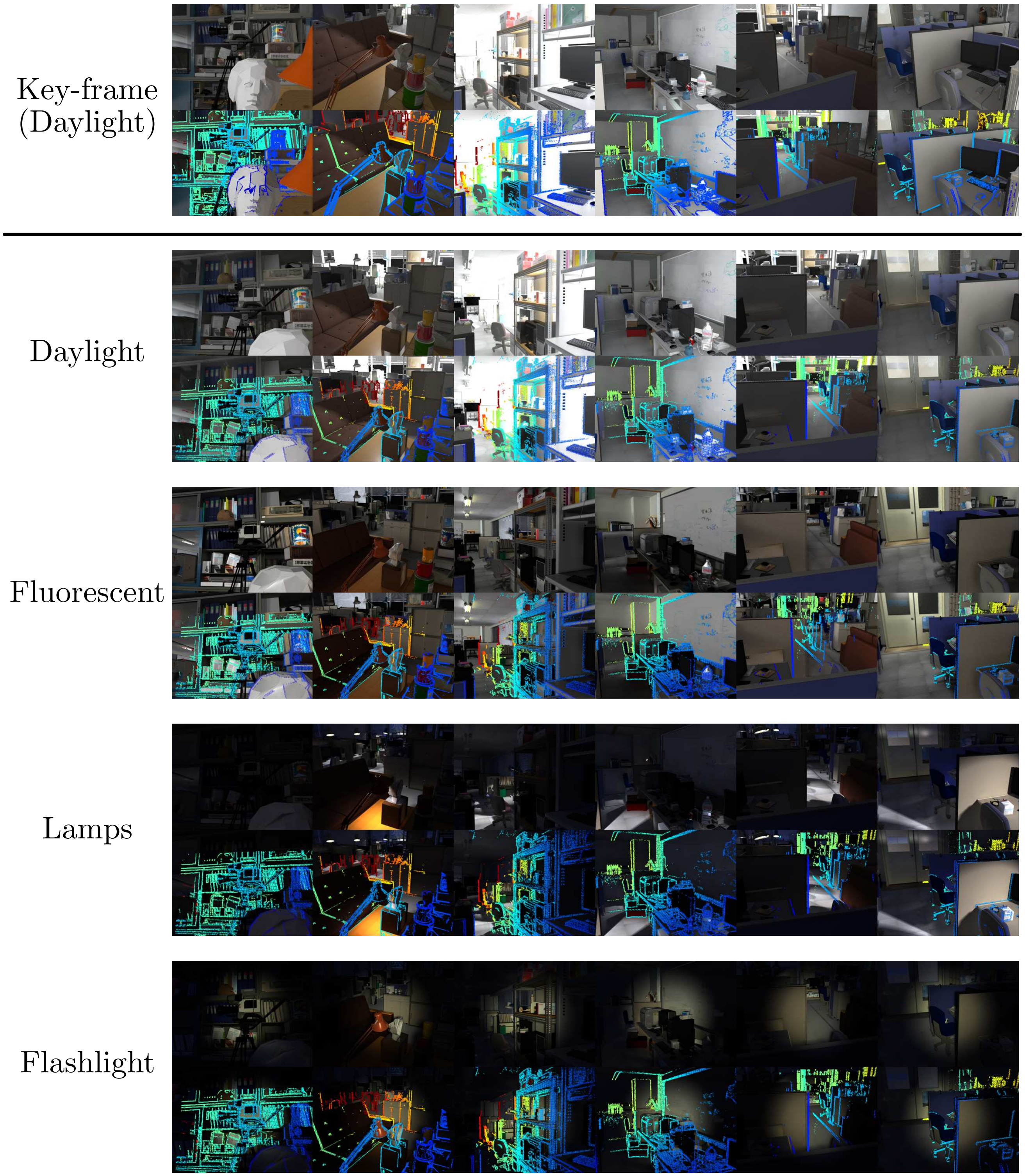

Figure 3. Sample depth maps projected into images during indoor localisation against a map built in daylight. Key-frame appearance and depth maps are shown at the top, and the results of the NID-SLAM tracking against the key-frame are shown for each of the four different conditions (daylight, fluorescent, lamps and flashlight), demonstrating robust tracking in the presence of indoor appearance change. Note that the images used for tracking are taken from a different location and view-point to the key-frame.

