3D Imagery for Infrastructure Management – Mobile Mapping meets the Cloud

S. Nebiker[a,b]

[a]Institute of Geomatics Engineering, FHNW University of Applied Sciences and Arts Northwestern Switzerland, Muttenz, Switzerland - stephan.nebiker@fhnw.ch
[b]iNovitas AG, Oberrohrdorferstrasse 1c, CH-5405 Baden-Dättwil, Switzerland

Photogrammetric Week 2017, Stuttgart

KEY WORDS: Mobile Mapping, 3D Imaging, Georeferencing, Image Matching, Cloud Services, Infrastructure Management

ABSTRACT:

Image-based mobile mapping systems enable the efficient acquisition of georeferenced image sequences, which can later be exploited in cloud-based 3D services. In order to provide a 360° coverage with accurate 3D measuring capabilities, a novel 360° panoramic stereo camera configuration is introduced. The paper furthermore presents investigations on georeferencing and matching strategies. The subsequent empirical tests of 3D measurements accuracies with the new 360° panoramic stereo camera configuration demonstrate relative measurement accuracies at the cm level and absolute accuracies in the range of 1 to 10 cm. The paper concludes with an overview of the broad spectrum of functionality of 3D image cloud services and of applications in the infrastructure domain.

1. INTRODUCTION

A functioning, up-to-date technical infrastructure, including e.g. utilities, traffic and communication, provides the foundation of a modern society. The enormous value of technical infrastructure and its great economic significance have been investigated, for example, in the Swiss national research programme NRP 54 "sustainable development of the built environment" (Schalcher et al., 2011). The study estimates the cost of rebuilding the technical infrastructure at 1.5 times the gross domestic product (GDP), or approx. 100'000 € per person (at 2008 prices). The yearly cost for maintaining the technical infrastructure is estimated at 3.5% of the GDP, equivalent to approx. 2'100 € per person. Digitisation and digital inventories of infrastructure assets, will play a key role in future infrastructure management – often referred to as "Infrastructure 4.0". Since the largest part of public technical infrastructure is located along traffic corridors such as roads or railways, mobile mapping technologies are ideally suited for efficiently and accurately capturing high-definition urban environments. These digital realities enable numerous measurement and inspection tasks and thus make a large part of costly and often dangerous fieldwork obsolete.

For the last 10 years, the mobile mapping market has been dominated by LiDAR-based systems. Yet, first successful mobile mapping experiments in the early 1990ies, such as the GPSVan (Novak, 1991) or the VISAT system (Schwarz et al., 1993), were all based on stereo imagery. After 10 years of LiDAR dominance, progress in imaging sensors, photogrammetric algorithms and cloud computing technologies have led to a comeback of image-based mobile mapping solutions. The Institute of Geomatics Engineering (IVGI) at University of Applied Sciences and Arts Northwestern Switzerland (FHNW) has been active in researching and developing stereovision-based mobile mapping methods and technologies since 2009 (Burkhard et al., 2012). The results of these research efforts are now in operational use for large-scale road and rail infrastructure management.

This overview paper first introduces the concept of geospatial 3D imagery and of geospatial 3D image spaces. It then addresses the following selected key issues: the design of a 360° panoramic stereo image acquisition system, image-based depth extraction strategies, image-based georeferencing strategies, measuring accuracies. The paper concludes with a discussion of functionality and applications.

