AUTOMATIC IMAGE-BASED 3D RECONSTRUCTION STRATEGIES FOR HIGH-FIDELITY URBAN MODELS - COMPARISON AND FUSION OF UAV AND MOBILE MAPPING IMAGERY FOR URBAN DESIGN STUDIES

W. Wahbeh¹*, G. Müller¹, M. Ammann¹, S. Nebiker¹

¹ Institute of Geomatics, FHNW University of Applied Sciences and Arts Northwestern Switzerland, 4132 Muttenz, Switzerland - (wissam.wahbeh, manuela.ammann, stephan.nebiker)@fhnw.ch, gianna.mueller@hotmail.com

Commission II, WG 4

KEY WORDS: Urban scene reconstruction, Photogrammetry, Mobile mapping, UAV, 3D modelling.

ABSTRACT:

Modelling environments for urban projects often involve virtual reconstructions of existing urban areas where generalised city models are not sufficient. The use of photogrammetry with aerial imagery from UAVs on the one hand, and imagery from street-level mobile mapping systems on the other, has advanced significantly in recent years. However, there are limitations when these two imagery types are used separately.

The main contribution of this paper is an end-to-end solution for creating large-scale and complex reality-based mesh models of urban environments. We outline a novel image-based 3D reality capturing process by combining street-level imagery from a backpack multi-camera mobile mapping system and nadir and oblique UAV imagery. This reconstruction is based on dense multi-view image matching method. A use case is presented with a goal to reconstruct an area for an urban project which aims to introduce an elevated cycle highway to facilitate traffic. The investigated area has a size of 0.64 square kilometres and is in the city of Basel in Switzerland. The MMS and UAV images were all oriented in a single photogrammetric project. Several reconstructions with different configurations were investigated and the results of the successful reconstruction with a geometric accuracy of a few cm are presented.

1. INTRODUCTION

Urban design is still mainly based on two-dimensional drawings, discounting many urban details that might influence the design and perception of the project. Virtual 3D modelling supporting the design is as important in urban projects as in architectural projects which are nowadays more often based on 3D reality capturing or building information models (BIM). Modelling environments for urban projects often involve virtual reconstructions of existing urban areas where generalised city models are not sufficient. However, creating urban-scale detailed reality-based models is a challenging task. Yet, realitybased mesh models are gaining increased interest, supported by software and hardware advancements. There are several applications for the use of these models, especially in the field of virtual reality to create walkable virtual cities (Schmohl et al., 2020) or for creating mobility planning simulations (Wahbeh et al., 2021).

Urban reconstruction methods and input data types are manifold. And despite the high volume of existing work, there are still many unsolved problems, especially when it comes to the development of fully automatic algorithms (Musialski et al., 2013). The use of photogrammetry with aerial imagery from UAVs on the one hand, and imagery from street-level mobile mapping systems on the other, has advanced significantly in recent years. However, there are limitations in terms of visibility and subsequently the completeness of the resulting model when these two imagery types are used separately. Recently, researchers have been trying to combine aerial oblique imagery and terrestrial imagery to optimise the urban models (Berrett et al., 2021; Toschi et al., 2017; Wu et al., 2018). The main contribution of this paper is an end-to-end solution for creating large-scale and complex reality-based mesh models of urban environments. We outline a novel image-based 3D reality capturing process by using a) street-level imagery from a backpack multi-camera mobile mapping system (MMS) developed inhouse (Blaser et al., 2018) and b) nadir and oblique UAV imagery. We compare the resulting models from UAV aerial imagery only, and from the fusion of UAV and MMS imagery. Based on comprehensive comparisons of the different modelling strategies we discuss the procedures in terms of geometry and texturing.

2. RELATED WORK

The motivation of the study presented in this paper is to support an urban planning project in Basel city with the goal of a raised bike highway in a complex area, which should facilitate the future circulation of cycling traffic (figure 1). This research is to be considered as further development of our related study in 2020 (Wahbeh et al., 2021). In the mentioned study we obtained highly detailed and accurate reconstructions of an urban environment, using only street-level mobile mapping images. This earlier study demonstrated the great potential of automatic reconstructions from street-level imagery but also showed the need for complementary aerial imagery in complex urban scenarios.