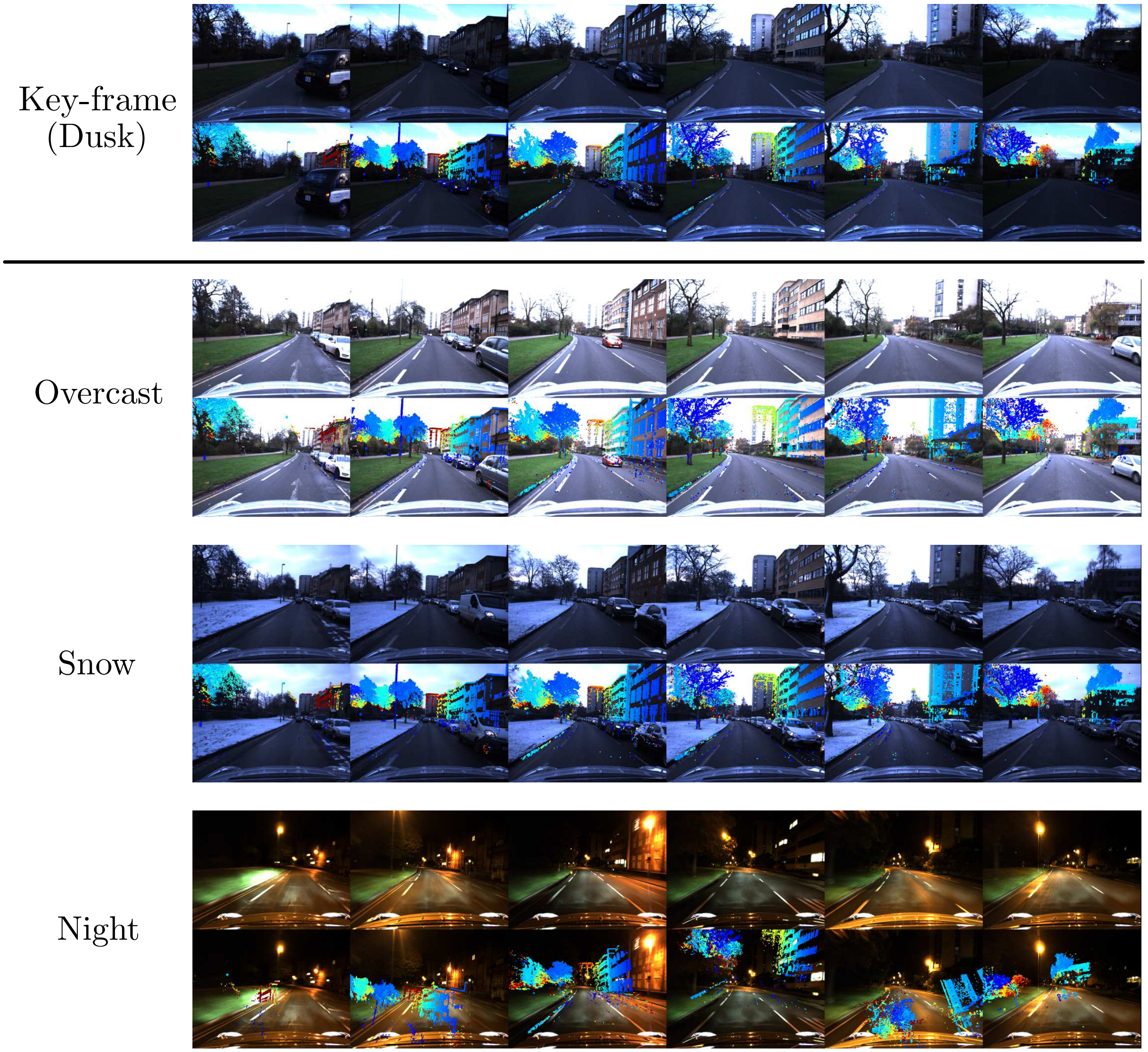

Figure 4. Sample depth maps projected into images during outdoor localisation against a map built at dusk. Key-frame appearance and depth maps are shown at the top, and the results of the NID-SLAM tracking against the key-frame are shown for each of three different conditions (overcast, snow and night), demonstrating robust tracking in the presence of outdoor lighting (overcast to dusk) and weather (snow to spring) changes. Note that NID-SLAM was not capable of localising at night relative to a map built under a different condition.