

OPEN URBAN AND FOREST DATASETS FROM A HIGH-PERFORMANCE MOBILE MAPPING BACKPACK – A CONTRIBUTION FOR ADVANCING THE CREATION OF DIGITAL CITY TWINS

S. Blaser¹, J. Meyer¹, S. Nebiker¹

¹ Institute of Geomatics, FHNW University of Applied Sciences and Arts Northwestern Switzerland, Muttenz, Switzerland - (stefan.blaser, jonas.meyer, stephan.nebiker)@fhnw.ch

Commission I

KEY WORDS: Mobile Mapping, Urban, Forest, Open Dataset, Localization, Georeferencing, SfM, SLAM, Smart City, Digital Twin

ABSTRACT:

With this contribution, we describe and publish two high-quality street-level datasets, captured with a portable high-performance Mobile Mapping System (MMS). The datasets will be freely available for scientific use. Both datasets, from a city centre and a forest represent area-wide street-level reality captures which can be used e.g. for establishing cloud-based frameworks for infrastructure management as well as for smart city and forestry applications. The quality of these data sets has been thoroughly evaluated and demonstrated. For example, georeferencing accuracies in the centimetre range using these datasets in combination with image-based georeferencing have been achieved. Both high-quality multi sensor system street-level datasets are suitable for evaluating and improving methods for multiple tasks related to high-precision 3D reality capture and the creation of digital twins. Potential applications range from localization and georeferencing, dense image matching and 3D reconstruction to combined methods such as simultaneous localization and mapping and structure-from-motion as well as classification and scene interpretation. Our dataset is available online at: <https://www.fhnw.ch/habg/bimage-datasets>

1. INTRODUCTION

The ongoing progress in digitalization leads to massive transformations and innovations in infrastructure management. Multiple domains require detailed and accurate 3D data for creating and updating smart city and digital twin solutions. Mobile mapping systems (MMS) hold the potential to provide such data in a cost-effective manner. From the first stereo image-based MMS in the early 1990s (Novak, 1991; Schwarz et al., 1993) they evolved into multi-stereo camera and panoramic camera configurations with an almost complete coverage (Meilland et al., 2015; Blaser et al., 2017).

Nebiker et al. (2015) discuss advantages of image-based Mobile Mapping (MM) over the widely distributed light detection and ranging (LiDAR)-based MM in terms of temporal coherence in the acquisition and of information density. Moreover, image-based MM allow to create 3D geospatial image spaces, which can be used e.g. for infrastructure management using cloud-based web applications. Blaser et al. (2017) show the great potential of automatically creating detailed 3D city models from street-level using imagery. As vehicle-based MMS have become well established, portable MMS have entered the market in recent years. Lehtola et al. (2017) provide a comparison of numerous state-of-the-art LiDAR-based indoor MMS. Blaser et al. (2018) present the development of a portable image-based indoor MMS and provide accuracy analysis in indoor environment with promising results within the centimetre range. Blaser et al. (2020) extended the portable MMS with a tactical grade inertial measurement unit (IMU) for indoor and outdoor use. They conducted performance evaluations using three independent georeferencing methods in challenging outdoor test sites not accessible to vehicles and achieved accuracies in the centimetre range. However, all georeferencing methods showed outliers. Combining various georeferencing methods or coupling multiple sensor data could further improve the accuracy as well as the reliability.

In recent decades and years, we also witnessed a paradigm shift towards open science. Open access publications, open source software as well as open datasets promoted transparency and comparability in science. Related sciences like computer vision or robotics experienced enormous progress, which was accelerated or even made possible thanks to the open science philosophy. Thus, we consider that our challenging datasets could help to accelerate the development of novel methods in the field of mobile 3D reality capture and smart city.

With this contribution, we

- publish two high-quality street-level datasets captured with a portable high-performance MMS from challenging environments in a city centre as well as in a forest for scientific use.
- describe our MMS and the resulting raw and pre-processed data and our accurate overall system calibration so that the data provided can be fully utilized.
- describe both test sites and their associated datasets.
- show our initial research and discuss the potential and possible applications for the datasets provided.

First, we discuss related work and related open datasets that already exist. Second, we describe our MMS, provide our overall system calibration, and specify resulting raw and pre-processed data. Third, we describe our test sites as well as our published datasets. Finally, we show initial research and point out its potential and show possible applications.

2. RELATED WORK

In autonomous driving, there exist numerous benchmark datasets (e.g. KITTI Vision Benchmark Suite, Waymo Open Dataset, ApolloScape Dataset, etc.). Choi (2019) provide a list of the most recent benchmark datasets and dataset collections for robotics. The KITTI Vision Benchmark Suite (Geiger et al., 2012) provides benchmark data and leader boards for numerous applica-

tions in autonomous driving. However, in these fields high frame rates and low latency are of greater importance than image quality, precise sensor synchronisation and exact overall calibration – aspects which are important for mapping applications.

