



FinnForest dataset: A forest landscape for visual SLAM

Ihtisham Ali^{a,*}, Ahmed Durmush^a, Olli Suominen^a, Jari Yli-Hietanen^a, Sari Peltonen^a, Jussi Collin^b, Atanas Gotchev^a

^a Faculty of Information Technology and Communication Sciences, Tampere University, Finland

^b JC Inertial Oy, Finland

ARTICLE INFO

Article history:

Received 6 February 2020

Received in revised form 24 July 2020

Accepted 28 July 2020

Available online 3 August 2020

Keywords:

Forest

Dataset

SLAM

Visual odometry

Navigation

Localization

Mapping

Stereo

Autonomous driving

Mobile robotics

Field robotics

Computer vision

ABSTRACT

This paper presents a novel challenging dataset that offers a new landscape of testing material for mobile robotics, autonomous driving research, and forestry operation. In contrast to common urban structures, we explore an unregulated natural environment to exemplify sub-urban and forest environment. The sequences provide two-natured data where each place is visited in summer and winter conditions. The vehicle used for recording is equipped with a sensor rig that constitutes four RGB cameras, an Inertial Measurement Unit, and a Global Navigation Satellite System receiver. The sensors are synchronized based on non-drifting timestamps. The dataset provides trajectories of varying complexity both for the state of the art visual odometry approaches and visual simultaneous localization and mapping algorithms. The full dataset and toolkits are available for download at: <http://urn.fi/urn:nbn:fi:att:9b8157a7-1e0f-47c2-bd4e-a19a7e952c0d>. As an alternative, you can browse for the dataset using the article title at: <http://etsin.fairdata.fi>.

© 2020 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The intense competition to develop a safe marketable self-driving car has motivated a huge amount of research in the field of autonomous vehicles. Coupled with the growing interest of companies to put self-driving cars on the road, various companies are also interested in introducing other forms of autonomous vehicles to automate various industrial processes such as mining, shipping, agriculture, and forestry. Irrespective of the industry and targeted operations, the autonomy of any machine is highly dependent on advancements in a number of vision technologies, such as object detection [1], reconstruction quality [2], scene perception [3]. However, the base capabilities of an autonomous vehicle that need most attention remain visual odometry, relocalization and mapping [4]. This requires testing and validation on all scenarios that a vehicle/machine can face in a simulated environment. A variety of public datasets are available that provide a good amount of data for testing in various conditions and locations. We will mention some of the well-known datasets in an attempt to measure the expanse of the collections and find a horizon. Most of these datasets focus on urban environments

(for example, [5–9]) in order to facilitate testing on public roads in urban areas. The earliest among these datasets that recorded urban environment are Ford Campus [10] and KITTI [11]. Being among the first public datasets in the field, these datasets contributed significantly towards testing and validation. The recent additions to the publicly available datasets are The Oxford RobotCar [12], KAIST Multi-Spectral Day/Night [13], and Complex Urban LiDAR Data Set [14]. All these datasets, when combined, provide a significant amount of testing data for urban environment with short and long trajectories at various speeds [10]. Moreover, they incorporate weather and seasonal changes [12], long term changes in urban structure [12] and gradual/sudden illumination variations [13]. However, all these datasets target indoor or outdoor urban environment.

In contrast, some unique datasets target entirely different environments to assist automation of other form of vehicles. Among these are Aqualoc Underwater [15], Canoe [16] and Underwater Caves SONAR and Vision Dataset [17]. These datasets comprise of data acquired for under water exploration and surface sailing conditions. On the other hand, a few public datasets target more domain specific terrains for their experimentation. In [18], authors recorded data in the Chile's largest underground production-active copper mine. The data was recorded for a length of approximately 2 km using Lidar, radar and stereo cameras fixed on

* Corresponding author.

E-mail addresses: ihtisham.ali@tuni.fi, ihtishamalikt@gmail.com (I. Ali).



Fig. 1. Recording platform. Our vehicle is equipped with four Basler HD cameras, a NovAtel GNSS antenna, and a KVH 1750 IMU with fiber optic gyro.

a robotic platform. The Devon Island rover navigation dataset [19] provides a dataset for testing rovers for planetary explorations. The dataset was recorded on Devon Island in the Canadian High Arctic, which is assumed to be analogous to Moon/Mars terrains due to the wide variety of geological features and microbiological attributes of the site.

The dataset that has high relevance to our work is the SFU Mountain dataset [20]. The study used a mobile ground based robot to traverse walking trails in the Burnaby Mountain, British Columbia, Canada. The dataset provides a semi-structured woodland terrain with different illumination and weather conditions and with changing vegetation, infrastructure, and pedestrian traffic. The dataset provides a good amount of data for visual odometry, however, lacks to present opportunities to test loop closure and re-localization.

In [21], the authors carried out brief experiments on the SFU Mountain dataset and their own dataset, Hillwood. The Hillwood dataset consists of photorealistic rendered and real forest video scenes. However, the Hillwood dataset only provide video recordings for testing without any ground truth information. In their conclusive remarks, the authors stressed upon the need and advantage of actual forest dataset with complete synchronized groundtruth poses [21].

In this paper, we present a new dataset that will target a real forest landscape recorded in the outskirts of Tampere, Finland. The goal is to provide testing data in order to facilitate the research towards increasing the autonomy of vehicles traversing rural areas and heavy machines working in the forest. Unlike urban settings, a terrain environment provides fewer discriminate landmarks and more repetitive textures in the scene. Presumably, such a situation strengthens visual odometry to some extent, however, affects adversely relocalization algorithms. This dataset provides semi-structured forest routes under different conditions (i.e. lighting, weather, vegetation, and infrastructure) in a highly self-similar natural environment. Furthermore, the sequences include scenes that best replicate the motions (i.e. stationary, sharp motion, bumps and potholes, slopes, and back-and-forth motion) and environments (i.e. log piles, close-up of trees, off-road routes) involved in actual forestry operations. The dataset includes unique trajectories to test both visual simultaneous localization and mapping (visual-SLAM) and visual odometry algorithms thoroughly. Moreover, each path is traversed in two different condition, namely sunny summer and snowy winter. The dataset provides images from 4 cameras and ground truth poses for each sequence in each condition using Global Navigation Satellite System (GNSS) and Inertial Measurement Unit (IMU) data fusion. We provide processed rectified images, calibration

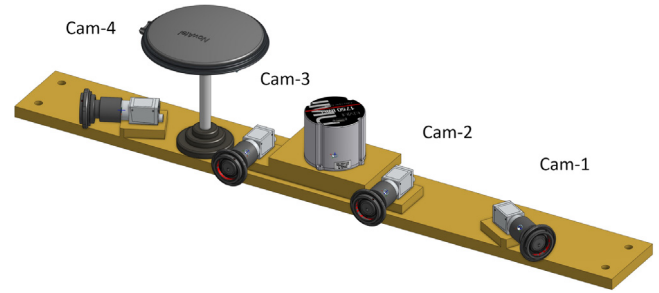


Fig. 2. Rendered 3D model of the sensor rig.

data and ground truth at three sampling rates i.e. 40, 13.33 and 8 Hz except for two sequences which are sampled at 20, 10 and 7 Hz. For simpler representation, here onwards, we will approximate 13.33 to 13 Hz in the manuscript. Additionally, we provide raw images (40 Hz) for most of the sequences and the calibration images to the public. For this purpose, we also provide development tools to process raw data and evaluation tools in order to facilitate benchmarking against the state-of-the-art.

We hope this dataset provides a good reference for rural, forest and general terrain environment in order to facilitate the researchers to mitigate the challenges faced in this field of research.

2. Recording platform and sensor configuration

The recording platform and the arrangement are illustrated in Fig. 1. The data was recorded using a sensor rig mounted on a vibration dampening platform affixed to the vehicle. The vibration dampening platform was affixed to the Jeep using strong suction cups. The rig houses all the sensors as shown in Fig. 2.

The sensor and hardware specifications are as follows:

- (i) 4 × Basler acA1920-50gc GigE camera with the Sony IMX174 CMOS Color sensor, Resolution (HxV) 1920 × 1200 (2.3 MPx), 84° HFoV, 59° VFoV, 6 mm Focal Length Lens, 20 cm baseline for each stereo pair.
- (ii) 1 × KVH 1750 IMU, fiber optic gyro, bias instability ≤ 0.05 °/h, 1σ, 200 Hz.
- (iii) 1 × NovAtel PwrPak7, OEM7 GNSS, 20 Hz.
- (iv) 1 × CC320 Machine Vision Timing Controller, 8 Digital Inputs of 5 V to 24 V at 3 mA to 20 mA, 8 Digital Outputs of 24 V and 20 mA.
- (v) 1 × Embedded system with Quad Core Intel Core i7 processor, 2 DDR4 with 64 GB memory, 6 GigE LAN with 4 PoE.

