Real-time Kinematic Ground Truth for the Oxford RobotCar Dataset

Will Maddern, Geoffrey Pascoe, Matthew Gadd, Dan Barnes, Brian Yeomans, and Paul Newman

Oxford Robotics Institute, Dept. Engineering Science, University of Oxford, UK
robotcardataset@robots.ox.ac.uk

***Abstract*—We describe the release of reference data towards a challenging long-term localisation and mapping benchmark based on the large-scale *Oxford RobotCar Dataset*. The release includes 72 traversals of a route through Oxford, UK, gathered in all illumination, weather and traffic conditions, and is representative of the conditions an autonomous vehicle would be expected to operate reliably in. Using post-processed raw GPS, IMU, and static GNSS base station recordings, we have produced a globally-consistent centimetre-accurate ground truth for the entire year-long duration of the dataset. Coupled with a planned online benchmarking service, we hope to enable quantitative evaluation and comparison of different localisation and mapping approaches focusing on long-term autonomy for road vehicles in urban environments challenged by changing weather.**

I. Introduction

For real-world autonomous driving systems, the challenges of reliable localisation and mapping in changing conditions using vision and Light Detection and Ranging (LiDAR) are well documented, and many impressive solutions have been proposed. However, most of these approaches are evaluated on small-scale datasets with only a few examples of challenging conditions, or medium-scale datasets with limited variation in driving conditions. The well-known KITTI dataset and associated benchmarks [1] uses data gathered in good conditions over a period of five days, and hence only represents a small fraction of the conditions an autonomous vehicle can expect to encounter over its operational lifetime.

In this paper we present an important prerequisite in the form of the underlying reference data for the localisation benchmark for autonomous vehicles which we are developing using the *Oxford RobotCar Dataset* [2]. This large-scale dataset consists of image, LiDAR, and Global Positioning System (GPS) data collected over a year of driving a repeated route in Oxford, UK, covering over 1000 km of total distance. A wide range of variation including illumination, weather, dynamic objects, seasonal changes, roadworks, and building construction were captured during the course of data collection. To build a localisation ground truth, we have post-processed the raw GPS and Inertial Measurement Unit (IMU) data with Global Navigation Satellite System (GNSS) base station recordings to produce a centimetre-accurate Real-time Kinematic (RTK) solution. We offer the corrected RTK solutions for a subset of traversals, and withhold the remaining ground truth with a view towards an online benchmarking service similar to the KITTI Vision Benchmark suite [1]. By providing this important prerequisite for such a benchmark where researchers can quantitatively evaluate and compare localisation and mapping approaches on a challenging large-scale dataset, we hope to accelerate development of long-term autonomy for future autonomous vehicles.

II. Related Benchmarks

A number of urban driving datasets have been made available [3, 4, 5, 6, 7, 8], including datasets that focus on specific challenging scenarios [9] and general apperance change over time [10], but none of these offer a benchmarking service for comparison of results.

The KITTI dataset [11] offers a comprehensive benchmarking suite for stereo, optical flow, visual odometry, object detection and tracking. However, KITTI data was collected only over a period of 5 days and does not contain challenging weather conditions, nor does it revisit the same location at different times for evaluating localisation. Similarly, the Cityscapes dataset and benchmark [12] contains stereo imagery from a wide range of locations but does not revisit locations at different times.

The most similar benchmark is the VPRiCE Challenge[1] which contains several different sequences to evaluate loop closure in challenging conditions, including matching across night to day and between seasons. However, the localisation metric for evaluation is precision vs recall, which does not incorporate the true 6DoF metric pose relative to the prior map, and the locations are only traversed twice. In contrast, the reference data presented in this paper evaluates full 6DoF pose estimation over 72 traversals of the route over the period of a year, totalling approximately 650 km of driving.