Over the last few decades, 3D reconstruction from image data has developed considerably. Different platforms such as aerial (Haala et al., 2015; Kang et al., 2019), UAV (Haala et al., 2011; Li et al., 2016; Schmohl et al., 2020) or mobile mapping imagery (Blaser et al., 2018; Wu et al., 2018) have been investigated. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLIII-B2-2022 XXIV ISPRS Congress (2022 edition), 6–11 June 2022, Nice, France



Figure 1. Reality-based 3D urban scene automatically reconstructed from MMS and UAV imagery with augmentation of planned bike highway.

The previous paper (Wahbeh et al., 2021) demonstrated the power of street-level MMS data for 3D reconstruction of urban environment. It also revealed some limitations because buildings could not be fully reconstructed as the roofs are often not visible from the street. The example of Stuttgart City Walk shows the high potential of UAV images for reconstruction (Schmohl et al., 2020). But reconstructions from UAV images also have their limitations, especially in the representation of the street space, in particular trees, street furniture, or balconies. For a complete coverage of an area, different shooting angles improve the texturing and modelling of urban scenarios (Wu et al., 2018). The MMS data complemented the aerial imagery with more detailed views of the facades, benefiting the reconstruction tremendously. And the aerial imagery was able to solve the problem of GNSS tracking loss for MMS in urban areas, as well as more accurately model the roofs that were difficult to see. The integrated georeferencing of airborne and ground-based stereo imagery was investigated by Nebiker et al. (2013) because direct georeferencing of MMS in urban areas is usually challenging due to limited GNSS reception. They were able to reduce the systematic error of the georeferencing of the MMS to approx. 2-2.5 cm in position and 1 cm in height with a co-registration of the airbone images. This accurate coregistration of aerial and terrestrial imagery is a key issue, when aiming at 3D reconstructions from both data sources.

3. MATERIALS AND METHODS

3.1 Urban Modelling Strategy

This study is based on dense multi-view image matching (DIM) which involves different sets of images from different cameras, lense types and camera models, points of view and distances to the object. The objective is to combine these image sets (figure 2) into a single photogrammetric reconstruction project. Earlier, unpublished attempts to connect nadir and MMS imagery had led to incomplete reconstructions of the image poses and thus to incomplete 3D reconstructions. The goal of using 45° oblique UAV images is to connect the 360° street level imagery with the nadir UAV imagery. The intermediate views including facades and roofs, strengthen the network geometry by dramatically increasing the number of tie points and by reducing strong viewpoint changes. The combination of terrestrial, oblique and nadir imagery is furthermore expected, to significantly improve the 3D reconstruction of roof edges, façade details and vegetation.



Figure 2. Simplified representation of the configuration of the three types of imagery used. Nadir UAV imagery, Oblique (45°) UAV imagery, and fisheye street-level imagery from mobile mapping system.

The photogrammetric reconstruction including the different image sets and ground control points produces accurately georeferenced images and a very detailed mesh model. The model is further processed outside the photogrammetric project to be improved and adapted to the project requirements and then (re-)textured from the georeferenced images. Figure 3 shows the outline of the workflow adopted.



Figure 3. Reconstruction and modelling workflow for urban 3D models from MMS, oblique and nadir UAV imagery

3.2 Data Capturing Systems

For a complete coverage of an area, we use our BIMAGE MMS with a multi-head 360° camera for street-level imagery and a UAV for nadir and oblique imagery.

The BIMAGE Backpack includes a multi-head panoramic camera FLIR Ladybug 5, GNSS- and IMU-based navigation unit, NovAtel SPAN CPT7, with tactical grade performance and two multi-beam LiDAR scanners Velodyne VLP-16. Thus, the system has state-of-the-art and high-end sensors, supports precise sensor synchronization and is accurately calibrated using state-of-the-art calibration techniques (Blaser et al., 2020). The BIMAGE backpack has proven useful for street-level urban 3D modelling in the work of Wahbeh et al. 2021).