For the sensors and their coordinate systems, we use the following notations,

- C₁ Camera 1
- C₂ Camera 2 (reference frame)
- C₃ Camera 3
- C₄ Camera 4
- I IMU
- G GNSS

All cameras were connected to the embedded computer. The cameras stored data on the computer while the IMU and GNSS data was recorded on the internal memory of the NovAtel Module. To minimize write latency into storage and to prevent losses, we used CAT7 cable and wrote on SSDs using parallel threads for all the cameras.

To obtain high quality images it was essential to control the exposure time of the cameras during the acquisition. To minimize the effect of motion blur, the exposure time was kept below

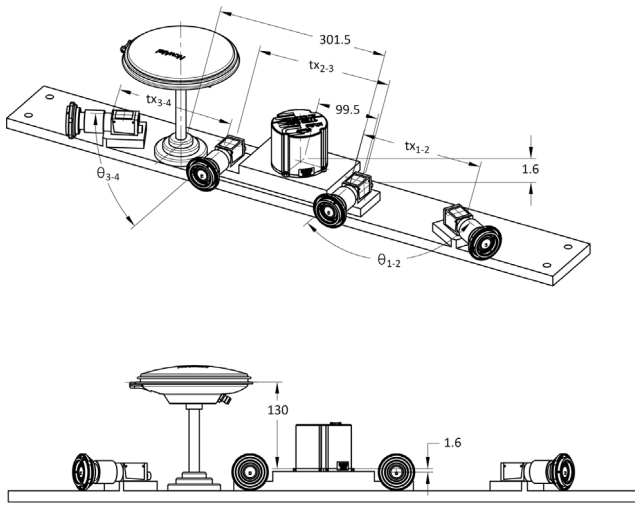


Fig. 3. The figure illustrates the mounting positions of the sensors with respect to each other from various views. The calibrated transformations are provided as part of the dataset.

10 ms. Moreover, to obtain images from four cameras for stereo analysis, it was of the utmost importance to enforce synchronicity. Hence, to acquire synchronized feed from four cameras at 40 fps on a Windows based platform, we utilized a special purpose triggering hardware known as Machine Vision Timing Controller. This timing controller or triggering device sent synchronized signals to all the cameras in order to enforce realtime consistent capture. Additionally, one trigger signal was sent to the NovAtel Module from the triggering device to generate timestamps. The NovAtel module was configured to store a timestamp in GPS time upon receiving the signal from the triggering device. The GPS time is more accurate and does not drift, unlike the clock on the Windows platform. This timestamp signal was sent at a delay of 1.5 ms. Even though this would have had a negligible effect, nonetheless, we compensated for this delay during the ground truth generation. The IMU and GNSS data are pre-synchronized by the NovAtel receiver. Hence, we have a precise synchronization among the cameras, IMU and GNSS data in effect. The raw data of GNSS and IMU are acquired at 20 Hz and 200 Hz, respectively. However, they are not the limits of the system. The maximum acquisition rate of the system is 100 Hz and 1000 Hz, respectively. For this study 20 Hz GNSS is used and interpolated to 50 Hz with the IMU data during post-processing by NovAtel Inertial Explorer software.

The sensor arrangement is illustrated in Fig. 3. It constitutes four cameras, a GNSS antenna and an IMU unit. The sensor rig has the middle cameras (C_2 and C_3) facing forward and houses the IMU unit in between them. The outward facing cameras (C_1 and C_4) are at nearly same angle from the forward direction. The motivation backing this camera arrangement is to test the effects of various camera configurations on the accuracy of joint perception. It is mostly observed in SLAM implementations that during forward motion, the view is dominated by consistently tracked areas of interest that are further away from the camera. This negatively affects the scale estimation for visual odometry. This is more apparent in monocular SLAM algorithms where the SLAM methods fail at certain point because the further points do not exhibit enough disparity change. The methods survive as long as the closest features are not lost due to motion blur. However, if the camera is fixed at an angle, instead of facing the forward direction, then the effective area in which the feature points exhibit disparity change increases as well.

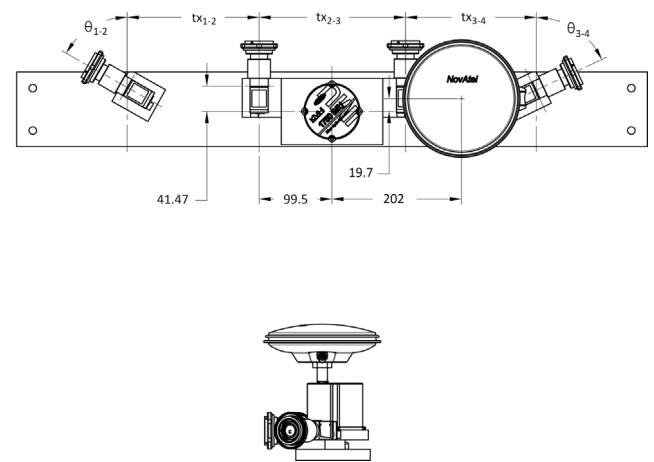


Fig. 4. The GPS trajectory of our recordings in the forest area in the outskirts of Tampere, Finland.

3. Data overview

Our primary contribution through publishing this dataset is to provide publicly accessible data recorded in forest for research towards Advanced driver-assistance systems (ADAS) and autonomous work machines. In general the dataset provides challenges by incorporating sequences that are recorded at various times of day and weather conditions. Moreover, the sequences have been recorded so that they present considerable challenges for both visual odometry and SLAM approaches. The area explored during the course of the recording sessions can be viewed in Fig. 4. The dataset comprises unique trajectories, most of which are recorded in two seasonal conditions. In winter, a part of the route was blocked due to heavy snow and could not be re-recorded in snowy conditions. An overview of the dataset is provided in Table 1. The dataset offers a total of 11 sequences. We provide the dataset at different sampling rates to facilitate testing. The original visual data was recorded at 40 Hz and later subsampled to facilitate testing. The subsampled versions are provided in the form of compressed image packages and Rosbags.