III. The Oxford RobotCar Dataset(s)

Our reference data and benchmark builds upon the *Oxford RobotCar Dataset* [2], one of the largest available datasets for autonomous driving research. It consists of over 20 TB of vehicle-mounted monocular and stereo imagery, 2D and 3D LiDAR, and inertial and GPS data collected over a year of driving in Oxford, UK. More than 100 traversals of a 10 km route illustrated in Figure 1 were performed over this period to capture scene variation over a range of timescales, from the 24 h day/night illumination cycle to long-term seasonal

[1] http://roboticvision.atlassian.net/wiki/spaces/PUB/pages/14188617

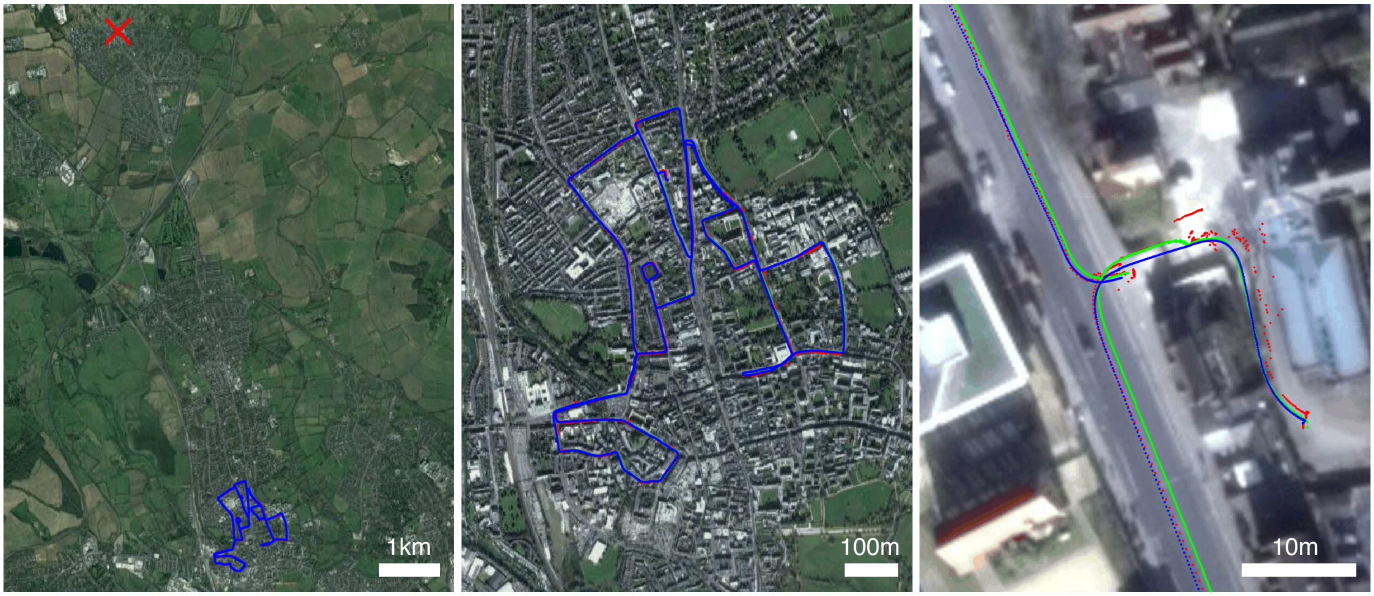

Figure 1. RTK ground truth for a single traversal from the *Oxford RobotCar Dataset*. Left: The GNSS base station location, shown with the red X, relative to the data collection route. Center: 80 traversals (72 released and 7 withheld) of the 10 km route have been processed for localisation ground truth. Right: Quality of the RTK position solution (blue) compared to the GPS-only (red) and GPS-inertial (green). A subset of seven RTK ground truth files have been withheld to form a planned localisation benchmark which is under development.

variations. For more details we refer the reader to [2].

