

PORTABLE IMAGE-BASED HIGH PERFORMANCE MOBILE MAPPING SYSTEM IN UNDERGROUND ENVIRONMENTS – SYSTEM CONFIGURATION AND PERFORMANCE EVALUATION

S. Blaser¹, S. Nebiker¹, D. Wisler¹

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

Commission I, ICWG I/IV

KEY WORDS: Mobile Mapping, Underground, Robot Operating System, SLAM, Sensor Orientation, Performance Evaluation

ABSTRACT:

The progression in urbanization increases the need for different types of underground infrastructure. Consequently, infrastructure and life cycle management are rapidly gaining in importance. Mobile reality capturing systems and cloud-based services exploiting georeferenced metric 3D imagery are already extensively used for infrastructure management in outdoor environments. These services minimise dangerous and expensive field visits or measurement campaigns. In this paper, we introduce the BIMAGE Backpack, a portable image-based mobile mapping system for 3D data acquisition in indoor environments. The system consists of a multi-head panorama camera, two multi-profile laser scanners and an inertial measurement unit. With this system, we carried out underground measurement campaigns in the Hagerbach Test Gallery, located in Flums Hochwiese, Switzerland. For our performance evaluations in two different tunnel sections, we employed LiDAR SLAM as well as advanced image-based georeferencing. The obtained absolute accuracies were in the range from 6.2 to 7.4 cm. The relative accuracy, which for many applications is even more important, was in the range of 2-6 mm. These figures demonstrate an accuracy improvement of the subsequent image-based georeferencing over LiDAR SLAM by about an order of magnitude. The investigations show the application potential of image-based portable mobile mapping systems for infrastructure inventory and management in large underground facilities.

1. INTRODUCTION

Underground construction and infrastructure gain importance with the ongoing progression in urbanization. This applies in particular to underground network infrastructures such as subways. Digitization trends on a broad scale lead to a rapid and massive transformation of the construction and real estate industry as well as infrastructure management. New methods and tools such as BIM (Building Information Modelling) and VDC (Virtual Design and Construction) in combination with new technologies for geospatial data capture and exploitation enable a paradigm shift in the way buildings and infrastructure are designed, tested, built, maintained and refurbished. The establishment of fully three-dimensional collaborative processes and workflows with stakeholders from multiple domains require accurate, detailed and up-to-date 3D geo data. New image-based mobile reality capturing techniques in combination with cloud technologies, such as presented by Nebiker et al. (2015), hold the potential to provide such data and services in a rapid, cost-efficient and user-friendly manner. As shown by Puente et al. (2013) current commercial mobile mapping systems (MMS) are dominated by LiDAR as primary sensors and with cameras as complementary sensors. However, first road- and railway-based mobile mapping experiments were based on stereo camera systems and date back to the early 1990ies (Novak, 1991; Schwarz et al., 1993). Since then image-based outdoor mobile mapping systems have evolved into multi-stereo systems (Cavegn & Haala, 2016) and into high-performance 360° stereo systems (Blaser et al., 2018b; Blaser et al., 2017; Meiland et al., 2015) with an unparalleled information richness and density.

In terms of positioning technologies, the vast majority of outdoor mobile mapping solutions currently rely on direct georeferencing using GNSS and INS. However, alternative

approaches are required for positioning and pose estimation in indoor or underground environments and in environments with poor GNSS coverage such as forests or urban canyons. Recent developments in indoor localization and mapping benefit from and build on methods and techniques, namely SLAM (simultaneous localization and mapping), from the robotics and more recently from the computer vision communities (Stachniss et al., 2016; Thrun, 2002). Indoor and underground mobile mapping not only requires new georeferencing strategies, but also new platforms. Solutions using carts, like Viametris iMMS (Thomson et al., 2013), work well in large unobstructed spaces but are not suitable for typical indoor environments with stairs, closed doors, obstructed floors due to ongoing construction etc. Thus, the focus in research and commercial development has shifted towards portable or 'personal' mobile mapping systems (Lethola et al., 2017; Nüchter et al., 2015). However, with very few exceptions, such as the image-based UltraCam Panther (Vexcel, 2018), most developments focus on LiDAR-based systems and point clouds. Griesbach et al. (2014) introduce an even smaller Integrated Positioning System (IPS) "DMT PILOT 3D" for underground localization and mapping, equipped with a stereo camera system and an INS. Other mapping sensors could also extend their system. Blaser et al. (2018a) introduce the BIMAGE backpack, a portable high-performance mapping research platform and system built on top of the Robot Operating System (ROS). The system is designed for the creation of 3D image data spaces and for the image-based building and infrastructure management (Nebiker et al., 2015). It features LiDAR SLAM-based real-time 3D mapping with a subsequent novel image-based georeferencing approach using relative orientation constraints leading to significant improvements in relative and absolute accuracies.

In this paper, we examine the potential for using the BIMAGE backpack in underground environments. As test area, we use the Hagerbach Test Gallery located in Flums Hochwiese, Switzerland. For the investigations, we collected data in various galleries with different stages of development.

First, we describe the image-based mobile mapping system *BIMAGE backpack*, which we use for our investigations. Further, we introduce the test site Hagerbach Test Gallery. In a performance evaluation, we examine the absolute and relative accuracies, show and apply different methods for significant accuracy improvement. Further, we discuss the results in terms of potential use.

2. IMAGE-BASED MOBILE MAPPING SYSTEM

The new image-based Mobile Mapping System (MMS) *BIMAGE backpack* should support kinematic 3D data acquisition in indoor environments. Similar to existing outdoor MMS (Burkhard et al., 2012; Blaser et al. 2018b), the indoor environments should be captured and represented as georeferenced 3D image spaces. Nebiker et al. (2015) describe the concept and the construction of geospatial 3D image spaces and discuss the advantages over 3D point clouds, namely in terms of handling and user-friendliness. Associated depth values for almost every pixel enable 3D measurements directly in the image with a single mouse click. With given exterior orientation parameters, the web application can calculate absolute 3D coordinates. Along with the depth information, 3D images can optionally contain further pixel-based information layers, e.g. quality of depth values or object classification.