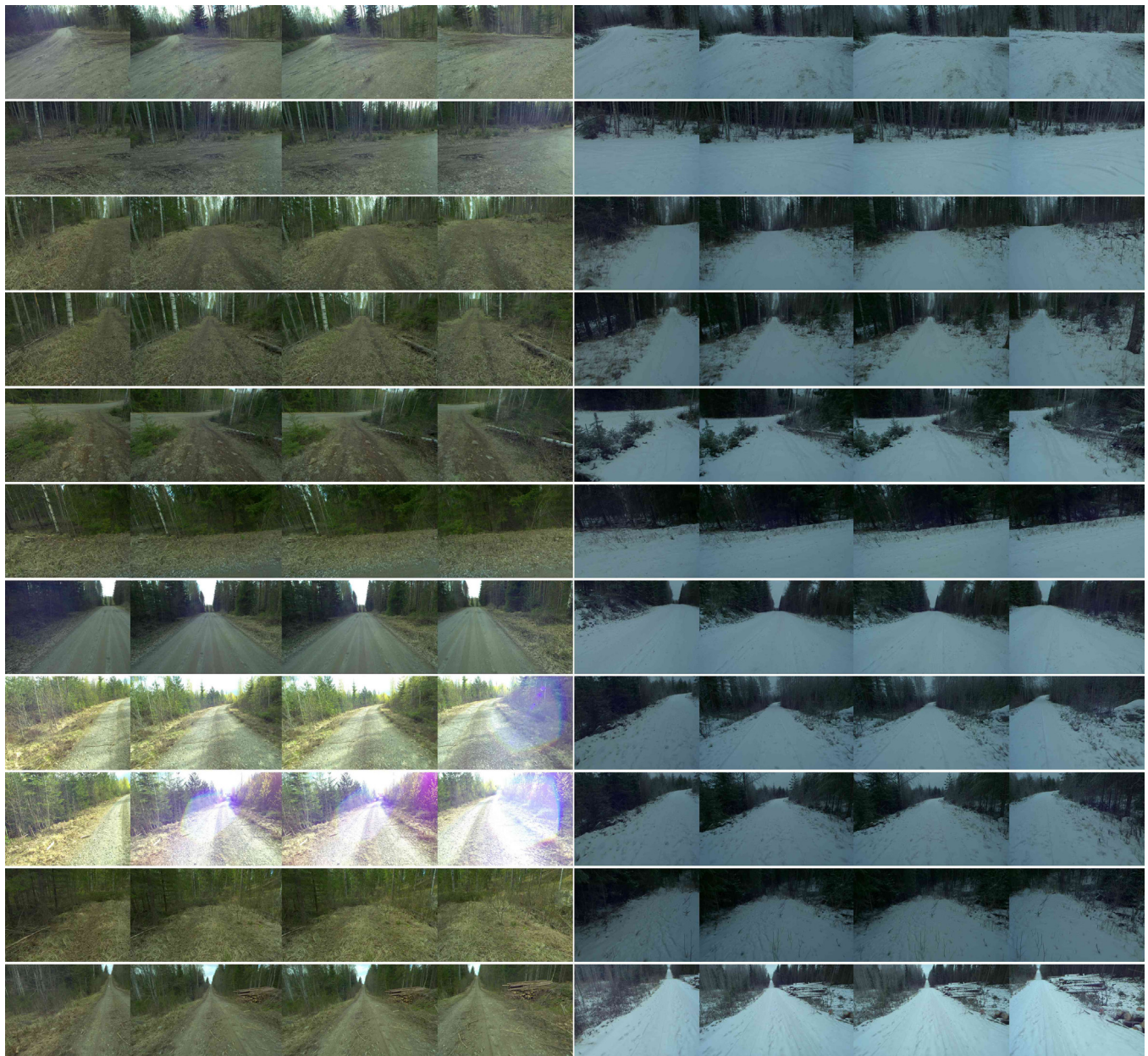


Fig. 5. Montage of images from all cameras arranged left to right illustrating the significant changes in appearance of same scene over seasonal and route changes.

Table 1

An overview of the nature of data in the FinnForest Dataset.

Seq. No	Frames			Distance (km)	Loop	Season	Time of Day
	40/20 Hz	13/10 Hz	8/7 Hz				
W01	27 630	9210	5526	1.29	Yes	Winter	Daylight, Overcast
W03	23 100	7700	4620	1.69	No	Winter	Daylight, Overcast
W04	37 010	12 337	7402	2.35	No	Winter	Daylight, Overcast
W05	57 288	19 096	11 458	4.74	No	Winter	Daylight, Overcast
W06	20 875	10 438	6959	3.59	No	Winter	Night
W07	43 780	21 890	14 594	6.48	No	Winter	Dusk, varying illumination
S01	27 960	9320	5592	1.29	Yes	Summer	Daylight, Sunny
S02	21 333	7111	4267	1.99	Yes	Summer	Daylight, Sunny
S03	15 000	5000	3000	1.69	No	Summer	Daylight, Overcast
S04	30 662	10 221	6133	2.32	No	Summer	Daylight, Sunny
S05	61 662	20 554	12 333	5.84	No	Summer	Daylight, Overcast

The number of frames at each sampling rate is provided against the sequence name in Table 1. Three of the sequences offer loop closure opportunities while the remaining sequences are aimed at

testing visual odometry. We have also tabulated the distance covered while traversing each path. The range of distance traveled varies from 1.3 km to 6.48 km. Information regarding the season



Fig. 6. Illustration of the drastic changes in appearance of the scene produced by different illumination and weather conditions.

and the illumination condition is also provided corresponding to each sequence. The seasonal name is also abbreviated in the name of each sequence for clarity. The dataset covers a variety of conditions with different illumination such as overcast, direct sunlight, dusk and night. However, we would like to state that the dataset does not offer sequences with rain and fog which would have provided further useful information for testing. Further details about the unique challenges of each sequence are provided in Section 7, Discussion.

Fig. 5 presents a montage of selected images illustrating the range of varying appearances of the environment encountered as a result of different season and routes. The left half of the montage constitutes the left to right camera images from the summer dataset while the right half of the montage shows the left to right images from the winter dataset of same scene from almost the same vehicle locations.

Fig. 6 illustrates the changes in appearance of the scene from almost similar camera perspective and location during both seasons and the challenges it brings about. The dark overcast in winter demands longer exposure time and slower vehicle motion to capture the details in the scene accurately. On the other hand, summer season presents challenges like overexposure, rain, puddle and flares in the scenes. The last row of images exhibit the conditions of a night and dusk time with varying illumination.

The high resolution and frame rate of the data recordings make it challenging to store the data on online data repositories. In order to make the usage of the data convenient for users, we have split the dataset into subset sequences. Each sequence can be downloaded and used independently as a .zip package at three sampling rates. Moreover, the most common configuration preferred for stereoscopic analysis is parallel, hence, we only provide the processed images from the forward facing stereo pair i.e. C_2 and C_3 . Nonetheless, the raw images from all the cameras

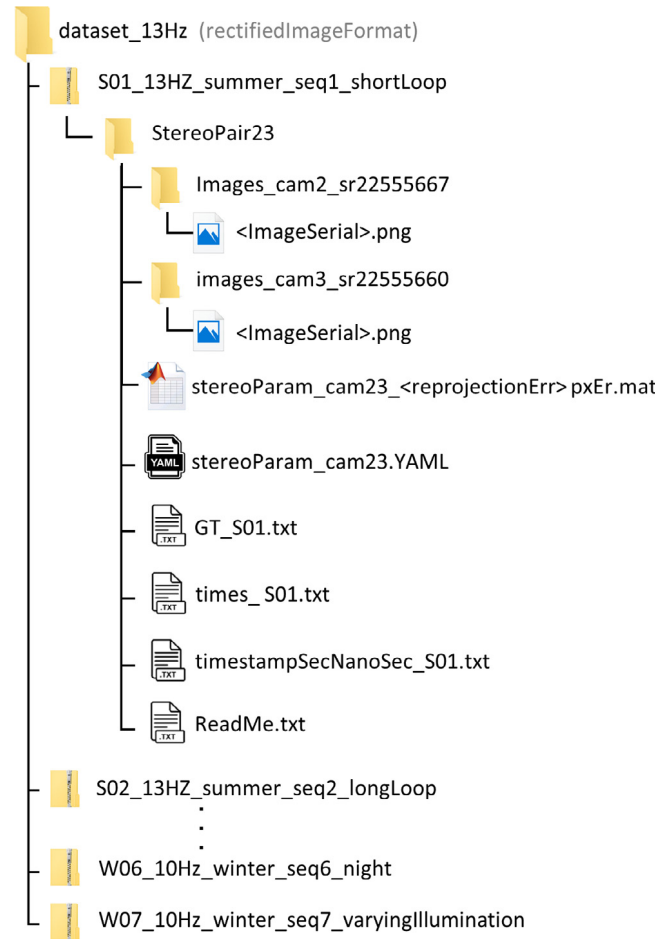


Fig. 7. Directory layout for a sub-sampled subset from the dataset. Extracting them will preserve the folder structure.

are provided in the dataset along with a toolkit to easily extract and process them in ready to use format. The MATLAB toolkit readily extracts the raw images into stereo pairs C_1 - C_2 , C_2 - C_3 and C_3 - C_4 .

The data structure or format for each sequence is illustrated in Fig. 7. The name of the folder constitutes the nature of the data and the rate at which it is sampled. Each sequence is self-contained and is provided with supporting files inside the compressed file format. The compressed file in turn constitutes sub folders, which correspond to the stereo pairs for forward facing cameras (C_2 and C_3). The Rosbag version contains the Rosbag file instead of the PNG image files for the cameras (C_2 and C_3). Additionally, the calibration files, timestamps and the ground truth poses are provided in the corresponding directories for the rectified cases. The ground truth data already corresponds directly to the images provided and does not need further matching or synchronization. Each row of the ground truth text file corresponds to a new reading of the ground truth pose of 3×4 matrix $[R|t]$ in the row first vectored form as shown below:

$$[R_{11} \ R_{12} \ R_{13} \ t_x \ R_{21} \ R_{22} \ R_{23} \ t_y \ R_{31} \ R_{32} \ R_{33} \ t_z]$$

4. Sensor calibration and ground truth

In this section, we will discuss two forms of calibration that are essential to use the data effectively.

4.1. Cam-to-cam calibration

The first calibration step is the camera-to-camera calibration, which is performed to compute the intrinsic parameters and extrinsic transformations for the cameras. In the dataset, we have included both the processed data (using the calibrations) and the raw data. The processed data from the cameras can be directly used with any SLAM pipeline using the provided calibration parameters. However, for researchers who wish to re-calibrate the cameras and process the raw data themselves, we have included the raw images along with the calibration images in the dataset.

The calibration images are provided as stereo pairs between the nearest two cameras. Special attention was given to calibration by recalibrating the cameras for each recording session. Although, the sensor setup was not altered, some minute numerical differences are possible. It is strongly recommended to use the calibration parameters from the calibration files and not the illustrations. The camera-to-camera calibrations are provided for the nearest camera pairs, namely C_1 - C_2 , C_2 - C_3 and C_3 - C_4 . These camera pairs are jointly calibrated using MATLAB stereo calibration toolbox for their intrinsic and extrinsic parameters based on the approach presented in [22]. The calibration information is provided in two forms, namely MATLAB stereo-parameters object file and a text file with excerpts of the object file along the dataset.