Most of the mobile mapping datasets published to date base on LiDAR point clouds and focus on a specific scientific issue. Tan et al. (2020) provide the Toronto-3D dataset for semantic segmentation of urban roadways. Their mobile lidar system (MLS) is precisely synchronized, so that they can provide coloured point clouds with additional precise global navigation satellite system (GNSS) time stamps for each LiDAR point. By contrast, Wang et al. (2019) introduce the ISPRS Benchmark on Multisensory Indoor Mapping and Positioning (MIMAP) using data from their self-developed Indoor MM backpack *XBeibao* (Wen et al., 2016). They use the Network Time Protocol (NTP) for sensor synchronization. However, they only synchronize the start time of the data acquisition and interpolate the subsequent data using the frame rate. Since the cameras and the smartphone were connected over Wi-Fi, significant time delays are to be expected. MIMAP consists of three benchmark data sets: Indoor Simultaneous Localization and Mapping (SLAM), Building Information Modelling (BIM) feature extraction and Indoor positioning.

Khoshelham et al. (2017) present the ISPRS Benchmark on Indoor Modeling. This benchmark includes five different indoor point cloud datasets without any sensor raw data and trajectory information. Each point cloud was collected with a different commercial indoor MMS. The benchmark includes the comparison of derived BIM models to their corresponding reference models.

Nex et al. (2015) introduce and describe the ISPRS Benchmark for Multi-Platform Photogrammetry. With respect to research activities in dense image matching (Cavegn et al., 2014), they proposed two benchmark areas ‘City center’ and ‘Zeche Zollern’ located in Dortmund, Germany. Both areas were captured with cameras and LiDAR scanners from different perspectives and ranges: from terrestrial and short range up to Unmanned Aerial Vehicle (UAV) and aircraft based.

3. SYSTEM DESCRIPTION

We captured both datasets with our portable *BIMAGE Backpack* MMS. Since our system is a self-developed non-commercial research prototype, there are no restrictions in accessing and providing raw sensor data and in describing the system design and configuration in detail. Something, which is typically not available from commercial systems due to intellectual property issues.

This chapter briefly introduces the non-commercial prototypical *BIMAGE Backpack* MMS, provides the overall system calibration parameters, and describes the output data as well as the data structure.

3.1 Hardware

Our portable MMS *BIMAGE Backpack* includes state-of-the-art and high-end sensors, such as the GNSS- and IMU-based navigation unit *NovAtel SPAN CPT7* with tactical grade performance, two multi-beam LiDAR scanners *Velodyne VLP-16* as well as the multi-head panoramic camera *FLIR Ladybug 5* (see Figure 1). Blaser et al. (2020) provide a detailed description of all components used.

Precise sensor synchronization is one of the key features. The internal clocks of both LiDAR scanners are synchronized with an electronic Pulse Per Second (PPS) to the reference clock of the

navigation unit. This allows to assign a precise acquisition timestamp to each LiDAR point. Furthermore, each panoramic camera trigger sends an electronic pulse to the navigation unit, which generates a precise reference time stamp. Consequently, each image is assigned with a precise acquisition timestamp.

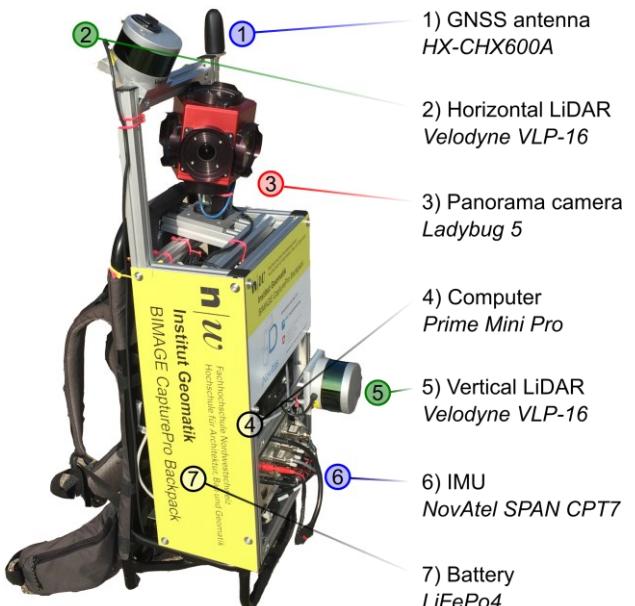


Figure 1. Portable high-performance MMS *BIMAGE Backpack* with multi sensor configuration. (Blaser et al., 2020)

3.2 Coordinate Frames and Overall System Calibration Parameters

The overall system calibration consists of a) the boresight alignment (BA), b) the relative orientation (RO), and c) the interior orientation (IO).