2. GEOSPATIAL 3D IMAGERY

In our earlier publication (Nebiker et al., 2015) we proposed to treat geospatial 3D imagery or "Geospatial 3D Image Spaces" as a new type of a native urban model, combining radiometric and depth information. We prefer the more general term 3D imagery over RGB-D imagery, since the concept is applicable to any combination of spectral channels, including near-infrared (NIR) or thermal infrared (TIR) imagery. We furthermore postulated a number of requirements, which should be fulfilled by such a model:

- it shall provide a high-fidelity metric photographic representtation of the urban environment, which is easy to interpret and which can be augmented with existing or projected GIS data
- the RGB and the depth information shall be spatially and temporally coherent, i.e. the radiometric and the depth observation should ideally take place at exactly the same instance (this could also be expressed as WYSIWYG = what you see is what you get)
- the depth information shall be dense, ideally providing a depth value for each pixel of the corresponding RGB image
- image collections are usually ordered, e.g. in the form of images sequences, for simple navigation and shall efficiently be accessed via spatial data structures
- the model shall support metric imagery with different geometries, e.g. with perspective, panoramic or fish eye projections
- the model shall be easy-to-use and shall at least support simple, robust and accurate image-based 3D measurements using enhanced 3D monoplotting
- the model shall provide measures to protect privacy

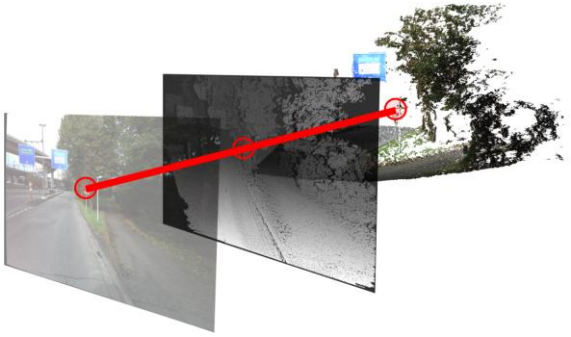


Figure 1. Principle of a geospatial 3D image with RGB image (left), co-registered dense depth map (middle) and its representation in object space as a coloured 3D point cloud (right).

3. ACQUISITION SYSTEM

Image-based mobile mapping systems have evolved from single stereo systems in the late 80ies and early 90ies (Novak, 1991; Schwarz et al., 1993) to multi-stereo systems (Cavegn & Haala, 2016; Meilland et al., 2015) or hybrid configurations consisting of stereo cameras and panorama cameras in combination with LiDAR sensors (Paparoditis et al., 2012). Specific 360° stereo-vision mobile mapping systems include the Earthmine Mars Collection System (Earthmine, 2014) with a vertical mast and four pairs of cameras, each pair forming a vertical stereo base.

The image-based mobile mapping system (MMS) (Figure 2) of the FHNW is being developed since 2009. It has evolved into a multi-stereo camera system mainly used for demanding road and rail infrastructure mapping and management projects (Burkhard et al., 2012). Our current research platform features a NovAtel SPAN inertial navigation system with a tactical grade IMU of the type UIMU-LCI and a L1/L2 GNSS kinematic antenna. In case of good GNSS coverage, these sensors provide an accuracy of horizontally 10 mm and vertically 15 mm with post-processing (NovAtel, 2012). Accuracies of roll and pitch angles are specified with 0.005° and heading with 0.008°.

Figure 2. Our mobile mapping research platform shown in a configuration with multiple stereo camera systems and a single panoramic camera.

The standard multiview stereo configuration (Figure 2) consists of a forward-looking stereo camera system with a base of 905 mm. These cameras have a resolution of 4008 x 2672 pixels (11 MP) at a pixel size of 9 µm, and a resulting field-of-view of 81° x 60°. Figure 2 also shows additional HD stereo cameras facing sideways and backwards, which were not part of the following studies.

4. NOVEL 360° PANORAMIC STEREO CAMERA SYSTEM

Our novel camera configuration approach allows for a complete 360° stereo image acquisition in heavily built-up urban environments. A forward looking and a backward looking stereo system, consisting of wide-angle pinhole cameras, cover the road with its infrastructure. In addition, two multi-head panoramic cameras tilted forward and backward by 90° each complete the configuration (see Figure 3). The individual heads of the panoramic cameras facing perpendicular to the driving direction build multiple stereo systems, which cover pavement, building façades and the entire overhead space.