4.2. Cam-to-IMU calibration

We calibrate the camera and IMU in order to obtain the external transformation between the camera and IMU unit. For this, a sequence was recorded in front of the calibration board, where the motion in all the six degrees of freedom was stimulated by moving along and around each axis. The relation between the camera and the IMU is then analogous to hand-eye calibration problem. For this, we utilize Kalibr toolkit [23] which estimates the spatial and temporal parameters of a camera system with respect to an intrinsically calibrated IMU. Since we have an accurate synchronization between the images acquired from the camera and the data from IMU/GNSS using the timestamps, we are not interested in the temporal relationship provided by the toolkit. However, the spatial parameters or the extrinsic transformation between the camera and the IMU is of interest to this work. We calibrate the IMU unit with the camera C_2 . We choose camera C_2 for calibration in order to be consistent with our ground truth coordinate system and the general approach of choosing a forward facing camera.

4.3. Ground truth quality evaluation

Acquiring ground truth information in an enclosed environment is a challenging step. The global accuracy of the ground truth solution is dependent on the availability of GNSS signals. In general, the strength and accuracy of GNSS signals are high in an open area, while poor signals are received in enclosed areas such as indoors, narrow city streets and forests. On the other hand, the local accuracy can be improved by fusing the information acquired through local sensing mechanisms such as IMU, Odometer, Radar, Lidar, Camera, etc. with the GNSS information for better results. As mentioned earlier, we utilize the NovAtel's PwrPak7TM module to acquire a ground truth solution through a tightly coupled pose estimation framework that uses GNSS and IMU information.

To assure the readers of the quality of the ground truth, we provide the estimated position accuracy in the form of standard deviations for the positions at every timestamp for all sequences

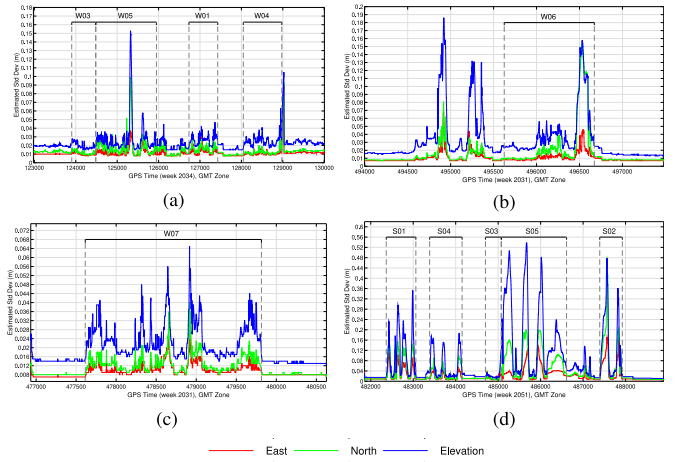


Fig. 8. Estimated position accuracy for ground truth poses (a) Winter sequences W01–W05 (b) Winter sequence W06 (c) Winter sequence W07 (d) Summer sequences S01–S05.

in Fig. 8. The graphs indicate the standard deviation in the estimated position in the North, East and Elevation/Height directions. The accuracy in North, East, and Elevations directly correspond to the accuracy in the local coordinate frame. The statistics in these figures are provided by the Inertial Explorer application used with the NovAtel's PwrPak7TM module.

To facilitate readers, we show the range of each sequence in the figures. It can be observed from Fig. 8(a–c) that the average standard deviation for the winter sequences (W01, W03–W07) is lower than 2 cm for East and North with occasional larger deviations. The spikes in deviation are obtained where the vehicles traverse a narrow path with trees densely covering the area around it for a longer period. In all the sequences, the errors in the East axis are the lowest followed by errors in the North. The largest deviations are found in the elevation, which is typical of such a system.

On the other hand, the summer sequences (S01–S05) exhibit slightly larger standard deviations (see Fig. 8(d)). Except for S02, the errors for all the summer sequences in East and North are lower than 15 cm and 20 cm, respectively. As before, the largest deviation is observed in the elevation. This is in the sequence S05 with a value of 0.54 m. The deterioration of the GNSS performance for the summer sequence is logical. In contrast to the winter sequence, which was recorded in December 2018, the summer sequence was recorded near the springtime of May 2019. In the springtime, the GNSS results can be affected by the foliage which can cause 24 to 35% attenuation at L-band [24]. The contributing factor to the attenuation of the signals is the combined effect of signal absorption and scattering from the conglomeration of tree canopies and trunks. In winter, the sparsity of foliage in the tree canopy provides for a larger interval of non-attenuating space, while that advantage is lost in springtime in the presence of dense foliage [25]. In the absence of the GNSS signal, the ground truth pose estimation system relies more on the information provided by the IMU. Nonetheless, considering the task at hand, the results obtained for the summer sequence are good and provide a valid reference for experimentation.

5. Development and evaluation toolkit

The dataset is accompanied by a set of MATLAB scripts that can be used for processing of raw data or evaluating the odometry obtained from user's algorithm against the ground truth poses. The dataset includes ready to be used information for easy access

to the researchers. Nonetheless, we provide a set of MATLAB tools for processing the data. Each data sequence is accompanied by a set of raw images. The raw data is of interest to the researchers who wish to re-calibrate the cameras using the set of calibration images provided with the dataset with their own or different calibration algorithms. The new calibration can then be used to process the raw images of the dataset. MATLAB script *readRaw_Debayer.m* and *readRaw_Rectify.m* can be used to read the raw images from a folder and write the debayered and rectified images onto another directory, respectively. The debayered color images can then be used with the provided calibration data or any newly computed calibration data using the calibration images.

An evaluation script is also provided as part of the toolkit to assess the results. The MATLAB script *mainEvaluate.m* can be used to evaluate the obtained visual odometry poses against the ground truth poses. Prior to using the script, the directories for the text file with the ground truth poses and the self-computed poses should be specified. The evaluation script computes relative pose error (relative translation and rotation errors) and absolute trajectory error (ATE) for each sequence and the overall errors for all sequences. The core reason for selecting these metrics is that relative pose error provides a good analysis of the local accuracy of the trajectory over a fixed distance. Relative comparison over fixed distances can measure the effect of drift more effectively and provide a better response for visual odometry. On the other hand, ATE provides a more coherent and globally consistent comparison using the absolute distances between the corresponding ground truth poses and the poses estimated by the assessed system.

The ATE can be obtained by computing the absolute distance between the estimated and the ground truth trajectory. For global consistency, it is essential that both trajectories are in the same reference coordinate system. If that is the case, then the ATE can be computed directly, otherwise, the alignment can be calculated in the form of a transformation matrix T in closed-form using Umeyama's method [26]. It is to be noted that ATE only considers the translational errors. The commonly used form of ATE is given as follows [27]

$$ATE_{rmse} = \sqrt{\frac{1}{N} \sum_{i=1}^N \|Tp_i - \hat{p}_i\|^2}, \quad (1)$$

where $T \in SE(3)$ transforms the trajectory p_i to the coordinate system of the ground truth poses \hat{p}_i . Additionally, the mean, standard deviation, minimum, median and maximum errors can be computed to analyze the performance from different perspectives.

As mentioned earlier, relative errors can provide more accurate local information about visual odometry errors. Kümmerle et al. [28] proposed to compute relative error over an interval followed by an average over all these errors. The interval was selected based on fixed distance. This is a good approach, however, the trajectory and orientation errors are amalgamated and form a joint error metric. Geiger et al. [11] took this concept and isolated the rotation and translation part. This enabled them to compute the rotation and translation error independent of each other. The isolated relative translation error (RTE) and relative rotation error (RRE) are defined as follows

$$RTE(\tau) = \frac{1}{|\tau|} \sum_{(i,j) \in \tau} \|(p_j \ominus p_i) \ominus (\hat{p}_j \ominus \hat{p}_i)\|_2 \quad (2)$$

$$RRE(\tau) = \frac{1}{|\tau|} \sum_{(i,j) \in \tau} \angle [(p_j \ominus p_i) \ominus (\hat{p}_j \ominus \hat{p}_i)], \quad (3)$$

where the interval τ corresponds to the set of image frames (i, j) that cover a specific length in the trajectory and p_i and \hat{p}_i are

the estimated and ground truth poses, respectively. The symbol \ominus denotes the inverse compositional operator explained in [28] and $\angle[\cdot]$ is the rotation angle for the rotation error.

6. Benchmarking

In this section, we discuss the nature of the trajectories planned and traversed during the dataset recording. Furthermore, we provide experimental results of using state-of-the-art visual SLAM methods on the FinForest dataset.