We also refer interested readers to the *Oxford Radar RobotCar Dataset* [8]. While focused on millimetre-wave Frequency-Modulated Continuous-Wave (FMCW) scanning radar, this dataset provides over 280 km of new sensor data common to the original *Oxford RobotCar Dataset* as well as additional 3D LiDAR data

IV. RTK Ground Truth

We have produced the localisation ground truth using low-level raw GPS and IMU data collected by the NovAtel SPAN-CPT mounted to the RobotCar. The SPAN-CPT is a high-accuracy inertial navigation system (INS) equipped with dual GPS antennas, fibre-optic gyroscopes and MEMS accelerometers[2]. The raw recordings were not released as part of the original dataset.

GNSS base station data was obtained from the UK Ordnance Survey[3], consisting of RINEX corrections updated at 1 Hz. The recordings were sourced from the static base station in Kidlington, UK, approximately 8.2 km from central Oxford. Fig. 1 illustrates the location of the base station relative to the data collection area. Crucially, the base station remained stationary for the entire duration of the data collection and hence provides a consistent position reference for RTK corrections.

We have post-processed the raw GPS, IMU and GNSS base station data using NovAtel Inertial Explorer[4] to form an optimised corrected RTK solution for all trajectories using tightly-coupled GNSS and IMU observations. The RTK solution provides corrected global position and orientation at 10 Hz. Fig. 1 illustrates the quality of the RTK corrected solution in comparison to the GPS-only and GPS-inertial solutions available at runtime. The estimated positioning errors of the RTK solution are typically less than 15 cm in latitude and longitude and less than 25 cm in altitude, and the orientation errors are less than 0.01° in pitch and roll and 0.1° in yaw.

V. Error Evaluation

To benchmark localisation performance we plan to evaluate two metrics: root mean square (RMS) position and orientation errors, and uncertainty estimation. The RMS errors are evaluated as follows:

$$\sqrt{\frac{1}{N}\sum_{k=1}^{N}\left(\hat{\mathbf{x}}_k-\mathbf{x}_k\right)^{\mathrm{T}}\left(\hat{\mathbf{x}}_k-\mathbf{x}_k\right)} \tag{1}$$

where $\hat{\mathbf{x}}_k \in \mathbb{R}^{3\times1}$ is the position or orientation estimate at time k and $\mathbf{x}_k$ is the ground truth position or orientation (from the RTK solution), for a total of N estimates.

Table I presents example results for the errors computed with GPS-only and GPS-inertial solutions relative to the RTK ground truth on one trajectory.

VI. Archives Amended

In total, we are releasing RTK solutions for 72 forays. We exclude from release seven *RobotCar seasons* runs for which careful ground truth pose is curated in [13], namely:

- `2014-12-16-09-14-09` – dawn

[2]https://www.novatel.com/products/span-gnss-inertial-systems/span-combined-systems/span-cpt

[3]http://www.ordnancesurvey.co.uk/gps/os-net-rinex-data/

[4]novatel.com/products/software/inertial-explorer/

Method	Position Error (m)				Orientation Error (deg)			
	Lat	Lon	Alt	**Total**	Roll	Pitch	Yaw	**Total**
GPS	2.88	1.78	8.02	**8.71**	-	-	-	**-**
GPS+Inertial	1.24	1.06	6.92	**7.11**	3.06	0.25	2.45	**3.93**

Table I

EXAMPLE GPS ERROR EVALUATION RELATIVE TO RTK GROUND TRUTH FOR A SINGLE ROUTE.

- `2015-02-20-16-34-06` – dusk
- `2014-12-10-18-10-50` – night
- `2014-12-17-18-18-43` – night+rain
- `2015-05-22-11-14-30` – overcast (summer)
- `2015-11-13-10-28-08` – overcast (winter)
- `2015-02-03-08-45-10` – snow
- `2015-03-10-14-18-10` – sun

Where solutions for the eighth `25 Nov 2014 - rain` run were not available. This decision was made in order to ensure that there is a third party benchmark which calculates performance against a hidden ground truth signal for a swathe of challenging conditions from the original *Oxford RobotCar dataset*.