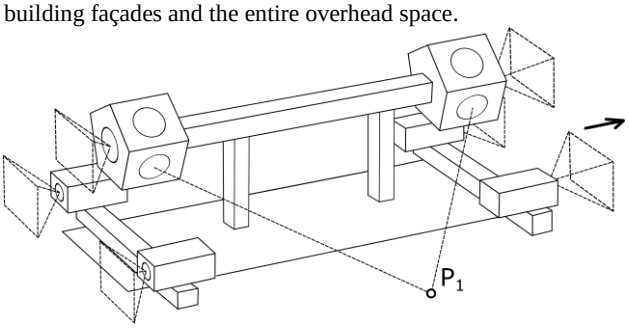


Figure 3. Novel configuration featuring several stereo camera systems for 360° RGBD mapping (patent pending)

The Ladybug5 panorama cameras are fisheye cameras for which the equidistant camera model (Abraham & Förstner, 2005) was used. The ROP between the left and right cameras of each stereo system were estimated defining the left cameras as origin. Furthermore, the ROP from each stereo system with respect to the front camera system were computed as well. In a second step, BA parameters between the front stereo system and the reference frame of the navigation system were determined in our outdoor calibration field. A detailed description of our calibration and image rectification procedure for the equidistant camera model is given in Blaser et al. (2017).

For our experiments, two tilted multi-head 360° panorama cameras of the type Ladybug5 were setup with a stereo base of 1584 mm for all the individual heads facing perpendicular to the driving direction. Each of the six camera heads of the Ladybug5 camera has a resolution of 2448 x 2048 pixels (5 MP) at a pixel size of 3.45 µm, a focal length of 4.3 mm and a field-of-view of about 113° x 94°.

5. TEST SITE AND DATA

Our study area is located at a very busy junction between five roads in the city centre of Basel, Switzerland (see Figure 4). It includes three tramway stops resulting in many overhead wires and it is surrounded by rather tall commercial properties that create a very challenging environment for GNSS positioning. Furthermore, a large number of moving objects in the form of pedestrians, cars, and tramways are typically present. The mobile mapping data for the subsequent tests were acquired in July 2014 and August 2015 (Cavegn et al., 2016; Blaser et al., 2017). In both campaigns, image acquisition intervals were in the order of 1 m between successive image exposures.

The following reference data was used in the subsequent investigations: four 360° terrestrial laser scans (TLS) using a Leica ScanStation P20 were performed in March 2015. The resulting point clouds have a 3D point accuracy of 1-2 cm. Ground control points (GCP) for image-based georeferencing were measured in March 2015, further check point coordinates for the evaluation of both relative and absolute accuracy were

determined using a total station in July 2016. Reference and check points have an absolute 3D point accuracy of < 1 cm.

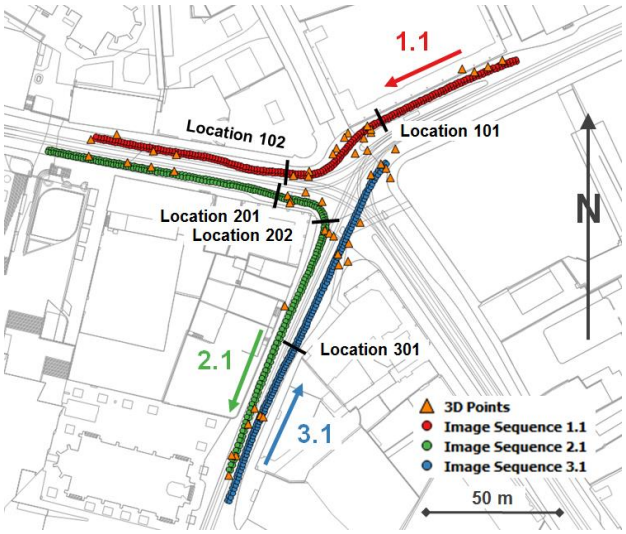


Figure 4. Base map of the test area with three different trajectories (map source: Geodaten Kanton Basel-Stadt)