All sequences start from and end at the same location. Each trajectory has been recorded with an intent to tackle different conditions. The first route (W01 and S01), shown in Figs. 9(a–c) and 10(a–c), comprises a short ellipse shaped trajectory that offers two repeated loop closures while traveling in the same direction and a third loop closure from the opposite direction. The terrain is rather harsh and mimics the uneven ground traversed by work machines. The second sequence (S02), shown in Fig. 10(d–f), offers another loop based trajectory for SLAM approaches. Unlike W01 and S01, this path is traveled only once and therefore forming a single closed loop. Moreover, as mentioned before, no winter recordings are available for this trajectory due to route blockage. The remaining sequences are more visual odometry oriented sequences. These sequences do not offer loop closures by traveling in the same direction. However, the same routes are traversed from the opposite direction, hence, offering an opportunity to explore relocalization possibilities while traveling from the opposite direction. The third, fourth and fifth sequence routes offer short, medium and relatively long trajectories for estimating visual odometry. The third route (W03 and S03), shown in Figs. 9(d–f) and 10(g–i), is the shortest of visual odometry sequences and offers the simplest case for testing. The fourth route (W04 and S04), shown in Figs. 9(g–i) and 10(j–l), offers more of an exploration type of trajectory with back and forth driving to mimic investigative movements of work machines. The fifth route (W05 and S05), shown in Figs. 9(j–l) and 10(m–o), is relatively long and provides more of a challenging odometry test course. Two more visual odometry sequences are provided in the winter condition W06 and W07 (see Fig. 9(m–o) and (p–r)) that are recorded in night and dusk time, respectively. In our opinion, sequence 3–7 are helpful for improving the autonomy of heavy work vehicles in such environments. The sequences mimic the movements of heavy machines that are more fixated on the task at hand in an exploratory manner.

It is noteworthy that the area traversed is deliberately kept limited in terms of displacement from the starting point. Unlike urban infrastructure, forest covered routes provide limited chances to record loop closure over large distances. Recording large distances without loop closure does not suit visual SLAM approaches, therefore, we focused on maintaining short distances with more information in terms of frame rate for improved accuracy. Moreover, at the given framerate, the data recorded is significantly high for the route traversed during the recordings.

Among the state-of-the-art visual SLAM implementations that rank high in the KITTI benchmarking suite [29], we chose ORBSLAM2 [4] and Stereo-Parallel Tracking and Mapping (S-PTAM) [30]. These studies provide open-source implementation of a stereo based visual SLAM method which facilitates the testing phase of our study. It is important to note that both ORBSLAM2 and S-PTAM are used in their standalone mode in order to process all the frames. S-PTAM in specific was not able to process all the incoming frames in its native ROS mode, where it attempts to simulate time-constrained real-time scenario. The implementation was not able to keep up with the incoming frames using the given computational resources. As a consequence, some of the frames were dropped in the ROS mode. To provide a fair and

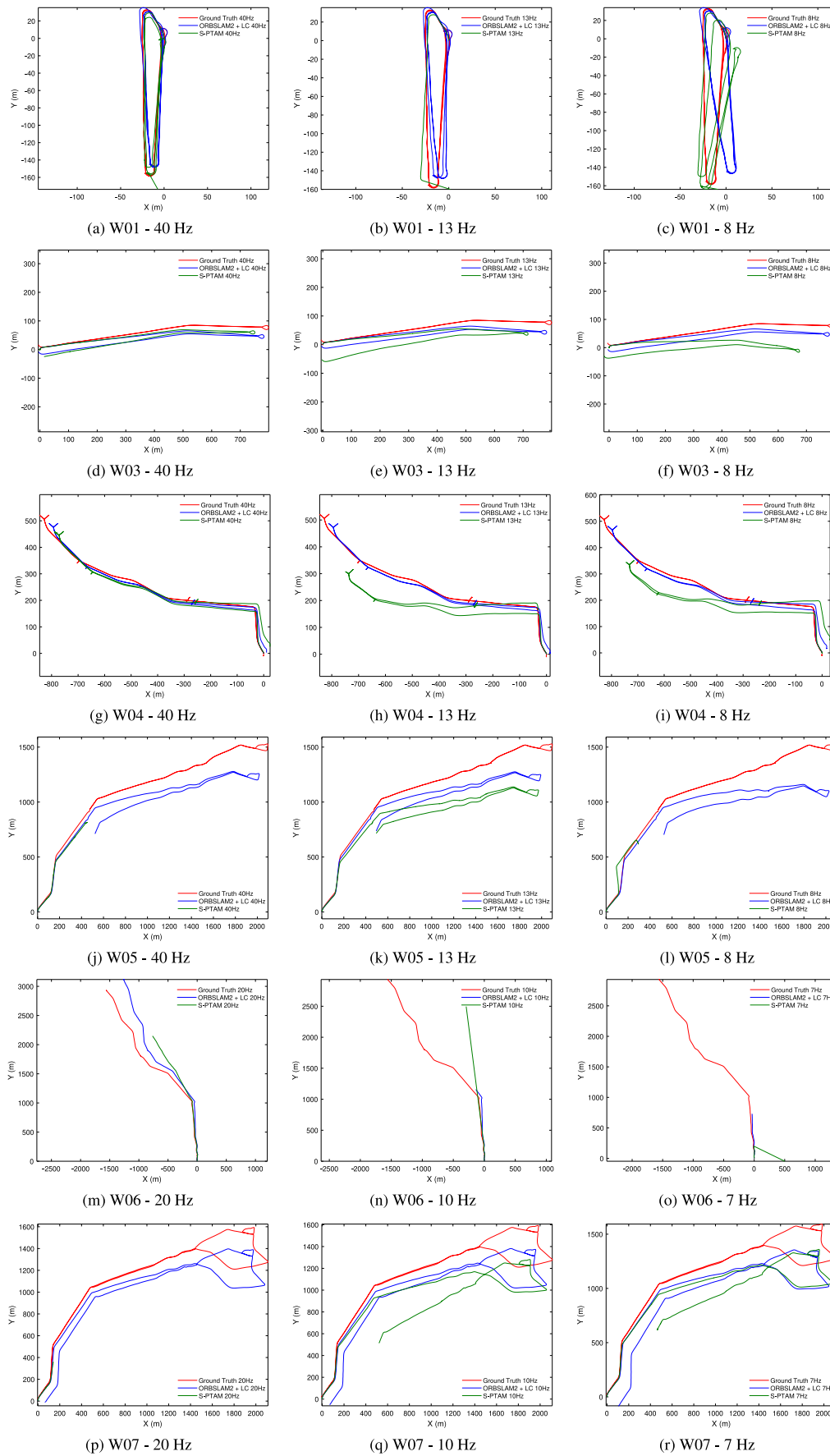


Fig. 9. Estimated trajectories plotted against the ground truth for the winter sequences in FinnForest dataset.