Figure 4. BIMAGE Backpack (Blaser et al., 2018).

The BIMAGE Backpack, as most MMS, supports direct georeferencing based on GNSS and IMU measurements. The system also supports advanced georeferencing methods, either based on LiDAR SLAM or employing image-based techniques. With state-of-the-art image-based georeferencing, the BIMAGE Backpack can achieve accuracies of a few cm even in city centres (Blaser et al., 2020).

The DJI Phantom 4 Pro is a quadcopter with a single frequency GNSS receiver and a consumer-grade Micro-Electromechanical System IMU for navigation based on predefined flight paths. The UAV includes the DJI FC6310 camera with an 8.8 mm nominal focal length, and a 1" CMOS 20 megapixel sensor with 2.41 x 2.41 µm nominal pixel size (DJI, 2022).



Figure 5. DJI Phantom 4 Pro (DJI, 2022).

3.3 Photogrammetric 3D scene reconstruction

3.3.1 SfM software and pipeline

The reconstruction of very large imagery datasets, acquired from different perspective with multiple cameras and different sensor models is very challenging. Thus, a suitable software needs to fulfil a number of requirements, namely:

- the support for different camera models, including fisheye models for the panoramic camera of the MMS
- georeferencing with ground control points
- masking of obstructed image areas
- ideally, also camera rig constraints to strengthen the image orientation process

In our case we used the commercial Structure-from-Motion (SfM) Software ContextCapture by Bentley for image orientation, georeferencing and 3D reconstruction. It supports most of the requirements listed above, including masking of the imagery to eliminate interfering objects like the frame of the

Backpack. The software can also combine different camera models such as the pinhole model for the DJI camera and fisheye models for the BIMAGE cameras, which are not supported by many SfM softwares.

Due to the integrated GNSS and IMU in the BIMAGE Backpack, image poses from direct georeferencing can be imported and introduced into the bundle adjustment.

Since the 3D reconstruction of very large image data sets is very computationally intensive, we used a high-end workstation with two 24 Core CPUs, 512 GB RAM and a Nvidia GeForce RTX 3090 graphics card.

3.3.2 Image orientation and georeferencing

The image orientation and georeferencing was carried out with three different image constellations. Once with the UAV images only, once with the MMS images only and finally using a combination of UAV and MMS. The calculation of the image orientation and georeferencing in ContextCapture can be controlled via "adjustment constraints" and "final rigid registrations". These two computations can be done using the control points or the image pose from the directed georeferencing. Three different strategies were investigated for exploiting the features of 'adjustment constraints' and 'final rigid registration'. One strategy was to use the image poses for 'adjustment constraints' and the control points for the 'final rigid registration'. The next strategy was to use the image poses and the control points for 'adjustment constraints' and the control points for the 'final rigid registration'. The third strategy was to use only control points for the 'adjustment constraints', thus no using image poses as observables. The strategies are summarized in table 1:

	Adjustment Constraints	Final Rigid Registrations
Variant 1	IP	CP
Variant 2	IP + CP	CP
Variant 3	СР	

Table 1. Summary of different settings with image pose from the directed georeferencing (IP) and control points (CP)

3.3.3 3D reconstruction

The large-scale 3D reconstruction process using all imagery (terrestrial, oblique and nadir) is also carried out in ContextCapture. The masks are used in order to avoid interfering objects in the tie points, in the following reconstruction, and in the subsequent texturing process. For the reconstruction, different accuracy levels can be selected. We chose the geometric accuracy of 1 pixel. All other settings were kept in default.

3.4 3D Urban Scene Modelling

Dense point clouds can be heavily affected by poor image quality or textureless areas, resulting in high frequency noise, holes, and uneven point density. These issues can be propagated during the mesh generation process. (Nocerino et al., 2020). Therefore, the generation of 3D mesh models of geometrically critical, and texture-poor objects, results in noisy meshes and some detached and isolated groups of polygons. This is often the case with very thin objects such as hanging cables, poles or possibly moving objects captured in several photos.