A prototypic indoor MMS, even for research purpose, should be portable and extensible in order to cover as many use cases as possible and to study different sensor configurations. Further, the mapping sensors should cover the indoor environment as completely as possible. Precise georeferencing is essential in mobile mapping. Most outdoor MMS use inertial navigation systems (INS) combining GNSS and IMU. The absence of GNSS in indoor environments precludes this sensor combination from being applied. A possible way to overcome the lack of GNSS is to replace GNSS with LiDAR-based simultaneous localization and mapping (SLAM) (Hess et al., 2016). Subsequent image-based georeferencing can further improve the LiDAR-based camera poses and transform them into a global reference frame as well (Cavegn et al., 2016).

2.1 System configuration

We mounted all sensors on a robust aluminium frame, which we fixed on a backpack carrying frame. The system contains a FLIR Ladybug5 multi-head panorama camera. The panorama camera consists of five radially arranged camera heads with a horizontal line of sight as well as a sixth camera head pointing upwards. The resolution of a single camera head is 5 MP. Further, we tilted the panorama camera by a few degrees to achieve a roughly horizontal camera plane when a person carries the backpack (see Figure 1, no. 2). The overlapping of images allows stitching them together to a panorama image, so that the surrounding frame is barely visible. For navigation, we use two Velodyne VLP-16 multi-profile LiDAR scanners as well as an XSens MTI-300 IMU. One of the two LiDAR scanners sits on top of the frame and is tilted by 30° to cover mainly the walls as well as some parts of the floor and the ceiling (see Figure 1, no. 1). The second, vertically mounted LiDAR scanner (see Figure 1, no. 5) covers mainly the floor and the ceiling. The second LiDAR is complementary to the

first one and gives additional geometric stability to the resulting fused point cloud. The 3D LiDAR SLAM algorithm registers both LiDAR point clouds to a voxel-based map, which is created simultaneously and used to estimate the relative system pose (see section 2.2). A 12 V battery with a capacity of 20 Ah (see Figure 1, no. 6) powers all system components. The current system weight is about 20 kg. Blaser et al. (2018a) give a detailed system overview with the main specifications of all components used.

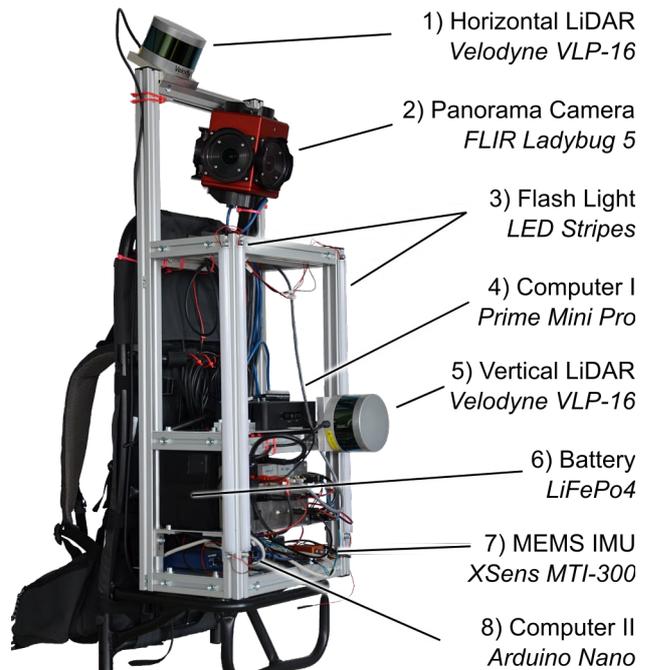


Figure 1. System overview demonstrating our backpack indoor mobile mapping system BIMAGE Backpack (Blaser et al., 2018a).

2.2 Sensor synchronization and acquisition software

Sensor synchronization of MMS platforms is essential and exerts a direct influence on the measuring accuracy. The reference time of the BIMAGE backpack is on Computer II Arduino Nano (see Figure 2). Computer II generates a pulse per second (PPS) and a NMEA string as well and synchronizes all navigation sensors continuously. The acquisition software on Computer I is able to generate and to send a trigger command to Computer II. Then, Computer II triggers the panorama camera as well as the flashlights with an electric pulse and sends the precise trigger time back to Computer I (see Figure 2).

The acquisition software follows a modular and flexible concept and bases on the Robot Operating System (ROS). ROS is an open-source robotic framework, which is widely used in robotic communities. It contains numerous existing tools and a comprehensive hardware support. Quigley et al. (2009) give an introduction of principles, paradigms and the functionality of ROS. Particularly noteworthy is the flexible graph-based communication concept consisting of nodes, messages, topics and services. A node describes a software module, which we represent as an ellipse in our software schema (see Figure 3). Nodes can communicate with each other by passing messages in a strictly predefined data structure through given topics. In our software schema, we represent topics as rectangles (see Figure 3). Either, a node can publish messages to one or more topics (red arrows) or can subscribe to one or more topics to receive

messages (blue arrows). By contrast, services offer synchronous communication. A node can advertise a service with predefined data structures of both request and response, similar to a web service. We represent services and topics as rectangles in our software schema (see Figure 3) but indicate the communication with services with green arrows.