Blaser et al. (2018) describe the BA, RO and IO panoramic camera calibration procedure and the results in more detail. In this contribution, we mainly provide the coordinate frame definition and the calibration parameters required for further data processing and data evaluation.

The BA describes lever-arm and misalignment from a specific sensor coordinate frame to the body frame. In case of a multi-camera configuration, the RO describes the lever-arm and the misalignment from a sub-ordered camera coordinate frame to the principal camera coordinate frame. Both BA and RO mathematically describe rigid body transformations with six parameters

$${}^b\mathbf{x}_i = {}^b\mathbf{T}_a({}^b\mathbf{R}_a, {}^b\mathbf{t}_a){}^a\mathbf{x}_i = \begin{bmatrix} {}^b\mathbf{R}_a & {}^b\mathbf{t}_a \\ \mathbf{0}^T & 1 \end{bmatrix} {}^a\mathbf{x}_i, \quad (1)$$

where the transformation ${}^b\mathbf{T}_a$ transforms a vector with homogeneous coordinates ${}^a\mathbf{x}_i$ from the coordinate frame a into the coordinate frame b ${}^b\mathbf{x}_i$ using the rotation matrix ${}^b\mathbf{R}_a$ and the translation vector ${}^b\mathbf{t}_a$.

In our MMS, we use right-handed coordinate frames and Euler angles about rotated axes

$$\mathbf{R}_a = \mathbf{R}_\omega \mathbf{R}_\phi \mathbf{R}_\kappa. \quad (2)$$

Figure 2 shows the orientation of the sensor coordinate frames. The body frame \mathbf{b} corresponds to the navigation coordinate frame

of the IMU. Its y-axis points in walking direction, the x-axis to the right and its z-axis points upwards. The panoramic camera sensor coordinate frame corresponds with the camera coordinate system of the principal camera head **cam0**. For cameras, we use the photogrammetric camera coordinate system definition with origin in the projection centre, where the x-axis points to the right, the y-axis points upwards and the z-axis points backwards to the viewing direction. Thus, principal camera head **cam0** of the panoramic camera points backwards to the moving direction (see Figure 2).

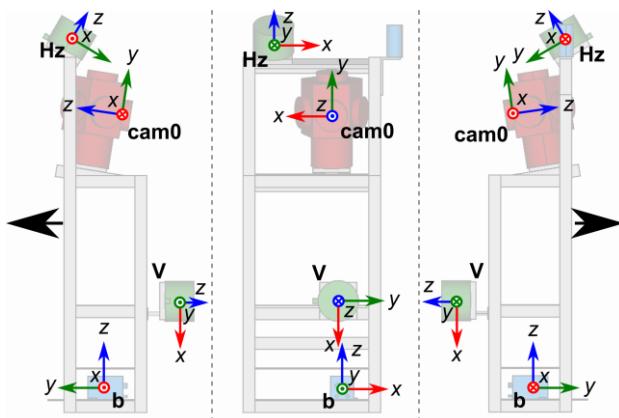


Figure 2. Coordinate frame definition outlines of the *BIMAGE Backpack* view from the left (left), back view (centre) and view from the right (right). The big black arrows (left and right) mark the moving direction. Bold labels **b** (body frame), **Hz** (horizontal LiDAR), **V** (vertical LiDAR) and **cam0** (panoramic camera) represent the coordinate frames and italic labels mark the coordinate axis. Point symbols in the coordinate frame origin represent forward-pointing axes, while cross symbols mark backward-pointing axes.

Sensor	Leverarm [m] $[x \ y \ z]^T$	Misalignment [deg] $[\omega \ \phi \ \kappa]^T$
Pano cam0	- 0.028	- 78.902755
	- 0.033	0.493794
	0.685	179.001458
LiDAR Hz	- 0.174	31.410335
	0.075	- 0.011460
	0.882	- 179.513889
LiDAR V	- 0.005	89.509669
	- 0.208	0.100370
	0.216	- 90.067824

Table 1. Boresight alignment parameters (BAs) of the *BIMAGE Backpack*. The BAs start from the sensors and point to the body frame.