6. GEOREFERENCING STRATEGIES AND RESULTS

The current state of technology is to use direct georeferencing for determining the trajectories and sensor orientations of kinematic multi-sensor platforms. Direct georeferencing – in combination with accurately calibrated multi-sensor configurations – is capable of delivering very high accuracies for sensor trajectories and for the resulting 3D point measurements. Earlier experiments by Burkhard et al. (2012) with our mobile mapping system under good GNSS coverage demonstrated, that absolute 3D point measurement accuracies of 3-4 cm horizontally and 2-3 cm vertically (1 sigma) can be achieved. However, in urban environments with poor GNSS coverage, direct georeferencing accuracies can severely degrade to several dm or m. In the case of image-based mobile mapping, the dense and highly redundant metric imagery can be used in integrated and image-based georeferencing, allowing to improve the absolute accuracies by an order of magnitude (Eugster et al., 2012).

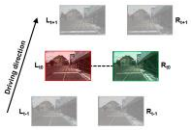
Cavegn et al. (2016) provide a systematic comparison of direct and image-based georeferencing in challenging urban environments. For the test field shown in Figure 4 they demonstrated typical direct georeferencing accuracies in the range of 10-50 cm with average residuals of 43 cm at the check-points. With their image-based georeferencing approach, based on bundle adjustment of just the forward-looking stereo imagery, the 3D point measurement accuracy was significantly improved, yielding average residuals of 4 cm.

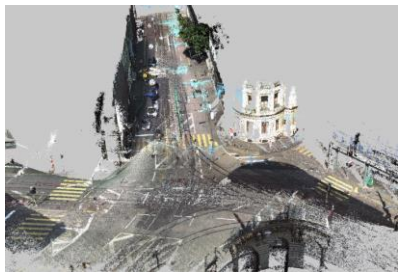
7. MATCHING STRATEGIES AND RESULTS

In the photogrammetric community, dense image matching software such as SURE (Rothermel et al., 2012) was originally developed to provide 3D point clouds or DEMs from standard airborne and close range terrestrial image blocks. In these scenarios, camera movement is primarily lateral to the viewing direction, resulting in lateral image overlaps and in stereo pairs with generally small differences in image scale. In mobile mapping scenarios, however, cameras often point in driving direction and camera motion is primarily in this direction, too.

The goal of our investigations was to extend image matching strategies to mono, stereo and multiview streetlevel image sequences. In order to exploit the potential of multi-view stereo matching in terms of higher completeness and potentially higher accuracy of the resulting depth information, existing dense image matching strategies and workflows needed to be adapted accordingly. Cavegn et al. (2015) provide an overview of matching strategies, outline the implementation of a corresponding workflow using SURE and present compare the results of the different strategies.

Configuration ‘one match image per base image’





Configuration ‘five match images per base image’

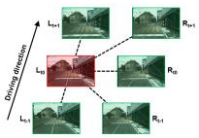


Figure 5. Matching configurations (left) with corresponding filtered and fused point clouds (right). Top: configuration ‘one match image per base image’; bottom: configuration ‘five match images per base image’

Figure 5 shows the resulting point clouds derived using the traditional stereo matching configuration (‘one match image per base image’) and a multiview matching configuration (‘five match images per base image’) of the front-looking stereo camera system. A closer look at the two point clouds reveals a significantly better completeness of the point cloud from multiview matching. A comparison with the terrestrial LiDAR reference data yields an RMSE of 14.7 cm and a mean difference of 2.3 cm for the first and a somewhat higher RMSE of 16.5 cm and mean difference of 2.7 cm for the second.