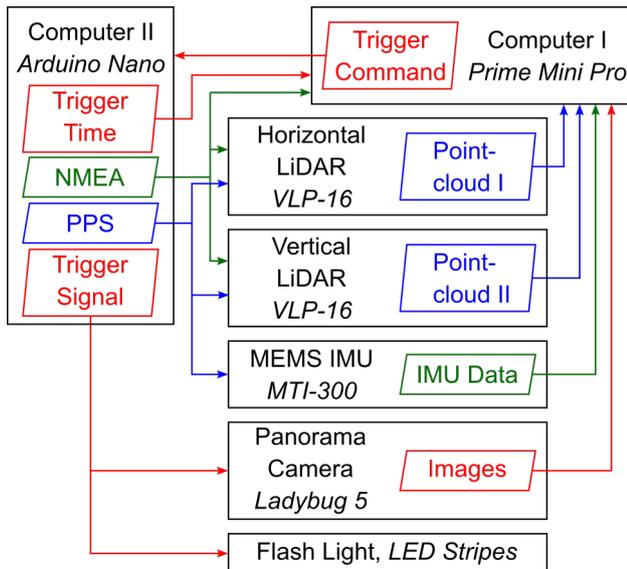


Figure 2. Schema with sensor synchronization and data flow (Blaser et al., 2018a).

A ROS bag node subscribes all messages from navigation sensors and writes the raw messages within a bag file on the hard disc (see Figure 3). Further, the Cartographer node – containing the 3D LiDAR SLAM algorithm Google Cartographer – subscribes all messages from navigation sensors as well. Hess et al. (2016) introduce the functionality of Google’s Cartographer exemplarily with their 2D LiDAR SLAM. Since the 3D LiDAR SLAM provides a real-time system pose, our conditional trigger node supports spatial trigger criteria, such as distance intervals and orientation differences, as well as non-spatial criteria such as time difference and triggers a command if necessary. The node of Computer II advertises a trigger command service, publishes the trigger time and logs all incoming messages from Computer II, which triggers the panorama camera hardware-based. The Panorama Camera node returns all six images from the camera heads strung together in one image. The Image Slicer node re-separates the images and different Image Writer node instances write the single images.

Our prototypic MMS initializes completely indoors. The starting position of the LiDAR SLAM defines the origin of the local coordinate frame. In addition, the LiDAR SLAM provides a map overview containing the trajectory; the fused LiDAR point cloud as well as the 2D visualization of the voxel map (see Figure 4, top). A further tab in the graphical user interface of our data acquisition software provides a preview of all camera heads (see Figure 4, bottom). Camera trigger constraints based on time, distance or angle can be set in this tab. Furthermore, the acquisition software provides four different camera configuration pre-sets for different lightning conditions. The camera configuration can be set manually as well. Blaser et al. (2018a) provide an even more detailed description of the implemented acquisition software.

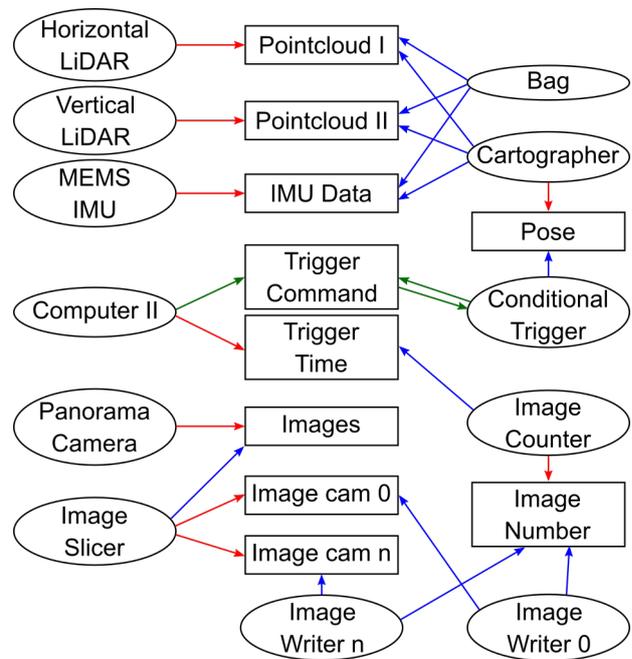


Figure 3. Schema with the implemented ROS software structure. ROS nodes are represented as circles and ROS topics as rectangles. Red arrows depict ROS publishers, blue arrows ROS subscribers and green arrows ROS services (Blaser et al., 2018a).

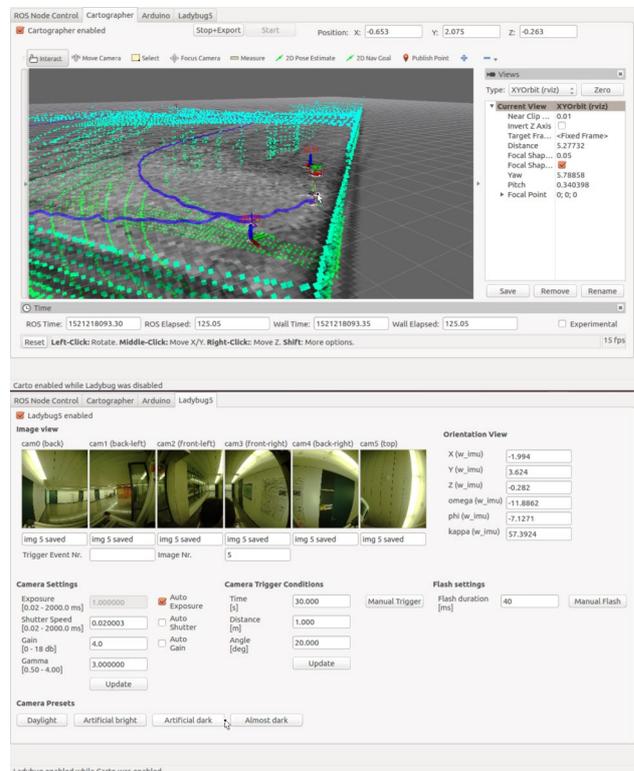


Figure 4. Screenshots of the data acquisition software. Top: Cartographer SLAM preview, Bottom: Image preview and conditional trigger settings (Blaser et al., 2018a).