Table 1 shows the BAs of the panoramic camera and of both LiDAR scanners, while Table 2 lists the ROs of the panoramic camera. To calculate camera poses, BAs and ROs can be concatenated e.g. as follows:

$${}^w\mathbf{T}_{cam\ n} = {}^w\mathbf{T}_b \ {}^b\mathbf{T}_{cam0} \ {}^{cam0}\mathbf{T}_{cam\ n}, \quad (3)$$

while ${}^b\mathbf{T}_{cam0}$ is the BA of the panoramic camera and ${}^{cam0}\mathbf{T}_{cam\ n}$ is the RO of a specific camera head, when we assume that the body frame pose ${}^w\mathbf{T}_b$ is given.

Sensor	Leverarm [m] $[x \ y \ z]^T$	Misalignment [deg] $[\omega \ \phi \ \kappa]^T$
Pano cam1	0.059	- 0.646642
	0.000	- 72.103993
	0.042	- 0.269886
Pano cam2	0.036	0.538770
	0.000	- 144.043410
	0.109	0.401385
Pano cam3	- 0.035	0.033551
	0.000	- 216.020014
	0.111	0.245717
Pano cam4	- 0.057	0.538346
	0.000	- 288.221386
	0.042	- 0.204128
Pano cam5	- 0.001	90.079865
	0.075	0.087070
	0.060	180.677179

Table 2. Relative orientation parameters (ROs) of the *Ladybug 5* panoramic camera. The ROs start from the mentioned camera heads (sensor) and point to the reference camera head (cam0).

By contrast, the IO describes the transformation from the image coordinate frame to the camera coordinate frame. For this purpose, we use the equidistant camera model (Abraham, Förstner, 2005) that appropriately models the fisheye distortions. Since we provide undistorted images to the equidistant camera model, the principal points of the camera heads correspond with its image centres. The image width amounts to 2048 pixels and the image height to 2448 pixels. The sensor pixel size is 3.45 μm . Table 3 lists the calibrated focal lengths of the individual panoramic camera heads.

Sensor	cam0	cam1	cam2	cam3	cam4	cam5
c [mm]	4.2607	4.2656	4.2698	4.2601	4.2505	4.2660

Table 3. Calibrated focal lengths c of the individual *Ladybug 5* panoramic camera heads.

3.3 Data Formats and Data Preparation

For both datasets, we provide the raw LiDAR and navigation data as well as the anonymized image data, to ensure free distribution without data protection issues.

Figure 3 shows the data acquisition frequency of the different sensors, which gives a first indication of the resulting data volume. The panoramic camera acquires image epochs consisting of six images from the individual camera heads with a frequency in the range of 0.5 to 2 Hz. The computer on the *BIMAGE Backpack* stores the raw images and the navigation unit generates precise timestamps. A first self-developed and python-implemented post-processing procedure undistorts the raw camera images to the equidistant camera model. Then, a second post-processing procedure using the open-source tool *Anonymizer* (understand.ai, 2019) detects and blurs personal data such as faces and car license plates.

Both horizontal and vertical LiDAR scanners acquire multi-profiles with a frequency of 10 Hz, whereby a total of 576'000 LiDAR points per second are recorded (see Figure 3). The slightly modified Robot Operating System (ROS)-based (Quigley et al., 2009) *Velodyne* driver (Withley, 2016) stores raw LiDAR sensor data packages within so-called rosbag files on the *BIMAGE Backpack* computer. With a post-processing step, using

a self-developed *Velodyne* driver extension, we calculate a precisely GNSS Seconds of Week (SOW) timestamped point cloud, which lies in the respective sensor coordinate frame.

The navigation unit *SPAN CPT7* records GNSS and IMU raw observations. In post-processing, we convert the proprietary data format to the GNSS RINEX data format and a CSV file containing the IMU raw observations. Furthermore, we also store the IMU raw data into a rosbag file.

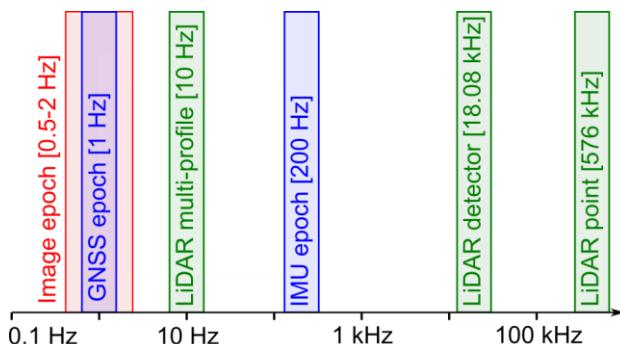


Figure 3. Frequency spectrum of the data recording by the *BIMAGE Backpack*. *Image epoch* denotes a high-resolution multi-head panoramic image consisting of 6 individual images. *GNSS* and *IMU epoch* represent one sensor reading. *LiDAR multi-profile* means one revolution of all multi-profile scanners, while *LiDAR detector* represents one point of one LiDAR profile and *LiDAR point* indicates one point of all LiDAR profiles and scanners.