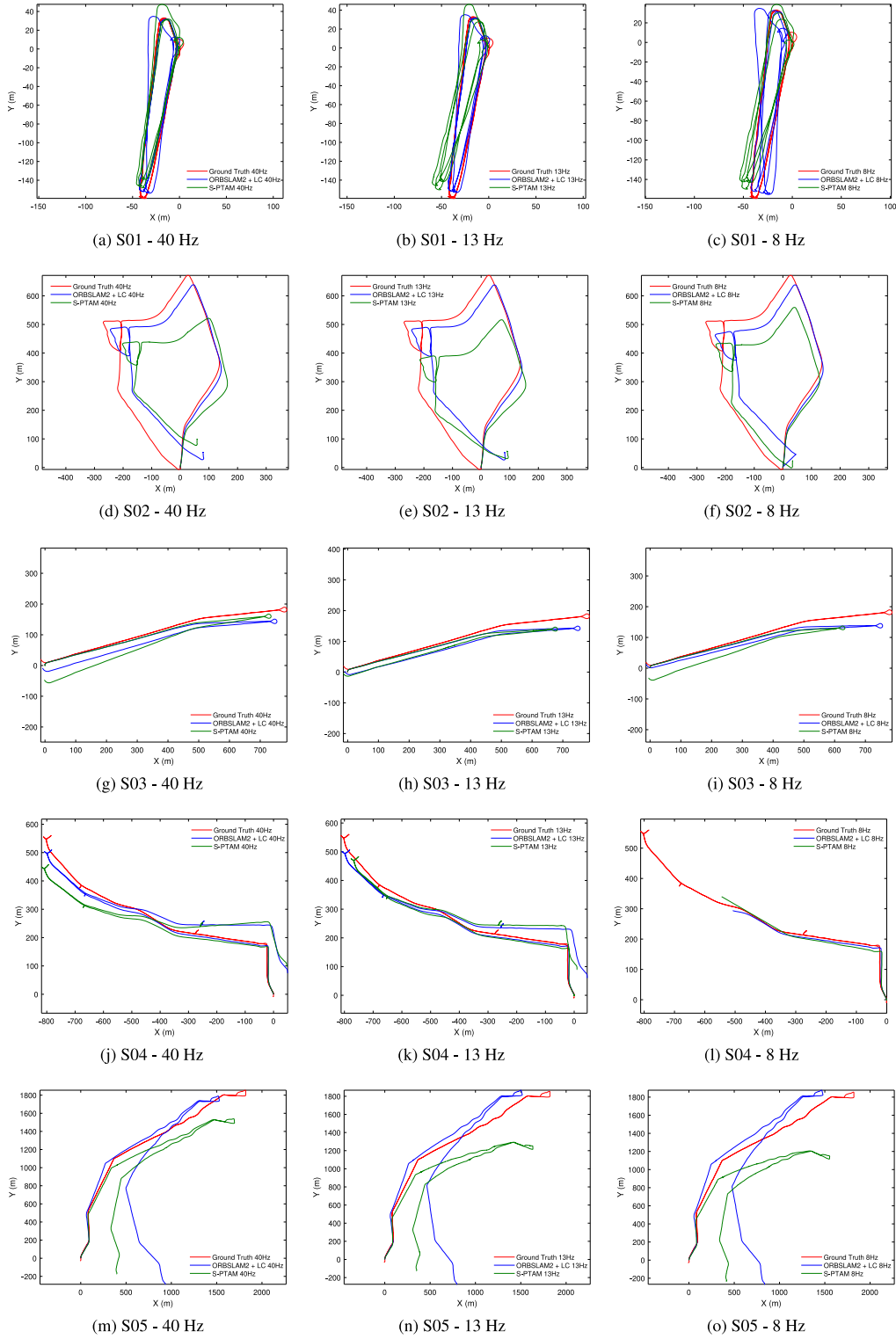


Fig. 10. Estimated trajectories plotted against the ground truth for the summer sequences in FinnForest dataset.

thorough comparison, we provide the results of both methods in their standalone mode with no time constraints for processing. In addition, S-PTAM was used without the loop closure capability due to compatibility issues of the implementation with new versions of dependencies. Except for the sequences with loops, the remaining majority visual odometry sequences should not be affected.

Nonetheless, ORB-SLAM2 and S-PTAM yield excellent results in a typical structured urban environment. These methods have been extensively tested in urban and indoors settings over KITTI, EuRoC, and Level 7 block-set datasets [4,30].

The results obtained with the aforementioned implementations over FinnForest dataset are plotted against the ground truth in Figs. 9 and 10 for all the sequences recorded with the

forward facing stereo pair (C_2 - C_3). Thorough quantitative result is tabulated in Tables 2 and 3. For all these experiments, a standard laptop with an Intel Core i7 @ 1.90 GHz processor and 32 GB RAM was used.

The primary aim of testing the dataset with state-of-the-art method is to educate the readers about the challenges provided by the dataset. Large drift and scale errors are observed for the visual odometry sequences, compared to the sequences with loops, in spite of short distances being covered. We will discuss the results obtained from experimentation in more detail in the next section.

7. Discussion

In this section, we discuss the experimental results using the new dataset and state our observations. Our remarks are intended to aid further research and experimentation with the given data.

7.1. Feature tracking in FinnForest

The forest environment provides unique challenges for tracking features. Due to self-similar and repetitive patterns, extracting correct matches and maintaining tracking with a low number of feature points is tricky. To avoid any obvious obstacles towards tracking, we recorded the data at low driving speeds around 25–30 km/h and low exposure time for image acquisition to avoid motion blur. Following the recommendation in [21], we include the skyline in the scene which is expected to be useful for navigation and augments to reliable features for matching. Furthermore, the forest view near the skyline significantly adds to the rotation accuracy (especially yaw and pitch) by providing features that are far away from the camera.

In most of the testing cases, we used 2000 feature points to track with ORB-SLAM2 and 1000 feature points with S-PTAM. The number of feature points selected was a compromise between the image resolution and memory management of the implementation. However, we observed that the selected parameters were suitable for testing most of the sequences and provided sufficient cross over candidates between frames for matching. During experimentation, we observed that S-PTAM required more tuning of the parameters compared to ORB-SLAM2, in which they were kept mostly the same for all experiments. ORB-SLAM2 uses ORB features which are both faster and more robust (due to rotation invariance) compared to features used by S-PTAM. S-PTAM uses the GFIT feature and BRIEF descriptors for matching. The descriptor is not invariant to rotation and as a result, the implementation requires parameter adjustment for various sequences of FinnForest dataset to maintain tracking on the parts of the route with harsh terrain.

As mentioned earlier, unlike urban routes, the path traversed while recording the dataset is a rough terrain. The combined effect of erratic motion, speed and data sampling introduce challenges for testing. It is apparent from the experimental results that the highest errors are observed in the visual odometry sequences while the errors are reduced and distributed in the sequences where loop closure has been achieved.

Effect of data sampling on tracking: The sampling of the dataset at lower rates is intended to facilitate testing and investigate a suitable data rate for real cases. Though lower frames per second (fps) are advantageous for testing purposes, information processing at lower fps can considerably compromise the visual odometry pipeline during real field operation. To exemplify the behavior, we take the experimental results of ORBSLAM2 on S02 at 8 Hz. ORBSLAM2 fails to continue its tracking of feature points when the vehicle hits a pothole and the scene observes a sharp motion. It is important to note, that ORBSLAM2 successfully

completed the same test sequence at higher frame rates (13 and 40 Hz). For further investigation, we significantly changed the parameters by increasing the feature points to 5000 and varying the FAST feature threshold between 4 and 18. However, the result remained the same. Surprisingly, S-PTAM successfully completed the test sequence S02 at 8 Hz when the feature points to detect were set to 1500.

On the contrary, ORBSLAM2 was able to handle a similar situation in W01 at 8 Hz with loosened parameters while S-PTAM failed to continue the tracking. However, none of the implementations were able to successfully complete the sequences W05, W06, and S04. A similar effect was observed in W07 at 7 Hz and the parameters were loosened again. This time ORB-SLAM was able to successfully process the entire sequence while S-PTAM failed.

Effect of motion on tracking: In some cases, the erratic motion due to terrain in combination with the scene is already too much even at a higher frame rate. In the case of W01, we observed that S-PTAM fails to complete the sequence at all sampling rates. The implementation fails while locally adjusting the poses that lie in the range where the sharp movements are observed. On the other hand, ORBSLAM2 was able to process the sequence with relative ease at 40 and 13 Hz without fine-tuning of the parameters. However, at 8 Hz the feature points used for tracking were increased, and the feature threshold lowered to maintain tracking even with ORBSLAM2.

A different cause is expected to be affecting S-PTAM while processing W05 at 8 Hz. The tracking failure occurs when the vehicle slows down to a momentary stationary state and restarts motion. We believe the source of the issue is the predictive feature search that fails to find matches. In both S-PTAM and ORBSLAM2, a motion model is used to predict the position of the map points on the latest image frame and find matches in the small neighborhood for tracking. In case, if the feature matches are not found in the small predicted neighborhood, ORBSLAM2 expands the search window as a fallback option. On the other hand, we believe, S-PTAM relies only on the decaying velocity model and does not expand its search neighborhood as a fallback option. As a result, a sudden change in velocity at a lower frame rate affects the tracking of feature points. This phenomenon is aggravated by the sub sampling since the same behavior is handled by S-PTAM at sampling of 13 Hz but fails at 8 Hz when change is more abrupt. By requesting more feature points to be detected in the new image frame, we can avoid the tracking failure altogether, however, poor matches with the map then lead to convergence issue in the local bundle adjustment step of S-PTAM.

Effect of illumination on tracking: The dataset includes various opportunities to test the robustness of visual SLAM implementation towards tracking and pose estimation in a scene with varying illumination. The notable opportunities regarding illumination change are provided by W07, S04, and W06. In W07, we observe gradual illumination change as it gets darker. The sequence was recorded at the dusk time and the illumination changes drastically between the start and end of the sequence. ORBSLAM2 did not face any issue in terms of tracking feature points in this sequence, however, S-PTAM faced considerable problems to maintain tracking at all sampling rates. S-PTAM also failed tracking at sampled data of 13 Hz, however, we have included the results since the failure point was close to the end of the sequence.

On the other hand, a more rapid change is observed in illumination due to direct sunlight in the sequence S04. At a sampling rate of 40 and 13 Hz, both ORBSLAM2 and S-PTAM can successfully process the sequence. However, at 8 Hz they fail at different points. The S-PTAM fails directly due to overexposure and flare observed in the scene while ORBSLAM2 fails due to fast erratic motion following the over-exposed scene in the recording.