2.3 Data post-processing workflow

For the data post-processing workflow, we use pre-calibrated parameters of the panorama camera. Blaser et al. (2018) describe the calibration procedure of the multi-head panorama camera Ladybug5. They estimate the interior orientation parameters (IOPs) of each camera head and the relative orientation parameters (ROPs) between the camera heads with a multi-constrained bundle adjustment. They use the equidistant camera model (Abraham & Förstner, 2005) for the camera heads. Further, we adopted boresight parameters (BAPs) between the panorama camera and the body frame, as well as between the laser scanners and the body frame from the construction plan. The analytical calibration of BAPs is still pending.

After data acquisition, we export the so-called cartographer state from LiDAR SLAM. In the following steps, we extract the trajectory from cartographer state and interpolate image events in order to get SLAM-based image poses (see Figure 5). Within a separate process, we undistort all raw images from the panorama camera heads using pre-calibrated IOPs.

At this stage, we can already build image-based web services with undistorted images and corresponding SLAM-based camera poses. Further, we can already carry out coordinate measurements with forward intersection using image measurements from several consecutive images.

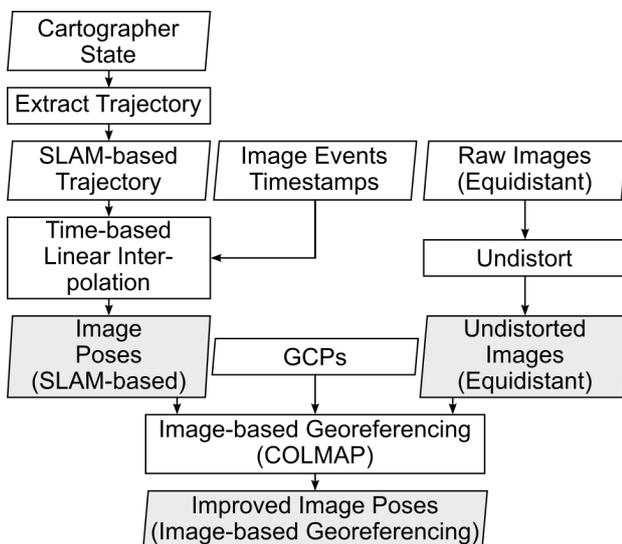


Figure 5. Flow chart indicating our data post-processing workflow (Blaser et al., 2018a).

Cavegn et al. (2018) extended the SfM pipeline COLMAP (Schönberger & Frahm, 2016) with Georeferencing capabilities by integrating prior EOPs and exploiting ROP constraints. In application scenarios with high accuracy requirements, we aim to improve LiDAR SLAM-based camera poses with a subsequent image-based georeferencing. Hence, we use LiDAR SLAM-based EOPs of the first camera as initial values, and fix ROPs of the other camera heads with the pre-calibrated values. Cavegn et al. (2018) describe the complete process of integrated Georeferencing used COLMAP in detail. Good lighting conditions and a structurally rich environment for proper and well-distributed feature detection are essential.

3. TEST SITE

The Hagerbach Test Gallery (VSH) constitutes a suitable research and development infrastructure for investigations and practical tests in underground mining. VSH was founded in 1970 and is located in Flums Hochwiese, Switzerland. The Amberg Group is the owner of VSH and mining companies can rent galleries to develop and test their innovations. In addition to a concrete mixer and other construction infrastructure, Amberg Group operates a laboratory for construction materials testing underground (VersuchsStollen Hagerbach AG, 2019). The VSH grew over the years to a widely branched and complex tunnel system with various stages of development and drilling diameter (see Figure 6).

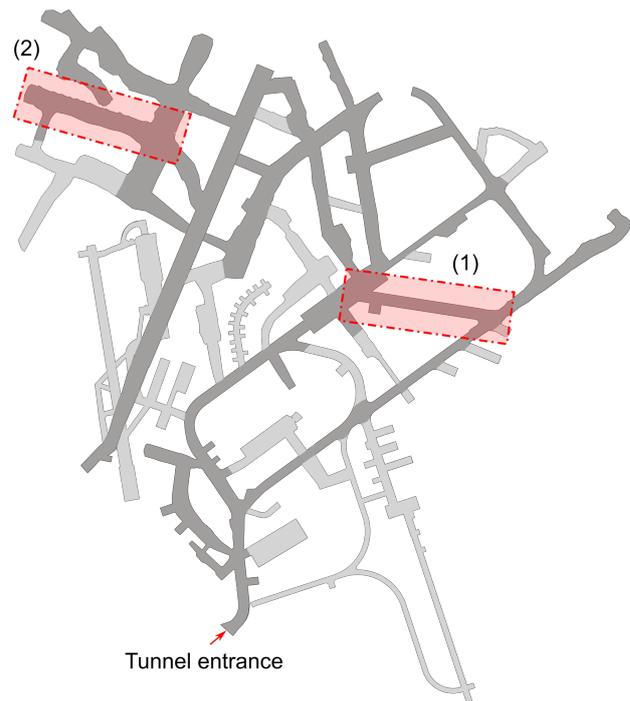


Figure 6. Ground plan of the Hagerbach Test Gallery (VSH). Dark grey areas depict accessible tunnels, while the access of light grey areas is restricted. Both test areas “safety tunnel” (1) and “SCAUT tunnel” (2) are delineated with red dash-dotted lines.

3.1 Reference data

Accurate reference data was acquired in a measurement campaign using high-end tachymetry. Rechsteiner & Wisler (2018) provide a detailed description of the reference point measurements as well as the geodetic adjustment. The VSH contains 127 reference points marked with screw bolts. They evaluated the entire reference point network using least-squares adjustment and with minimal constraints in order to prevent error influences of reference points from former measurement campaigns. The maximum standard deviation in planar space amount to 4.1 mm and 0.9 mm in vertical respectively.