4. DATASET DESCRIPTION

In this section, we describe the two backpack MMS datasets that we publish. We captured both datasets in challenging environments. The first dataset is from a test site in the city centre and the second dataset is from another test site in a forest.

Both datasets contain a) image data with undistorted equidistant and anonymized images from the individual panoramic camera heads as well as the image timestamps, b) LiDAR data represented as timestamped point clouds in the sensor coordinate frame, c) navigation data with GNSS raw observations from the *BIMAGE Backpack* as well as from the reference station and IMU raw data.

Our *BIMAGE* dataset website provides more detailed information about the data structure and the data formats. Furthermore, the website provides download access to the datasets for scientific use:

<https://www.fhnw.ch/habg/bimage-datasets>

4.1 City Centre

The first dataset was acquired in the city centre of Basel, Switzerland. The 800 m loop-shaped trajectory was recorded in 24 minutes. It includes different road and path widths including a large square with good GNSS reception for system initialization (see Figure 4, Image 1). By contrast, it also includes narrow alleys only accessible to pedestrians with steps and slopes up to 16 % (see Figure 4, Image 2). Wide pedestrian promenades with shops on both sides dominate other parts of the trajectory (see Figure 4, Image 3). Image 4 in Figure 4 shows the main traffic axis through the city centre with busy tram and bicycle traffic.

The dataset ‘city centre’ contains 721 panoramic images, approx. 840 million LiDAR points, GNSS data as available and IMU data. We provide 15 ground control points (GCPs) arranged in groups of three and 18 check points (CPs) along the first loop of the trajectory (see Figure 4). Most of the GCPs and CPs are well-defined natural reference points, but some were marked with photogrammetric targets. Fricker and Weber (2019) provide a detailed description of the reference point measurements by tachymetry and show a 3D standard deviation below 5 mm.

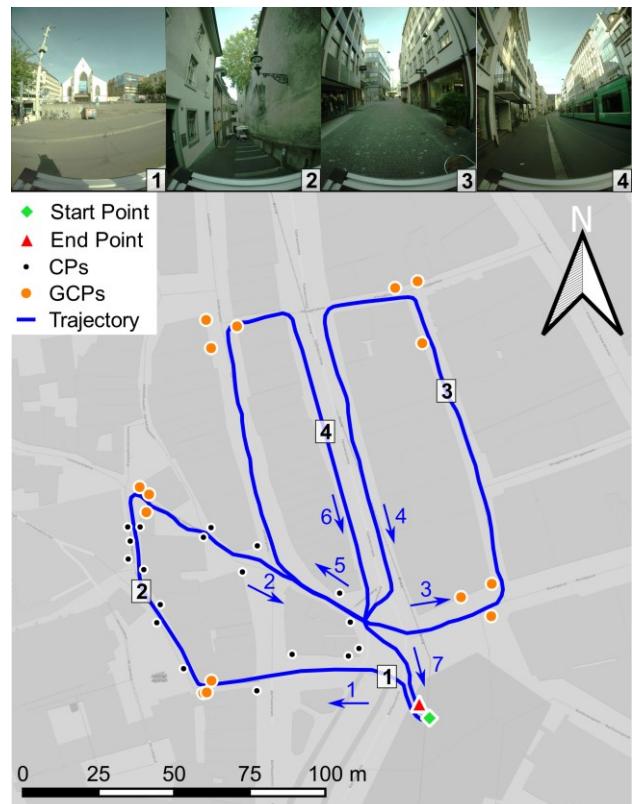


Figure 4. Map from the ‘city centre’ dataset with images showing typical environmental conditions. We extended this map from Blaser et al. (2020) by the check points (CPs) and ground control points (GCPs) that we publish with this dataset.

4.2 Forest

The second dataset was acquired in a partially dense forest. Its trajectory length amounts to 740 m and the data capture required 25 minutes. It also incorporates an area with good GNSS reception at a nearby highway exit for system initialization (see Figure 5, Image 1). Furthermore, the forest path leads through a road underpass (see Figure 5, Image 2). Narrow paths only accessible to pedestrians with dense vegetation at ground level dominate the scenery in images 3 and 6 of Figure 5. In addition, the trajectory also includes drivable forest roads with less dense vegetation (see Figure 5, Images 4 and 5).

The ‘forest’ dataset includes 843 panoramic images, approx. 850 million LiDAR points and navigation data in the scope of the first dataset. We provide 15 GCPs arranged in groups of three and 8 CPs along the first segment of the trajectory (see Figure 5). All points are marked with photogrammetric targets and fixed either on trees or on driven-in pillars. Fricker and Weber (2019) describe the reference point measurements by tachymetry with closed polygons as well as the geodetic evaluation, which shows a 3D standard deviation of 5 mm.