Table 2

Quantitative results of ORBSLAM2 for the FinnForest dataset at different sampling rates.

Data Sampling	40/20 Hz			13/10 Hz			8/7 Hz		
Seq. No	ATE (rmse)	RTE (%)	RRE (deg/m)	ATE (rmse)	RTE (%)	RRE (deg/m)	ATE (rmse)	RTE (%)	RRE (deg/m)
W01	3.35	2.1785	0.00014197	3.6738	2.3016	0.00019584	3.4914	2.4092	0.00021607
W03	12.266	9.1805	0.00012107	12.025	9.2299	0.00013344	12.249	9.1962	0.0001253
W04	17.421	7.7753	9.7778e-05	17.244	7.8332	9.935e-05	20.666	7.6746	0.00011482
W05	55.422	9.2678	0.0001298	56.323	9.4365	0.00013865	75.715	9.7977	0.00022451
W06 ^a	21.789	32.14	0.00020608	TL	TL	TL	TL	TL	TL
W07 ^a	37.933	7.2208	0.00011185	34.324	7.2193	0.00013786	48.88	7.2107	0.00016175
S01	4.3677	1.9672	0.00022474	3.8189	1.917	0.00019894	6.2793	2.1462	0.00027508
S02	26.132	4.2061	0.00017796	26.874	4.2181	0.0001728	TL	TL	TL
S03	12.633	5.873	0.00020877	10.986	5.6022	0.00018197	9.8899	5.5459	0.00018258
S04	30.053	5.5827	0.0001988	26.825	5.4608	0.00018299	TL	TL	TL
S05	228.88	9.4575	0.00025169	191.52	8.8165	0.00020505	200.81	8.9426	0.00021338

^aIndicates that the data is subsampled at 20/10/7 Hz.

TL: Tracking lost.

Table 3

Quantitative results of S-PTAM for the FinnForest dataset at different sampling rates.

Data Sampling	40/20 Hz			13/10 Hz			8/7 Hz		
Seq. No	ATE (rmse)	RTE (%)	RRE (deg/m)	ATE (rmse)	RTE (%)	RRE (deg/m)	ATE (rmse)	RTE (%)	RRE (deg/m)
W01	TL	TL	TL	TL	TL	TL	TL	TL	TL
W03	19.709	10.166	0.00011828	27.663	12.63	0.0004809	28.369	14.819	0.00063508
W04	25.852	9.4934	0.00014839	45.091	14.9	0.00071498	48.944	14.914	0.00073208
W05	TL	TL	TL	79.774	11.312	0.00011181	TL	TL	TL
W06 ^a	TL	TL	TL	TL	TL	TL	TL	TL	TL
W07 ^a	TL	TL	TL	102.54	8.319	0.00019895	TL	TL	TL
S01	7.3247	2.883	0.00018821	9.4022	4.0914	0.00066569	8.652	3.9342	0.00030787
S02	34.391	9.2735	0.0005317	44.68	11.63	0.00061402	34.752	9.2786	0.00020271
S03	21.779	7.0644	0.00025365	31.418	11.105	0.00025333	47.392	14.82	0.00031883
S04	31.891	7.1297	0.00023556	39.749	9.703	0.00019259	TL	TL	TL
S05	130.41	10.182	0.00022272	171.55	14.517	0.00032586	201.65	17.9	0.00038022

^aIndicates that the data is subsampled at 20/10/7 Hz.

TL: Tracking lost.

The Night sequence, W06, is especially challenging for both implementations. Neither of the implementations could process the sequences under normal parameter settings. ORBSLAM2 was able to process the sequence at 40 Hz with relaxed parameters after the FAST feature threshold was reduced to 4 to avoid losing the track of features. S-PTAM is not able to process the W06 sequence at any sampling rate. Even after the feature threshold is reduced, S-PTAM fails to converge at local bundle adjustment. This is expected since the scene in view is limited to a few meters of the snow-covered road. As a result, the poses estimated do not agree over a longer duration and fail to converge at bundle adjustment.

7.2. Loop closure

The dataset provides three sequences with loop closure opportunities. Among these, S01 and W01 repeat the same route twice in one direction and the third time in the opposite direction. This means that ORBSLAM2 can identify the loop closure opportunity at any time of the second lap of the drive. During experimentation we observed that ORBSLAM2 successfully closes the loop and distributes the errors for the aforementioned sequences. In contrast, ORBSLAM2 fails to close the loop for the sequence S02, even though, enough overlap of the start and end scenes is provided. Oddly, ORBSLAM2 can re-localize itself at the end of the sequence S02 that is processed at 8 Hz after losing track of the feature points. A closure can be observed due to re-localization in Fig. 10(f) in the trajectory estimated by ORBSLAM2. We believe that sparser keyframes formed at 8 Hz provided more decisive information compared to the same sequence at higher fps, where the relocalization was not observed.

7.3. Drift

A drift in scale and rotation can be observed in the estimations provided by ORBSLAM2 and S-PTAM for all of the visual odometry sequences. This effect of drift becomes stronger as the sample rate drops down from 40 to 8 Hz. The effect is most apparent in S03 and W03 (see Fig. 9(d–f) and (g–i)).

7.4. Seasonal effect

Seasonal changes have an apparent effect on various aspects of this dataset. As discussed earlier, the ground truth accuracy reduced in the summertime compared to the wintertime due to considerably higher foliage effect in the summer. An added challenge from the perspective of recording was that, while traversing the forest, different parts of the forest provided different levels of shade from the sun due to the density of the trees in that specific part. This created a challenge to avoid over or underexposure of the scenes since we used a fixed aperture. These effects are more obvious in the sequence S04.

The winter sequences, on the other hand, were adequately exposed since most of the recordings are in overcast. In addition, there was enough texture on the ground due to tire tracks in the snow. ORBSLAM2 handled tracking very well with evenly distributed points on the snow-covered ground. S-PTAM focused more on the obvious texture from the trees. Most of the feature points from the snow-covered road are discarded by S-PTAM as false matches.

7.5. Effect of ground truth precision

It is important to note that the experimentation is independent of the precision level of the ground truth position since

the test algorithms did not use the IMU and GNSS information. However, the effect of the ground truth precision indeed has to be considered when comparing the experimental results against the ground truth poses. The ground truth precision for each sequence is shown using Fig. 8 and discussed in Section 4.3. In the context of benchmarking, we can say that we are more confident in the comparison performed in Tables 2 and 3 for the winter sequences (W01–W07) than the summer sequences (S01–S05) since the precision of ground truth position for winter sequences is comparatively higher. Nonetheless, the precision of the ground truth is high enough in both cases for valid analysis of visual odometry/SLAM algorithms.

It is important to remember that the visual odometry/SLAM algorithms may give different responses for the same trajectory recorded under different condition, as we discussed throughout Section 7. Therefore, arguing that one result is better than the other without comparing to the provided ground truth is not an objective conclusion.

8. Summary

In this paper, we have presented a novel dataset that offers a forest-like environment in various light and weather conditions for visual odometry and SLAM systems to process. The dataset provides synchronized and processed image frames from 4 cameras that can be used independently or as stereo pairs. Moreover, raw data is also provided to encourage further examination into the system. We believe this dataset will prove immensely useful towards enlarging the spectrum and diversity of the testing data for autonomous vehicles, especially, autonomous heavy work machines. We hope that this dataset will provide new challenges and inspire exploration of new possibilities for autonomous vehicles/machines.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This Work was partially supported by Business Finland, Finland under the project Spatial Sensing for Machines (SSFM, grant number 1822/31/2016). This work was carried out with the support of Centre for Immersive Visual Technologies (CIVIT) research infrastructure, Tampere University, Finland.