8. PERFORMANCE EVALUATION

Despite the widespread use of mobile mapping systems, there are few systematic studies on the relative and absolute 3D measurement accuracies provided by the different systems. A number of authors investigate the precision and accuracy of mobile terrestrial laser scanning systems (e.g. Barber et al., 2008; Haala et al., 2008; Puente et al., 2013). Barber et al. (2008) and Haala et al. (2008) demonstrate 3D measurement accuracies under good GNSS conditions in the order of 3 cm. Only few publications investigate the measurement accuracies of vision-based mobile mapping systems. Burkhard et al. (2012) obtained absolute 3D point measurement accuracies of 4-5 cm in average to good GNSS conditions using a stereovision mobile mapping system. In this section, we present selected results from an empirical accuracy analysis based on measured 3D points and 3D distances based on imagery captured with the new 360° panoramic stereo configuration. Detailed results can be found in Blaser et al. (2017).

8.1 Empirical 3D Measurement Accuracies

For our empirical accuracy analysis of measurements in fisheye stereo images, 16 check points on a façade (Figure 6, left) and 12 check points on a pedestrian crossing (Figure 6, right) were defined. Check points were selected as pairs at different ranges, thus forming reference 3D distances. The point pairs on the façade lie more or less at the same distance whereas the 3D distances on the pedestrian crossing vary in viewing direction.

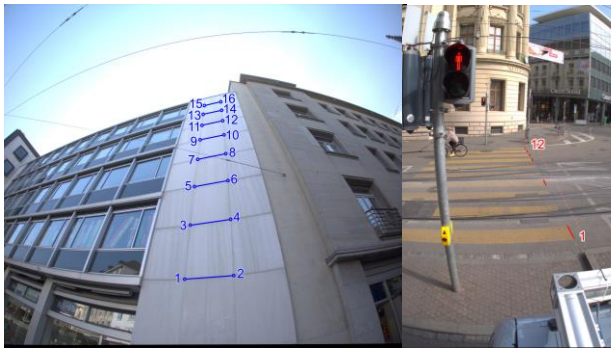


Figure 6. Layout of reference distances for the accuracy analysis (left: architectural use case, right: road infrastructure use case)

The results of the relative accuracy investigations are depicted in Figure 7. The empirical standard deviations of the distances on the façade are in the range of 0.5 to 1.5 cm while those on the pedestrian crossing vary between 0.5 and 6.5 cm. It can be seen that the standard deviation of the distances on the pedestrian crossing increases with the object distance, whereas no significant object distance-based standard deviation increase occurs for the façade. It is believed that the poorer results for the distances of the pedestrian crossing must partly be attributed to the point definition uncertainties. The absolute 3D point measurement accuracies presented in Blaser et al. (2017) were showing standard deviations in the range of 3.5 – 13.0 cm for the façade points and 1.0 – 6.5 cm.

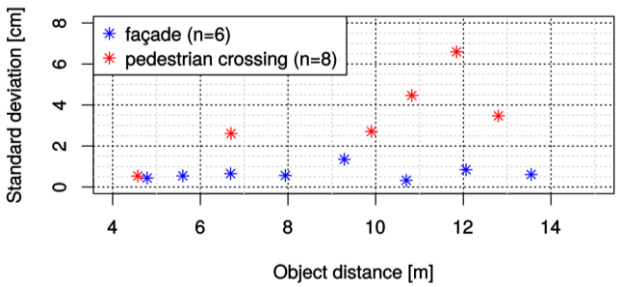


Figure 7. Accuracy of 3D distances in relation to object distance

8.2 Automated Reconstruction of 3D City Models

The main advantage of our novel camera configuration is its nearly 360° stereo coverage and its capability to capture highly overlapping image sequences. In the reconstruction experiments by Ackermann and Studer (2016), all images captured by all sensors were included. Distortion and principal point errors were previously corrected in order to facilitate the 3D reconstruction process. The city model was created with the 3D reconstruction software ContextCapture from Bentley, which also supports a fisheye camera model. The result of the automatic 3D reconstruction process was a highly detailed 3D city model shown in Figure 8. Even the detailed façades of the buildings from the classicist epoch are represented almost perfectly.