3.2 Hagerbach Test Gallery

For our investigations, we chose two separate test sites with different shaping. Our first test site named “safety tunnel” (see Figure 6, (1)) is designed as a road tunnel with different extension levels (see Figure 7, top). Our second test site (see

Figure 6, (2)) is the tunnel of the Swiss Center of Applied Underground Technologies (SCAUT). In contrast, the “SCAUT tunnel” contains an innovative test setup for the automatic installation of emergency exits as well as railway infrastructure (see Figure 7, bottom).



Figure 7. Test site “safety tunnel” (1) is developed as a road tunnel with different stages of development (top). In contrast, the test site “SCAUT tunnel” (2) contains the infrastructure for a railway tunnel as well as a crosscut to the rescue tunnel (bottom).

4. PERFORMANCE EVALUATION

In both test sites “safety tunnel” (1) and “SCAUT tunnel” (2), we carried out data acquisition with the BIMAGE backpack. In order to overcome the low lighting conditions, we fixed additional spotlights sideward on the backpack. In addition, we set the camera trigger constraints to a distance interval of 2 m. The data acquisition procedure of the first test site “safety tunnel” is loop-shaped with an overlap area (see Figure 8, left). Further, we captured the second test field “SCAUT tunnel” with an additional tunnel after a sharp turn.

The LiDAR SLAM also provides a 3D point cloud with co-registered LiDAR data. Plane projections of the 3D point cloud are useful for a first visual assessment of the SLAM estimation (see chapter 4.1).

For quantitative investigations, we compare image-measured points with reference points. Within four consecutive images with corresponding camera poses, we measure target points and estimate relative 3D point coordinates with bundle adjustment-based forward intersection using a self-developed Python tool. Afterwards, we transform the 3D points into the reference frame performing a 6DoF coordinate transformation.

4.1 SLAM-based Georeferencing

The loop-shaped data acquisition of our first test site “safety tunnel” (see Figure 8, left) starts nearby the fork at the top right (S) and terminates at the fork bottom right (E). There is an overlap area between start and endpoint needed for loop-closure. By contrast, the trajectory in our second test site “SCAUT tunnel” (see Figure 8, right) begins in the first curve at the top left (S), turns at the end of the “SCAUT tunnel” and returns to near the starting point (E).

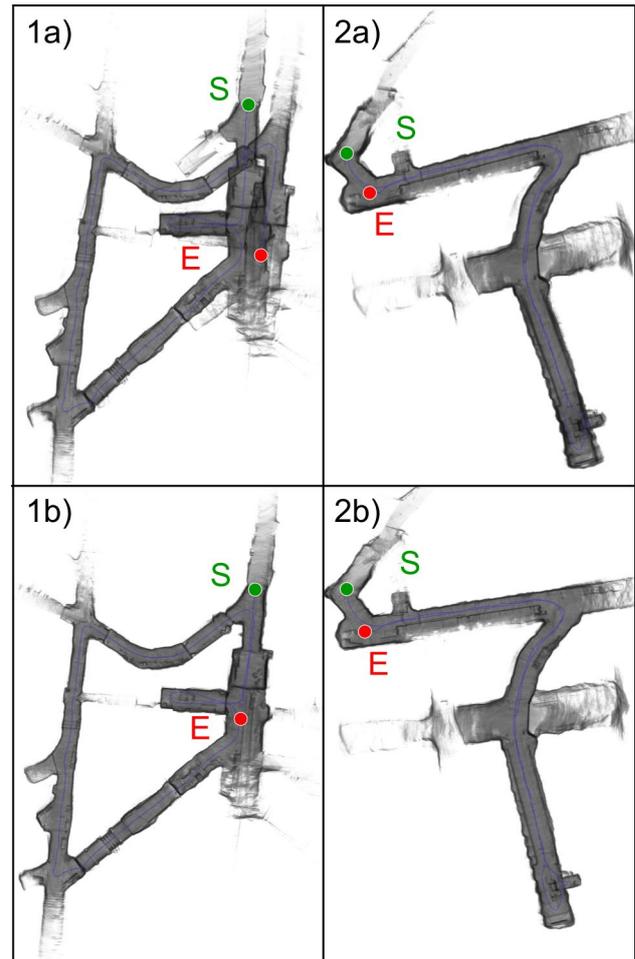


Figure 8. Point cloud projections to the XY-plane from 3D LiDAR SLAM algorithm. Point clouds from test site “safety tunnel” (1) are left and those from test site “SCAUT tunnel” (2) are right. We processed the point clouds at the top (a) with the original SLAM parameter set and the point clouds at the bottom (b) with the optimized parameter set.

The projection of the 3D point cloud resulting from LiDAR SLAM onto the XY-plane yield a first visual quality indicator of the SLAM estimation (see Figure 8). First, we performed the SLAM algorithm with the standard configuration, whereby a significant loop-closure error occurred in the test field “safety tunnel” (see Figure 8a). We could overcome the loop-closure error with continuous SLAM parameter optimization, namely by increasing the rotation weight of the Ceres Scan matcher and by increasing the number of node optimizations. Rechsteiner & Wisler (2018) describe the LiDAR SLAM parameter optimization in detail. Similar to the first test site “rescue tunnel”, we performed the LiDAR SLAM for the “SCAUT tunnel” with

both parameter sets. The point cloud projection based on the optimized SLAM parameter set (see Figure 8, 2b) shows less noise compared to the point cloud projection based on the original SLAM parameter set (see Figure 8, 2a).