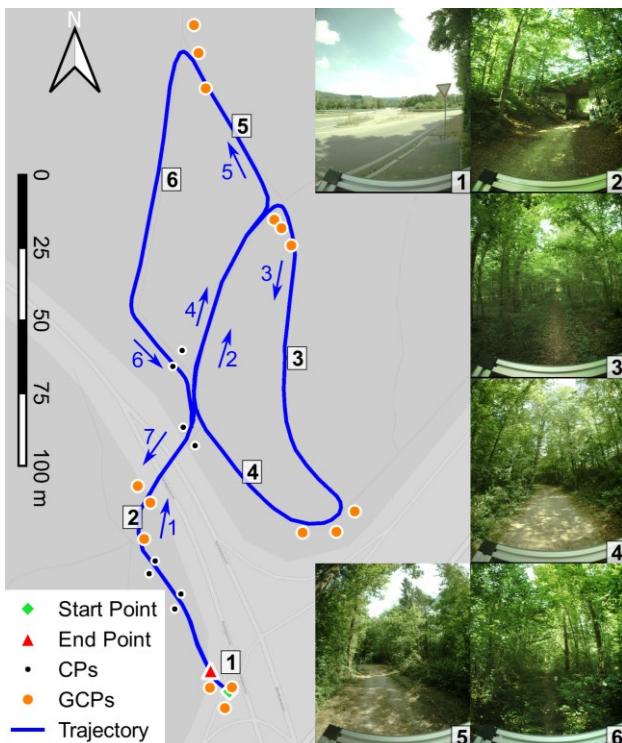


Figure 5. Map from the ‘forest’ dataset with images showing typical environmental conditions. We extended this map from Blaser et al. (2020) by the check points (CPs) and ground control points (GCPs) that we publish with this dataset.

5. APPLICATIONS AND FIRST EXPERIMENTS

Both high-quality datasets provide a great potential for evaluating and improving methods for multiple tasks related to high-precision 3D reality capture and the creation of digital twins. Potential applications range from localization and georeferencing, dense image matching and 3D reconstruction to combined methods such as SLAM and Structure-from-Motion (SfM) as well as classification and scene interpretation.

In this section, we show and discuss first experiments using the published datasets in the fields of georeferencing, SLAM and 3D reconstruction.

5.1 Georeferencing

In the past, Blaser et al. (2020) successfully conducted investigations on the localization and georeferencing using both datasets. They investigated and compared the three different georeferencing methods a) direct georeferencing, b) SLAM-based georeferencing and c) image-based georeferencing.

They processed the direct georeferencing using data from the navigation unit *SPAN CPT7* using tightly coupled GNSS and IMU sensor data fusion with a Kalman filter using the *Waypoint Intertial Explorer* software. By contrast, for LiDAR SLAM-based georeferencing they independently processed raw IMU and LiDAR data with the 3D SLAM algorithm *Google Cartographer* (Hess et al., 2016). Finally, they introduced the LiDAR SLAM-based georeferencing poses and the panoramic camera images into the SfM-pipeline *Agisoft Metashape* and performed the image-based georeferencing using a camera rig constrained bundle-adjustment.

They achieved median GCP and CP coordinate differences between 45.2 cm and 100.2 cm using direct georeferencing, between 21.0 cm and 36.6 cm using SLAM-based georeferencing and between 4.3 cm and 13.4 cm with image-based georeferencing.

However, there is a great potential for further accuracy and robustness improvements by combining and coupling different georeferencing methods or by developing novel methods combining different sensor data. Nevertheless, georeferencing forms the basis for mapping and further applications and products and has a direct influence on its accuracy. Thus, a more accurate georeferencing enables accurate reconstruction and mapping.

5.2 Simultaneous Localization and Mapping

SLAM has great potential as it is not only an alternative georeferencing method for areas with poor GNSS coverage, but also simultaneously generates a map in near real-time. Especially LiDAR SLAM is promising on both datasets because the LiDAR acquisition data frequency is higher, and the processing effort lower compared to images and visual SLAM. Thus, we also obtain a point cloud when performing the SLAM-based georeferencing. Figure 6 shows the resulting point cloud that incorporates the trajectory, which is projected on the XY-plane. Not only the street-level is clearly visible, but the point cloud also partially depicts commercial indoor areas, which are visible from the street-level. If the point cloud has an accuracy analogous to SLAM-based georeferencing, it is not only sufficient for completeness checks during data acquisition but could also be used as a tool for higher-level urban planning.



Figure 6. XY-plane from the resulting point cloud of the LiDAR SLAM using the city centre dataset.