VII. Software Development Kit

The original Software Development Kit (SDK)[5] has been updated to allow existing methods to use the RTK solutions where Visual Odometry (VO) or Inertial Navigation System (INS) poses were already used interchangeably. Specifically, the building of pointclouds and projection of laser scans into cameras images can now be performed with the RTK solutions using the corresponding Python scripts:

- `build_pointcloud.py` and
- `project_laser_into_camera.py`

as well as MATLAB functions:

- `BuildPointcloud.m` and
- `ProjectLaserIntoCamera.m`.

The reader is referred to the dataset documentation[6] and the SDK documentation for more information.

VIII. Conclusions

We have presented the prerequisite reference data towards the planned *Oxford RobotCar Long-Term Autonomy Benchmark*, a new dataset for evaluating long-term localisation and mapping approaches for autonomous vehicles in dynamic urban environments. We expect to offer the benchmark as part of the *Oxford RobotCar Dataset* website[7] in the near future.

References

[1] A. Geiger, P. Lenz, and R. Urtasun, "Are we ready for autonomous driving? The KITTI Vision Benchmark suite," in *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*. IEEE, 2012, pp. 3354–3361.

[2] W. Maddern, G. Pascoe, C. Linegar, and P. Newman, "1 year, 1000 km: The Oxford RobotCar dataset," *The International Journal of Robotics Research*, vol. 36, no. 1, pp. 3–15, 2017.

[3] G. Pandey, J. R. McBride, and R. M. Eustice, "Ford campus vision and LIDAR data set," *The International Journal of Robotics Research*, vol. 30, no. 13, pp. 1543–1552, 2011.

[4] D. Pfeiffer, S. Gehrig, and N. Schneider, "Exploiting the power of stereo confidences," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2013, pp. 297–304.

[5] J.-L. Blanco-Claraco, F.-Á. Moreno-Dueñas, and J. González-Jiménez, "The Málaga urban dataset: High-rate stereo and LiDAR in a realistic urban scenario," *The International Journal of Robotics Research*, vol. 33, no. 2, pp. 207–214, 2014.

[6] "Waymo open dataset: An autonomous driving dataset," 2019.

[7] Z. Yan, L. Sun, T. Krajnik, and Y. Ruichek, "Eu long-term dataset with multiple sensors for autonomous driving," *arXiv preprint arXiv:1909.03330*, 2019.

[8] D. Barnes, M. Gadd, P. Murcutt, P. Newman, and I. Posner, "The Oxford Radar RobotCar Dataset: A Radar Extension to the Oxford RobotCar Dataset," *arXiv preprint arXiv: 1909.01300*, 2019. [Online]. Available: https://arxiv.org/pdf/1909.01300

[9] S. Meister, B. Jähne, and D. Kondermann, "Outdoor stereo camera system for the generation of real-world benchmark data sets," *Optical Engineering*, vol. 51, no. 02, p. 021107, 2012.

[10] H. Badino, D. Huber, Y. Park, and T. Kanade, "Real-Time Topometric Localization," in *International Conference on Robotics and Automation (ICRA)*, St Paul, Minnesota, USA, May 2012.

[11] A. Geiger, P. Lenz, C. Stiller, and R. Urtasun, "Vision meets robotics: The KITTI dataset," *The International Journal of Robotics Research*, vol. 32, no. 11, pp. 1231–1237, 2013.

[12] M. Cordts, M. Omran, S. Ramos, T. Rehfeld, M. Enzweiler, R. Benenson, U. Franke, S. Roth, and B. Schiele, "The Cityscapes dataset for semantic urban scene understanding," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 3213–3223.

[13] T. Sattler, W. Maddern, C. Toft, A. Torii, L. Hammarstrand, E. Stenborg, D. Safari, M. Okutomi, M. Pollefeys, J. Sivic *et al.*, "Benchmarking 6dof outdoor visual localization in changing conditions," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 8601–8610.

[5] https://github.com/ori-mrg/robotcar-dataset-sdk

[6] https://robotcar-dataset.robots.ox.ac.uk/documentation

[7] http://robotcar-dataset.robots.ox.ac.uk