4.2 Image-based Georeferencing

In order to perform subsequent image-based georeferencing, we introduce undistorted images from all individual camera heads into the SfM-Pipeline Agisoft PhotoScan. Additionally, we introduced all corresponding camera poses to the first camera and fixed the ROPs between the first and all other camera heads. Further, we did not introduce any target points. Hence, the optimized camera poses from the bundle-adjustment completely refer to the initial LiDAR SLAM-based camera poses. Processing time and resulting reprojection errors are similar to Cavegn et al. (2018).

4.3 Results and discussion

Using a) LiDAR SLAM-based image poses, processed with standard SLAM configuration, b) LiDAR SLAM-based image poses, processed with optimized SLAM-parameters as well as c) improved image poses from subsequent image-based georeferencing, we measure signalized target points within four consecutive images and estimate 3D coordinates with a bundle adjustment-based forward intersection.

4.3.1 Relative accuracies

First, we analyse the empirical standard deviations from forward intersection. They represent the precision of 3D coordinate observations. The relative orientation accuracies in-between consecutive images affect the precision of 3D coordinate observations the most. Cavegn et al. (2018) investigate on the influence from ROP calibration on 3D measurements. They do not observe any significant influence from ROP calibrations on 3D measurements. Hence, in this contribution we also use image measurements from different camera heads. The precision of a 3D coordinate observation indicates the achievable relative coordinate measurement accuracy and is a good measure for the accuracy of a 3D distance.

Our investigations show a significant improvement of the precision of 3D measurements. It ranges from 76-123 mm for case a) using SLAM-poses determined with standard parameters and improves to 61-81 mm for case b) using SLAM-poses determined with optimized parameters. Case c) using camera poses from image-based georeferencing significantly improves the precision of 3D measurement to 2-6 mm (see Figure 9 and Table 1). Using SLAM poses, the precision of 3D measurements was worse by factor two in test field “SCAUT tunnel” (2) compared to test field “safety tunnel” (1). We assume the larger diameter of test field “SCAUT tunnel” and thereby the lower point cloud density as a reason. Following the fact that a lower point cloud density leads to a poorer feature detection in the LiDAR SLAM.

By contrast, when using poses from image-based georeferencing for 3D coordinate observation, a narrower tunnel impedes bundle adjustment, due to worse intersection conditions.

4.3.2 Absolute accuracies

In order to analyse absolute accuracies, we transform each dataset containing 3D measurements to the ground truth coordinate frame using the same four ground control points (GCPs) (see Figure 10). Further, we estimate the differences between the transformed 3D measurements and ground truth. For each dataset, we separately report mean differences for

GCPs, CPs within the test site (without extrapolation) as well as for all available CPs (see Table 1).

RMSE of CPs between image-based georeferencing and ground truth inside our first test field “safety tunnel” amount to 62 mm and inside our second test field “SCAUT tunnel” to 74 mm. Outside the test site “SCAUT tunnel” steadily rising differences up to 540 mm occur (see Figure 10). Reasons could be erroneous image matching or extrapolation effects from 3D coordinate transformation. Outside the “safety tunnel”, we determined differences within the similar amount as outside the “SCAUT tunnel”. As for the “SCAUT tunnel”, we can mention the same error influences as in the “safety tunnel”. Moreover, the lighting conditions outside the “safety tunnel” were significantly worse.

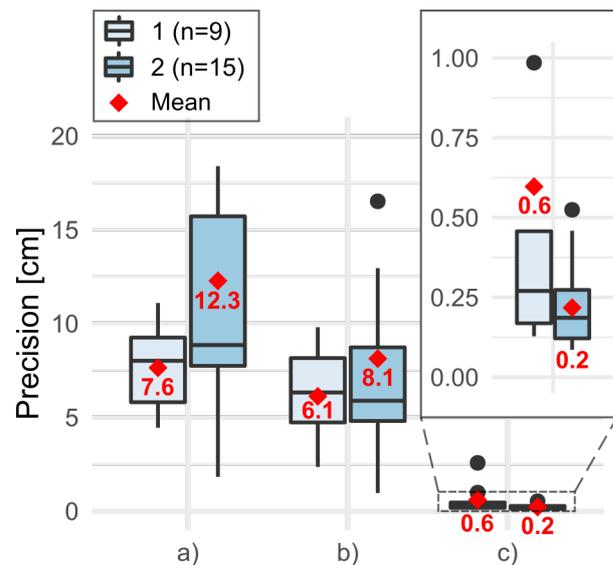


Figure 9. Boxplot with the precision of the 3D coordinate observations. Bright blue boxplot represent precisions from test site 1) “safety tunnel” and dark blue boxplot represent data set from test site 2) “SCAUT tunnel”. Precisions of datasets a) based on LiDAR SLAM-based poses processed with standard SLAM configuration and b) with optimized configuration. Precisions of datasets c) based on image-based georeferencing.

Our investigations show an accuracy increase by factor eight, from 2.83 m to 0.35 m in test field “safety tunnel” using b) SLAM poses processed with optimized parameters instead of using a) SLAM poses processed with original parameters. The improvement amounts to factor three, respectively from 0.72 m to 0.23 m in the other test field “SCAUT tunnel” (see Table 1). The reason for the improvements are eliminated systematic deviations as a result of the automatically detected loop-closures (see chapter 4.1).

After image-based georeferencing, the RMSE within both test sites range from 62 mm to 74 mm. They are similar to accuracies which have been reported with image-based street-level MMS in difficult urban environments (Blaser et al., 2017). It should be noted that the VSH offered equally challenging conditions to the BIMAGE backpack. On the one hand, difficult geometrical conditions for the LiDAR SLAM algorithm with hardly any long edges and extended planes. On the other hand, disadvantageous lighting conditions for image-based georeferencing. In Blaser et al. (2018a) we achieved better results by factor two under simpler conditions.