The LiDAR point cloud from the forest dataset (see Figure 7) also clearly shows single trees and fine-grain structures, so that the point cloud has great potential for forest applications. The improvement of SLAM algorithms has enormous potential, as more available accurate real-time 3D information opens numerous other applications.

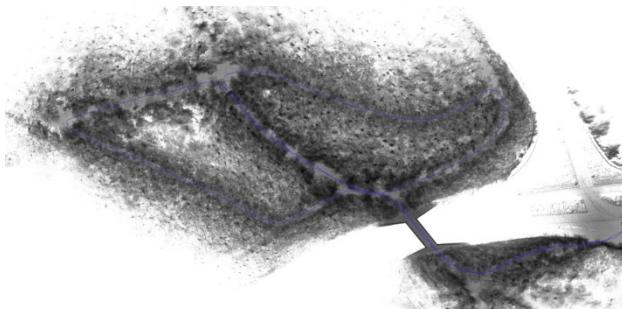


Figure 7. XY-plane from the resulting point cloud of the LiDAR SLAM using the forest dataset.

5.3 3D Reconstruction

Furthermore, we performed a 3D reconstruction using the city centre dataset based on the image-based georeferenced image poses. We used the 3D reconstruction software *ContextCapture* from Bentley, which also supports fisheye camera models.

The result of the automatic 3D reconstruction process was a highly detailed 3D city model from street-level perspective which is similar to previous investigations (Blaser et al., 2017) with vehicle-based MMS, despite the lower image density of the *BIMAGE Backpack*.

Completeness and level of detail of the automatic reconstruction are remarkable (see Figure 8). However, the reconstruction accuracy decreases with increasing height. Nevertheless, such a street level dataset has great potential for reality-based VR traffic simulations (Wahbeh et al., 2021) or to complement aerial-based city models.



Figure 8. Samples of the automatically reconstructed 3D city model of the ‘city centre’ dataset using street-level backpack MM data

6. CONCLUSION AND OUTLOOK

In this paper, we provided two high-quality datasets captured with the *BIMAGE Backpack* MMS in challenging urban and forest environments. We further described our *BIMAGE Backpack* MMS in detail, provided the overall system calibration parameters and specified the resulting raw and pre-processed data so that the datasets can be fully used for future investigations. We then described both test sites ‘city centre’ and ‘forest’ and their associated data sets. The quality of these data sets has been thoroughly evaluated and demonstrated. Blaser et al. (2020), for example, achieved georeferencing accuracies in the centimetre range using these data sets in combination with image-based georeferencing.

Both datasets can be used for developing, testing, and improving digital twin-related tasks (e.g. georeferencing, SLAM, SfM, 3D reconstruction, classification, and scene interpretation). In the future, we aim to provide contests in various fields, possibly in cooperation with interested other groups and universities.

ACKNOWLEDGEMENTS

The research project PReCIS (20200331_04_PReCIS) was supported by Forschungsfonds Kt. Aargau and by our industrial partner iNovitas AG.

REFERENCES

- Abraham, S., Förstner, W., 2005. Fish-eye-stereo calibration and epipolar rectification. *ISPRS J. Photogramm. Remote Sens.* 59(5), 278-288. doi.org/10.1016/j.isprsjprs.2005.03.001
- Blaser, S., Cavegn, S., Nebiker, S., 2018. Development of a Portable High Performance Mobile Mapping System Using the Robot Operating System. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, IV-1, 13-20. doi.org/10.5194/isprs-annals-IV-1-13-2018
- Blaser, S., Meyer, J., Nebiker, S., Fricker, L., Weber, D., 2020. Centimetre-Accuracy in Forests and Urban Canyons – Combining a High-Performance Image-Based Mobile Mapping Backpack with new Georeferencing Methods. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, V-1-2020, 333-341. doi.org/10.5194/isprs-annals-V-1-2020-333-2020
- Blaser, S., Nebiker, S., Cavegn, S., 2017. System Design, Calibration and Performance Analysis of a Novel 360° Stereo Panoramic Mobile Mapping System. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, IV-1/W1, 207-213. doi.org/10.5194/isprs-annals-IV-1-W1-207-2017
- Cavegn, S., Haala, N., Nebiker, S., Rothermel, M., Tutzauer, P., 2014. Benchmarking High Density Image Matching for Oblique Airborne Imagery. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XL-3, 45-52. doi.org/10.5194/isprsarchives-XL-3-45-2014
- Choi, S., 2019. Awesome Robotics Datasets. <https://sunglok.github.io/awesome-robotics-datasets/> (22.01.2021).
- Fricker, L., Weber, D., 2019. Portables Mobile Mapping im Aussenraum. Bachelor Thesis. FHNW University of Applied Sciences and Arts Northwestern Switzerland, Switzerland (unpublished).