Figure 8. Samples of the reconstructed 3D city model (Ackermann & Studer, 2016)

9. FUNCTIONALITY AND APPLICATIONS OF 3D IMAGE CLOUD SERVICES

Thanks to the great progress in both Cloud and Web technologies, geospatial 3D imagery of entire states or countries can now be offered as Cloud Services. The functionality of such Cloud Services is discussed in detail in Nebiker et al. (2015). From a perspective of infrastructure management, the following functionality is particularly useful:

- 3D measurements such as distances, areas, height differences
- 3D digitizing of visible infrastructure elements
- augmentation of visible & invisible infrastructure object (e.g. for completeness checks or referencing underground objects)
- staking-out / GIS-to-field functionality using tablets, smartphones and in the future: AR glasses

Since 3D imagery can be captured with just about any mobile platform, 3D imagery cloud services can incorporate imagery captured from the air, on waterways as well as along roads (Figure 9), railways (Figure 10) etc.

Figure 9. Utilities management (Ingesa Oberland AG)

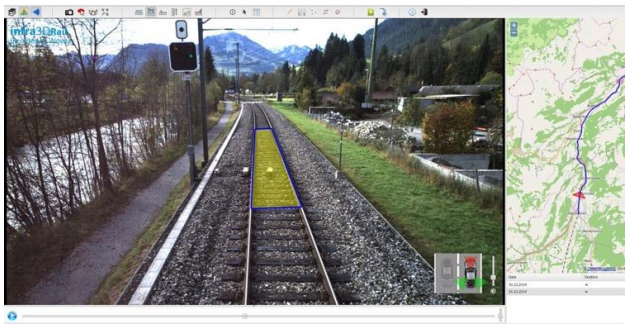


Figure 10. Rail infrastructure management (BLS)

10. CONCLUSIONS AND OUTLOOK

In this paper, we discussed a number of important issues in creating dense and accurate geospatial 3D imagery and in exploiting this 3D imagery as cloud services. Geospatial 3D imagery offers a powerful alternative to LiDAR-based (coloured) 3D point clouds. 3D imagery can be acquired with much lighter and cheaper platforms, e.g. with light-weight multicopter UAVs. The highly redundant imagery can further be used for image-based georeferencing, which yields absolute georeferencing accuracies at the level of a few centimetres – with further improvements to be expected. In 3D imagery, the radiometric and depth information is spatially and temporally coherent, which ensures the WYSIWYG (what you see is what you get) principle. 3D image services are very intuitive and easy to use – also by the large majority of non-geospatial experts. And due to simple 3D monoplotting, traditional field measurement tasks involving the identification of specific points or edges remain as simple as previously in the field.

In the domain of infrastructure management, the combination of 3D imagery and cloud services are leading to a paradigm change. The cloud services will transfer a large part of the traditional field visits and fieldwork to the office. Routine fieldwork will no longer require trained surveyors but will be carried out in the office by the infrastructure domain experts. Already now, there is a wide spectrum of functionality and applications offered by 3D image cloud services.

Our ongoing and future work includes large-scale image-based georeferencing with the exploitation of calibrated multi-view camera configurations for further accuracy and performance improvements. We will further pursue the application of Machine Learning and Deep Learning to additional information extraction tasks towards high-end scene understanding. Last but not least, geospatial 3D imagery also have a great potential for indoor mapping and Building Information Management (BIM). Here, we are investigating the combination of LiDAR SLAM and advanced image-based georeferencing approaches for the generation of accurate and robust indoor 3D image spaces.

ACKNOWLEDGEMENTS

This work was financially supported by the Swiss Commission for Technology and Innovation CTI as part of the projects infraVIS and BIMAGE.

REFERENCES

Abraham, S. & Förstner, W., 2005. Fish-eye-stereo calibration and epipolar rectification. *ISPRS J. Photogramm. Remote Sens*, 59(5), pp. 278–288.