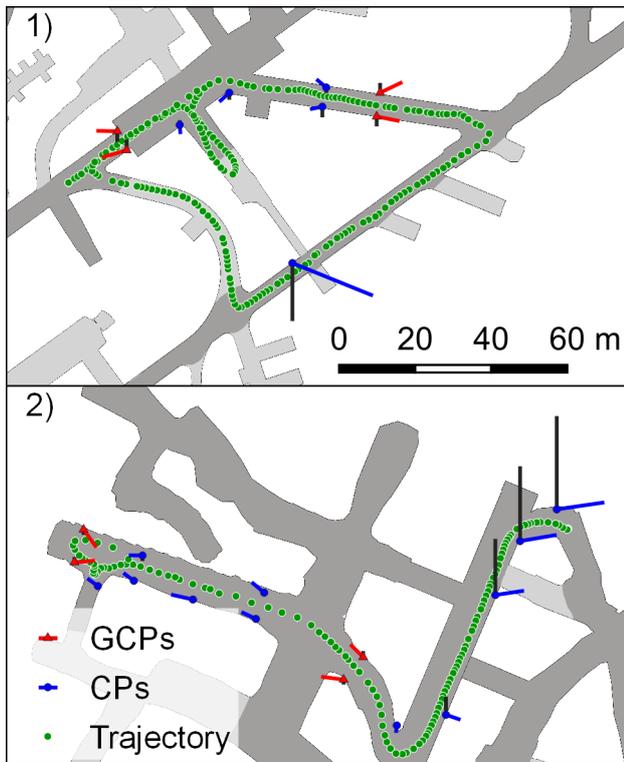


Figure 10. Overview plan of image-based georeferencing. Green points represent camera locations. We transformed all datasets into the reference frame with a 6DoF coordinate transformation using the red ground control points (GCP). Red and blue vectors represent the coordinate differences in XY-plane to the reference points. Black vectors represent the differences in altitude. All vectors are scaled by factor 50.

Method	Test site	Precision [mm]	Accuracy [mm]		
			GCPs	CPs central	All CPs
SLAM a)	1)	76	194	279	2'839
	2)	123	450	985	724
SLAM opt. b)	1)	61	135	165	347
	2)	81	259	300	233
Image-based c)	1)	6	61	62	246
	2)	2	51	74	222

Table 1. Precision and accuracy values for ground control points (GCPs), check points within the central area (CPs central) and all check points (All CPs) from both a) standard and b) optimized LiDAR SLAM-based image poses as well as from c) image-based georeferencing. The table contains results of both test sites 1) “safety tunnel” and 2) “SCAUT tunnel”. Precision indicates the mean RMSE of forward intersection of a single point measurement. Accuracy shows the RMSE of residuals to tachymetry.

5. CONCLUSION AND OUTLOOK

In this contribution, we carried out kinematic data acquisition underground with the presented portable image-based MMS *BIMAGE Backpack* in the two differently developed test sites “safety tunnel” and “SCAUT tunnel”. We examined both SLAM-based and image-based georeferencing with qualitative

and quantitative methods. 3D point clouds, resulting from the LiDAR SLAM algorithm and projected onto the XY-plane gave a first visual indication of the quality of the SLAM estimation. Using this indicator, we continually improved the SLAM configuration so that the algorithm closed the loops. For the quantitative investigations, we measured signalised target points within four consecutive images and calculated 3D coordinates with bundle adjustment-based forward intersection. The empirical standard deviation of forward intersection indicates the precision of a 3D coordinate measurement. We obtained standard deviations of 3D coordinates using SLAM-based camera poses in the range of 61 to 123 mm and using improved camera poses with image-based georeferencing in the range of 2 to 6 mm. This corresponds to an increase in relative accuracy by more than an order of magnitude. These results show that with relative measurements, e.g. 3D distances, centimetre accuracy can also be achieved in demanding underground environments. It should be noted that relative measurements cover a large part of everyday applications.

In order to compare the estimated 3D coordinates with ground truth, we transformed the 3D coordinates into the ground truth coordinate frame. We achieved an accuracy increase by factor three to eight, using the optimized SLAM camera poses. The subsequent image-based georeferencing showed a further increase in accuracy by a factor of two to four. We achieved absolute accuracies in the range of 62 to 74 mm. These accuracies are similar that those obtainable with an image-based street-level MMS in challenging urban conditions.

Our investigations show a great potential for use in various underground application areas. These range from maintenance and inventory to the creation of a digital twin using 3D geospatial images. A current development step involves the robust generation of depth maps. The additional depth layer enables 3D measurement directly in the image with just one mouse click. As a next step, we will prototypically develop interfaces to BIM. In future, existing BIM objects should be augmented in the 3D images, which allows, for example, checks for data completeness. Furthermore, we envisage to construct and edit BIM objects directly within the image-based 3D web service.

ACKNOWLEDGEMENTS

The applied research project BIMAGE (18493.2 PFES-ES) was supported by Innosuisse, the Swiss Innovation Agency and by iNovitas AG our industrial partner. Geomatics students from University of Applied Science and Arts Northwestern Switzerland (FHNW) carried out all measurements during a student project. Amberg Group supported this student project materially and financially.

REFERENCES

- Abraham, S., & Förstner, W., 2005. Fish-eye-stereo calibration and epipolar rectification. *ISPRS J. Photogramm. Remote Sens.*, 59(5), pp. 278–288. doi.org/10.1016/j.isprsjprs.2005.03.001.
- Blaser, S., Cavegn, S. & Nebiker, S., 2018a. Development of a Portable High Performance Mobile Mapping System using the Robot Operating System. In: *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, Karlsruhe, Germany, Vol. IV-1/1, pp. 13–20. doi.org/10.5194/isprs-annals-IV-1-13-2018.