- Geiger, A., Lenz, P., Urtasun, R., 2012. Are we ready for autonomous driving? The KITTI vision benchmark suite. *2012 IEEE Conference on Computer Vision and Pattern Recognition*, Providence, RI, 2012, 3354–3361.
doi.org/10.1109/CVPR.2012.6248074
- Hess, W., Kohler, D., Rapp, H., Andor, D., 2016. Real-Time Loop Closure in 2D LiDAR SLAM. *IEEE International Conference on Robotics and Automation (ICRA)*, Stockholm, Sweden, 1271–1278. doi.org/10.1109/ICRA.2016.7487258
- Khoshelham, K., Díaz Vilariño, L., Peter, M., Kang, Z., Acharya, D., 2017. The ISPRS Benchmark on Indoor Modelling. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLII-2/W7, 367–372. doi.org/10.5194/isprs-archives-XLII-2-W7-367-2017
- Lehtola, V. V., Kaartinen, H., Nüchter, A., Kaijaluoto, R., Kukko, A., Litkey, P., Honkavaara, E., Rosnell, T., Vaaja, M. T., Virtanen, J.-P., Kurkela, M., Issaoui, A. El, Zhu, L., Jaakkola, A., Hyppä, J., 2017. Comparison of the Selected State-Of-The-Art 3D Indoor Scanning and Point Cloud Generation Methods, *Remote Sens.*, 9(8). 796. doi.org/10.3390/rs9080796
- Meilland, M., Comport, A. I., Rives, P., 2015. Dense Omnidirectional RGB-D Mapping of Large-scale Outdoor Environments for Real-time Localization and Autonomous Navigation. *Journal of Field Robotics*, 32(4), 474–503. doi.org/10.1002/rob.21531
- Nebiker, S., Cavegn, S., Loesch, B., 2015. Cloud-Based Geospatial 3D Image Spaces—A Powerful Urban Model for the Smart City. *ISPRS Int. J. Geo-Information*, 4(4), 2267–2291. doi.org/10.3390/ijgi4042267
- Nex, F., Gerke, M., Remondino, F., Przybilla, H.-J., Bäumker, M., Zurhorst, A., 2015. ISPRS Benchmark for Multi-Platform Photogrammetry. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, II-3/W4, 135–142. doi.org/10.5194/isprsannals-II-3-W4-135-2015
- Novak, K., 1991. The Ohio State University Highway Mapping System: The Stereo Vision System Component. *Proc. of the 47th Annual Meeting of the Institute of Navigation*, Williamsburg, VA, 121–124.
- Quigley, M., Conley, K., Gerkey, B., Faust, J., Foote, T., Leibs, J., Berger, E., Wheeler, R., Ng, A., 2009. ROS: an open-source Robot Operating System. *ICRA workshop on open source software*, 3(3.2), 5.
- Schwarz, K. P., Martell, H. E., El-Sheimy, N., Li, R., Chapman, M. A., Cosandier, D., 1993. VIASAT - A mobile highway survey system of high accuracy. *Proceedings of the Vehicle Navigation and Information Systems Conference*, Ottawa, Canada, 476–481.
- Tan, W., Qin, N., Ma, L., Li, Y., Du, J., Cai, G., Yang, K., Li., J., 2020. Toronto-3D: A Large-scale Mobile LiDAR Dataset for Semantic Segmentation of Urban Roadways. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 202–203.
doi.org/10.1109/CVPRW50498.2020.00109
- understand.ai, 2019. Anonymizer. Software.
<https://github.com/understand-ai/anonymizer> (14 April 2021).
- Wang, C., Dai, Y., El-Sheimy, N., Wen, C., Retscher, G., Kang, Z., Lingua, A., 2019. Progress on ISPRS Benchmark on Multisensory Indoor Mapping and Positioning. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLII-2/W13, 1709–1713. doi.org/10.5194/isprs-archives-XLII-2-W13-1709-2019
- Wahbeh, W., Ammann, M., Nebiker, S., Van Eggermond, M., Erath, A., 2021. Image-Based Reality-Capturing and 3D Modelling for the Creation of VR Cycling Simulations. To appear in: *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*
- Wen, C., Pan, S., Wang, C., Li, J., 2016. An indoor backpack system for 2-D and 3-D mapping of building interiors. *IEEE Geoscience and Remote Sensing Letters*, 13(7), 992–996. doi.org/10.1109/LGRS.2016.2558486
- Withley, J., 2016. Velodyne Driver. ROS Wiki https://wiki.ros.org/velodyne_driver (19 April 2021).