When the model is large, which is normally the case in urban reconstructions, a reconstruction of the model in parts can be recommended to optimise the calculation process. In order to split the reconstruction area, it is advisable to ensure sufficient overlap so that there are no uncovered parts.

463

Consequently, the photogrammetric reconstruction results in a rough mesh that should be edited to produce a geometrically clean model as the basis for an urban planning proposal.

Photogrammetry software normally supply tools to clean the mesh by performing basic operations. However, the use of external software dedicated to 3D modelling offers many more possibilities, especially when it comes to replacing parts of the mesh with new structured models, adding missing parts, optimise mesh subdivision and improving its topology.

Unfortunately, most software dedicated to architectural, urban and BIM 3D modelling loses accuracies by using large coordinates. Therefore, a transformation to a temporary local coordinate system is essential when editing the model with 3D modelling software.

Modelling steps are highly dependent on the type and goal of modelling, such as the integrity of the mesh, the level of detail required, the smoothness of the surface, and the optimisation of the size of the model by resampling the mesh. Different modelling steps could be executed to improve the geometrical quality of the model before getting it textured building the texture in the photogrammetric project. The workflow including some fundamental steps to clean the mesh is illustrated by the graph in figure 6.



Figure 6. Workflow shows the main steps to improve the model geometry using 3D modelling software (continues line boxes)

4. CASE STUDY

4.1 Test location

The area under investigation in this study is concerned with an urban project to introduce an elevated cycle highway to facilitate traffic. The area has a size of 0.64 square kilometres and is located in the city of Basel in Switzerland. The objective of the reconstruction is to provide a metric basis for the design and public communication of the project since the public opinion is to be involved in the decision to approve such a project.

4.2 Data Acquisition

The data acquisition took place on the 22nd of October 2021. It consisted of several mobile mapping campaigns shown in figure 8 and several UAV flight missions illustrated in figure 9.

The mobile mapping campaigns were carried out using the BIMAGE backpack (Blaser et al., 2018) mounted on our electric mobile mapping system (eMMS) (figure 7). The survey route was about 7.5 km in length. This included 4.5 km on roads (orange trajectories in figure 8) and 3 km on footpaths and sidewalks (yellow trajectories in figure 8). In around 1.5 hours, a total of 27'600 fisheye images had been captured with the Ladybug multi-head panoramic imaging system. The average distance between subsequent image acquisition positions is around 2.5 m. The spatial resolution of the imagery is approx. 0.5 cm/pixel at 5 m distance.



Figure 7. BIMAGE backpack mounted on our electric mobile mapping system (eMMS)

The following UAV missions were carried out using a DJI Phantom 4 with 2.7 cm ground sampling distance for nadir images and around 4 cm for oblique images. The overlap between the UAV images was about 80%. The different flight missions are illustrated in figure 9 and include:

- Two nadir flight missions (figure 9, top left) with a GSD of 2.7 cm
- Five oblique flight missions with an image inclination of 45° and an average GSD of around 4 cm. The focus of the first two oblique mission (figure 9, top right) was the central bridge, crossing some 16 rail tracks. The focus of the other oblique missions (figure 9, bottom left and bottom right) were the main and prominent building facades along the planned future bike highway

Overall, 811 nadir and oblique RGB images had been acquired in the seven flight missions.



Figure 8. Mobile mapping campaigns with the BIMAGE backpack: vehicle-based acquisition along the roads (orange trajectories) and along footpaths and sidewalks (yellow trajectories). GNSS control and check points are shown as red markers.



Figure 9. UAV flight missions. The lines describe the path, and the coloured surface is the photographed area. In the top left image: UAV flight with nadir view, the remaining images: UAV flight with oblique view (45°).