- Blaser, S., Nebiker, S., & Cavegn, S., 2018b. On a Novel 360° Panoramic Stereo Mobile Mapping System. *Photogramm. Eng. & Remote Sens.*, 84(6), pp. 347–356. doi.org/10.14358/PERS.84.6.347.
- Blaser, S., Nebiker, S., & Cavegn, S., 2017. System Design, Calibration and Performance Analysis of a Novel 360° Stereo Panoramic Mobile Mapping System. In: *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, Hannover, Germany, Vol. IV-1/W1, pp. 207–213. doi.org/10.5194/isprs-annals-IV-1-W1-207-2017.
- Burkhard, J., Cavegn, S., Barmettler, A., & Nebiker, S., 2012. Stereovision Mobile Mapping: System Design and Performance Evaluation. In: *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, Melbourne, Australia, Vol. XXXIX, Part B5, pp. 453–458. doi.org/10.5194/isprsarchives-XXXIX-B5-453-2012.
- Cavegn, S., Blaser, S., Nebiker, S. & Haala, N., 2018. Robust and Accurate Image-Based Georeferencing Exploiting Relative Orientation Constraints. In: *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, Riva del Garda, Italy, Vol. IV-1/2, pp. 57–64. doi.org/10.5194/isprs-annals-IV-2-57-2018.
- Cavegn, S., & Haala, N., 2016. Image-Based Mobile Mapping for 3D Urban Data Capture. *Photogramm. Eng. & Remote Sens.*, 82(12), pp. 925–933. doi.org/10.14358/PERS.82.12.925.
- Cavegn, S., Nebiker, S., & Haala, N., 2016. A systematic comparison of direct and image-based georeferencing in challenging urban areas. In: *Int. Arch. of the Photogramm. Remote Sens. Spatial Inf. Sci.*, Prague, Czech Republic, Vol. XLI, Part B1, pp. 529–536. doi.org/10.5194/isprs-archives-XLI-B1-529-2016.
- Griesbach, D., Baumbach, D. & Zuev, S., 2014. Stereo-Vision-Aided Inertial Navigation for Unknown Indoor and Outdoor Environment. In: *8th IEEE International Conference on Indoor Positioning and Navigation (IPIN)*, Busan, Korea, pp. 709–716. doi.org/10.1109/IPIN.2014.7275548.
- Hess, W., Kohler, D., Rapp, H., & Andor, D., 2016. Real-Time Loop Closure in 2D LIDAR SLAM. In: *IEEE International Conference on Robotics and Automation (ICRA)*, Stockholm, Sweden, pp. 1271–1278. doi.org/10.1109/ICRA.2016.7487258.
- 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., El Issaoui, A., Zhu, L., Jaakkola, A. & Hyypä, J., 2017. Comparison of the Selected State-Of-The-Art 3D Indoor Scanning and Point Cloud Generation Methods. *Remote Sens.* 9(8). 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. *J. F. Robot.*, 32(4), pp. 474–503. doi.org/10.1002/rob.21531.
- Nebiker, S., 2017. 3D Imagery for Infrastructure Management – Mobile Mapping meets the Cloud. In: *The 56th Photogrammetric Week*, University of Stuttgart, Institute for Photogrammetry, Stuttgart, Germany.
- 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), pp. 2267–2291. doi.org/10.3390/ijgi4042267.
- Novak, K., 1991. The Ohio State University Highway Mapping System: The Stereo Vision System Component. In: *Proc. of the 47th Annual Meeting of the Institute of Navigation*, Williamsburg, VA, pp. 121–124.
- Nüchter, A., Borrmann, D., Koch, P., Kühn, M., & May, S., 2015. A Man-Portable, Imu-Free Mobile Mapping System. In: *Int. Ann. Photogramm. Remote Sens. Spat. Inf. Sci.*, La Grande Motte, France, Vol. II-3/W5, pp. 17–23. doi.org/10.5194/isprsannals-II-3-W5-17-2015.
- Puente, I., González-Jorge, H., Martínez-Sánchez, J., & Arias, P., 2013. Review of mobile mapping and surveying technologies. *Measurement*, 46(7), pp. 2127–2145. doi.org/10.1016/j.measurement.2013.03.006.
- 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), p. 5.
- Rechsteiner, F. & Wisler, D., 2018. Erstellung eines digitalen Zwillings des Versuchsstollens Hagerbach mittels des BIMAGE Backpacks und anderer 3D-Sensoren und Plattformen. Bachelor Thesis. FHNW University of Applied Sciences and Arts Northwestern Switzerland, Muttentz, Switzerland (unpublished).
- Schönberger, J. L., & Frahm, J.-M., 2016. Structure-from-Motion Revisited. In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, USA, pp. 4104–4113.
- 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. In: *Proceedings of the Vehicle Navigation and Information Systems Conference*, Ottawa, Canada, pp. 476–481.
- Stachniss, C., Leonard, J. J., & Thrun, S., 2016. Simultaneous Localization and Mapping. *Springer Handbook of Robotics*, pp. 1153–1176.
- Thomson, C., Apostolopoulos, G., Backes, D., & Boehm, J., 2013. Mobile Laser Scanning for Indoor Modelling. In: *Int. Ann. Photogramm. Remote Sens. Spat. Inf. Sci.*, Antalya, Turkey, Vol. II-5/W2, pp. 289–293. doi.org/10.5194/isprsannals-II-5-W2-289-2013.
- Thrun, S., 2002. Robotic mapping: A survey. *Exploring Artificial Intelligence in the New Millennium*, 1, pp. 1–35.
- VersuchsStollen Hagerbach AG, 2019. Hagerbach Test Gallery. http://www.hagerbach.ch/downloads/VSH_brochure_english_online.pdf (10. January 2019).
- Vexcel, 2018. UltraCam Panther. http://www.vexcel-imaging.com/products__trashed/ultracam-panther/ (17. January 2019).