References

- [1] R. Benenson, M. Omran, J. Hosang, B. Schiele, Ten years of pedestrian detection, what have we learned? in: *European Conference on Computer Vision*, Springer, 2014, pp. 613–627.
- [2] A. Geiger, M. Roser, R. Urtasun, Efficient large-scale stereo matching, in: *Asian Conference on Computer Vision*, Springer, 2010, pp. 25–38.
- [3] J. Long, E. Shelhamer, T. Darrell, Fully convolutional networks for semantic segmentation, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 3431–3440.
- [4] R. Mur-Artal, J.D. Tardós, Orb-slam2: An open-source slam system for monocular, stereo, and rgb-d cameras, *IEEE Trans. Robot.* 33 (5) (2017) 1255–1262.
- [5] J. Engel, V. Usenko, D. Cremers, A photometrically calibrated benchmark for monocular visual odometry, 2016, arXiv preprint [arXiv:1607.02555](https://arxiv.org/abs/1607.02555).
- [6] N. Carlevaris-Bianco, A.K. Ushani, R.M. Eustice, University of Michigan North Campus long-term vision and lidar dataset, *Int. J. Robot. Res.* 35 (9) (2016) 1023–1035.
- [7] A.Z. Zhu, D. Thakur, T. Özarslan, B. Pfrommer, V. Kumar, K. Daniilidis, The multivehicle stereo event camera dataset: An event camera dataset for 3D perception, *IEEE Robot. Autom. Lett.* 3 (3) (2018) 2032–2039.
- [8] H. Jung, Y. Oto, O.M. Mozos, Y. Iwashita, R. Kurazume, Multi-modal panoramic 3D outdoor datasets for place categorization, in: *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS, IEEE, 2016*, pp. 4545–4550.
- [9] M. Cordts, M. Omran, S. Ramos, T. Rehfeld, M. Enzweiler, R. Benenson, U. Franke, S. Roth, 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.
- [10] G. Pandey, J.R. McBride, R.M. Eustice, Ford campus vision and lidar data set, *Int. J. Robot. Res.* 30 (13) (2011) 1543–1552.
- [11] A. Geiger, P. Lenz, C. Stiller, R. Urtasun, Vision meets robotics: The KITTI dataset, *Int. J. Robot. Res.* 32 (11) (2013) 1231–1237.
- [12] W. Maddern, G. Pascoe, C. Linegar, P. Newman, 1 year, 1000 km: The Oxford RobotCar dataset, *Int. J. Robot. Res.* 36 (1) (2017) 3–15.
- [13] Y. Choi, N. Kim, S. Hwang, K. Park, J.S. Yoon, K. An, I.S. Kweon, KAIST multi-spectral day/night data set for autonomous and assisted driving, *IEEE Trans. Intell. Transp. Syst.* 19 (3) (2018) 934–948.
- [14] J. Jeong, Y. Cho, Y.-S. Shin, H. Roh, A. Kim, Complex urban dataset with multi-level sensors from highly diverse urban environments, *Int. J. Robot. Res.* 38 (6) (2019) 642–657.
- [15] M. Ferrera, J. Moras, P. Trounev-Peloux, V. Creuze, D. Dégez, The aqualoc dataset: Towards real-time underwater localization from a visual-inertial-pressure acquisition system, 2018, arXiv preprint [arXiv:1809.07076](https://arxiv.org/abs/1809.07076).
- [16] M. Miller, S.-J. Chung, S. Hutchinson, The visual-inertial canoe dataset, *Int. J. Robot. Res.* 37 (1) (2018) 13–20.
- [17] A. Mallios, E. Vidal, R. Campos, M. Carreras, Underwater caves sonar data set, *Int. J. Robot. Res.* 36 (12) (2017) 1247–1251.
- [18] K. Leung, D. Lühr, H. Houshiar, F. Inostroza, D. Borrmann, M. Adams, A. Nüchter, J. Ruiz del Solar, Chilean underground mine dataset, *Int. J. Robot. Res.* 36 (1) (2017) 16–23.
- [19] P. Furgale, P. Carle, J. Enright, T.D. Barfoot, The Devon Island rover navigation dataset, *Int. J. Robot. Res.* 31 (6) (2012) 707–713.
- [20] J. Bruce, J. Wawerla, R. Vaughan, The SFU mountain dataset: Semi-structured woodland trails under changing environmental conditions, in: *IEEE Int. Conf. on Robotics and Automation 2015, Workshop on Visual Place Recognition in Changing Environments*, 2015.
- [21] J. Garforth, B. Webb, Visual appearance analysis of forest scenes for monocular slam, in: *2019 International Conference on Robotics and Automation, ICRA, IEEE, 2019*, pp. 1794–1800.
- [22] Z. Zhang, A flexible new technique for camera calibration, *IEEE Trans. Pattern Anal. Mach. Intell.* 22 (2000).
- [23] P. Furgale, J. Rehder, R. Siegwart, Unified temporal and spatial calibration for multi-sensor systems, in: *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems, IEEE, 2013*, pp. 1280–1286.
- [24] K.-D. Park, J. Won, The foliage effect on the height time series from permanent GPS stations, *Earth Planets Space* 62 (11) (2010) 849–856.
- [25] J. Goldhirsh, W.J. Vogel, Handbook of Propagation Effects for Vehicular and Personal Mobile Satellite Systems, Vol. 1274, NASA Reference Publication, 1998, pp. 40–67.
- [26] S. Umeyama, Least-squares estimation of transformation parameters between two point patterns, *IEEE Trans. Pattern Anal. Mach. Intell.* 13 (4) (1991) 376–380.
- [27] M. Salas, Y. Latif, I.D. Reid, J. Montiel, Trajectory alignment and evaluation in SLAM: Horn's method vs alignment on the manifold, in: *Robotics: Science and Systems Workshop: The Problem of Mobile Sensors*, 2015, pp. 1–3.
- [28] R. Kümmerle, B. Steder, C. Dornhege, M. Ruhnke, G. Grisetti, C. Stachniss, A. Klein, On measuring the accuracy of SLAM algorithms, *Auton. Robots* 27 (4) (2009) 387.
- [29] The KITTI vision benchmark suite 2020, 2020, http://www.cvlibs.net/datasets/kitti/eval_odometry.php, Cvlibs.net.
- [30] T. Pire, T. Fischer, G. Castro, P. De Cristóforis, J. Civera, J. Jacobo Berles, S-PTAM: Stereo parallel tracking and mapping, *Robot. Autonom. Syst. (RAS)* 93 (2017) 27–42, [http://dx.doi.org/10.1016/j.robot.2017.03.019](https://doi.org/10.1016/j.robot.2017.03.019).



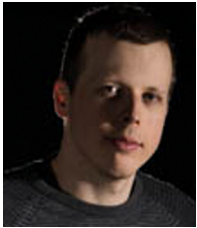
Ihtisham Ali received his B.Sc. in Mechatronics Engineering from the University of Engineering and Technology, Pakistan (2014) and his M.Sc. in Automation Engineering from Tampere University, Finland (2017). Currently, he is a doctoral researcher in 3D Media Group at Tampere University. He has worked on several industrial projects pertaining to machine automation using visual cues. His research interest is focused on computer vision and robotics specifically object pose estimation, 3D reconstruction, and visual SLAM.



Ahmed Durmush received the B.Sc. degree in Control Systems Engineering from Istanbul Technical University (2012). He is developing the systems required by the 3D Media group at CIVIT for light field capture and reconstruction and other research areas.



Sari Peltonen received the M.Sc. degree in mathematics from the University of Tampere, Finland, in 1996. She received the Ph.D. degree in signal processing from the Tampere University of Technology in 2000. She is currently University Lecturer in signal processing at the Unit of Computing Sciences in Tampere University. Her research interests include robust estimation, image processing and tomographic image reconstruction.



Olli Suominen graduated with both B.Sc. and M.Sc. (Tech) in Information Technology from Tampere University of Technology (2011/2012) with a major in Signal Processing. Since then, he has been a Ph.D. student in 3D Media Group at the Laboratory of Signal Processing at TUT and managing the construction and development of [Centre for Immersive Visual Technologies](#). After starting from the B.Sc. thesis using only one camera (Depth Image Based Rendering) and an M.Sc. thesis using two cameras (Stereo Depth Estimation) he has now scaled up to 40 cameras for

the Ph.D. with research interests in multi-camera systems, 3D reconstruction, multimodal sensor fusion, SLAM and light field capture. He currently focuses on applications in heavy mobile work machines, leading several industry driven research projects and developing relations with the industry.



Jussi Collin is CEO of Nordic Inertial (JC Inertial Oy) and adjunct professor at Tampere University. He received the M.Sc. and Dr.Tech. degrees from the Tampere University of Technology, Tampere, Finland, in 2001 and 2006, respectively, specializing in inertial navigation algorithms. His research interests are in modern machine learning methods and their industrial applications in the field of inertial sensing.



Dr.Tech Jari Yli-Hietanen has background in Signal Processing research. He has worked with various topics including robotic cognition, machine vision and natural language processing. Currently, he is developing IT services for research support at Tampere University.



Atanas Gotchev received his M.Sc. degrees in radio and television engineering (1990) and applied mathematics (1992), his Ph.D. degree in telecommunications (1996) from the Technical University of Sofia, and the D.Sc.(Tech.) degree in information technologies from the Tampere University of Technology (2003). He is a Professor of Signal Processing and Director of the Centre for Immersive Visual Technologies at Tampere University. His recent work concentrates on the algorithms for multi-sensor 3-D scene capture, transform-domain light-field reconstruction, and

Fourier analysis of 3-D displays.