Ackermann, R. & Studer, D., 2016. Stereo-Panorama-Aufnahmen im urbanen Raum - Untersuchungen mit einem neuen 360°-Stereo-Panoramasystem. Bachelor Thesis. FHNW University of Applied Sciences and Arts Northwestern Switzerland, Muttenz, Switzerland (unpublished).

Barber, D., Mills, J. & Smith-Voysey, S., 2008. Geometric validation of a ground-based mobile laser scanning system. *ISPRS J. Photogramm. Remote Sens.*, 63(1), pp. 128–141.

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. Spat. Inf. Sci.* IV-1/W1, 207–213.

Burkhard, J., Cavegn, S., Barmettler, A. & Nebiker, S., 2012. Stereovision Mobile Mapping : System Design and Performance Evaluation. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.*, Melbourne, Australia, Vol. XXXIX, Part B5, pp. 453–458.

Cavegn, S., Haala, N., Nebiker, S., Rothermel, M., Zwölfer, T., 2015. Evaluation of Matching Strategies for Image-based Mobile Mapping. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* II-3/W5, 361–368.

Cavegn, S. & Haala, N., 2016. Image-Based Mobile Mapping for 3D Urban Data Capture. *Photogramm. Eng. Remote Sens.*, 82(12), pp. 925–933.

Cavegn, S., Nebiker, S. & Haala, N., 2016. A Systematic Comparison of Direct and Image-Based Georeferencing in Challenging Urban Areas. In: *Int. Arch. Photogramm. Remote Sens.*, Prague, Czech Republic, Vol. XLI, Part B1, pp. 529–536.

Earthmine, 2014. Earthmine Mars Collection System. www.earthmine.com/html/products_mobile.html (27.3.17).

Eugster, H., Huber, F., Nebiker S. & Gisi, A., 2012. Integrated Georeferencing of Stereo Image Sequences Captured with a Stereovision Mobile Mapping System – Approaches and Practical Results from a Practical Perspective. In: *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, Melbourne, Australia, Vol. XXXIX, Part B1, pp. 309–314.

Haala, N., Peter, M., Kremer, J. & Hunter, G., 2008. Mobile LiDAR Mapping for 3D Point Cloud Collection in Urban Areas - A Performance Test. In: *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, Beijing, China, Vol. XXXVII, Part B5. pp. 1119–1124.

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.

Nebiker, S., Cavegn, S., Eugster, H., Laemmer, K., Markram, J. & Wagner, R., 2012. Fusion of Airborne and Terrestrial Image-based 3D Modelling for Road Infrastructure Management – Vision and First Experiments. In: *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, Melbourne, Australia, Vol. XXXIX, Part B4, pp. 79–84.

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.

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.

NovAtel, 2012. Span-SE User manual. No. OM-20000124. Calgary, Alberta, Canada.

Paparoditis, N., Papelard, J.-P., Cannelle, B., Devaux, A., Soheilian, B., David, N. & Houzay, E., 2012. Stereopolis II: A Multi-Purpose and Multi-Sensor 3D Mobile Mapping System for Street Visualisation and 3D Metrology. *Rev. française photogrammétrie télédétection*, 200, pp. 69–79.

Puente, I., González-Jorge, H., Martínez-Sánchez, J. & Arias, P., 2013. Review of mobile mapping and surveying technologies. *Measurement*, 46(7), pp. 2127–2145.

Rothermel, M., Wenzel, K., Fritsch, D. & Haala, N., 2012. SURE - Photogrammetric Surface Reconstruction from Imagery. In: *Proceedings LC3D Workshop*, Berlin, pp. 1–21.

Schalcher, H.-R., Boesch, H.-J., Bertschy, K., Sommer, H., Matter, D., Gerum, J., Jakob, M., 2011. Was kostet das Bauwerk Schweiz in Zukunft und wer bezahlt dafür? [Fokusstudie NFP 54]. vdf, Zürich.

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, pp. 476–481.