In addition, 32 natural control and check points (red markers in figure 8) were recorded on 17^{th} of March 2022 with the Leica GNSS GI14. The GNSS points have an 3D position accuracy of 1.8 cm (max. 4.5 cm).

Exposures and Images UAV	811
Exposures MMS	4'600
Images MMS	27'600
Control Points	20
Check Points	12

 Table 2. Summary of the mapping campaign with the MMS
 UAV and GNSS

All in all, there are 28'411 images from the Backpack and the UAV in this use case, although we did not use all of them for the reconstruction. The exact numbers can be seen in section 4.3.

4.3 Image orientation and georeferencing

As we restricted the area, only 8'389 images were required. Thus, 810 UAV images and 7'579 MMS images were used for further processing. The check and control points vary depending on the variant, because certain points were only visible from one platform (listed in table 3). The check and control points are distributed throughout the area for complete coverage.

	# images	# check points	# control points
UAV	810	13	8
MMS	7'579	13	11
UAV+MMS	8'389	20	12

 Table 3. Summary of processing time

The results of the georeferencing are described in section 5.

4.4 3D Reconstruction

We compare the resulting models using different combinations of images. The investigated modelling scenarios include UAV imagery only, MMS imagery only, and fusion of both data sources. Based on comprehensive comparisons of the different modelling strategies we subsequently propose a workflow to optimise the models in terms of geometry and texturing to be suitable for visualisations and immersive VR experience. Moving objects like cars or pedestrian are automatically filtered through the use of highly redundant imagery.

With the accuracy level of 1 pixel for reconstruction, the geometric accuracy of the combined model ranges from 0.10 to 46 cm per pixel with an average of 2.44 cm. The more images and cameras are processed, the longer it takes to process them. The processing time of the UAV was approx. 6 hours: 9 min for alignment and 5 hours for the 3D reconstruction. For the MMS data the processing time was 22 hours: 1 hour for the alignment and 21 hours for the 3D reconstruction. For the combination of MMS and UAV, overall processing time was 98 hours: 2 hours for the alignment and 96 hours for the 3D reconstruction.

	# images	Aligning [min]	Reconstruction [h]
UAV	810	9	5
MMS	7'579	61	21
UAV+MMS	8'389	65	33

 Table 4. Summary of processing time

4.5 3D Model

The best models of the three scenarios were imported into the modelling software for analysis. from the first comparison it became immediately clear that the most suitable model is the one made by combining UAV and MMS images from the point of view of completeness and level of detail. Figure 10 shows a comparison of the imported raw models. The reconstruction has produced models with a very high geometric accuracy and level of detail. The model was refined semi-automatically and manually adapted to the needs of the project for immersive visualisation using the modelling software Autodesk 3dsMax. Figure 11 illustrate some model optimisation steps.



Figure 10. Illustrations of different 3D reconstructions of an identical sub-scene. Left: fusion UAV and MMS images, middle: UAV images only, right: MMS images only. From top to bottom: the distribution of the images on the plan, textured mesh front view, and a wireframe top view.



Figure 11. The mesh after import into the modelling software. From left to right. a few steps to obtain a clean model well bounded from the fragmented mesh with detached polygons.

5. RESULTS AND DISCUSSION

5.1 Results

The MMS and UAV images were all oriented in a single photogrammetric project. Several reconstructions with different configurations were carried out during the study. After several trials it was determined that the image orientation and georeferencing using only the control points for adjustment constraints gave the best results. The model with the best control point root mean square error (RMSE) combining UAV and MMS images produced a 3D mesh with an average resolution in object space of 1.9 cm/pixel. UAV imagery only produced an average resolution of 3.6 cm/pixel and MMS only an average resolution of 1.8cm/pixel. The reprojection error and RMSE of this reconstruction are listed in table 5.

The textured 3D mesh was produced from the best model. After some manual editing to clean it up, it was textured using the same oriented images from the three image sets used for the reconstruction. Figure 12 shows a rendering of the final model. In this rendering no additional direct lighting effects are used, therefore the lighting and shadows visualised are the real lighting effects at the time of the data acquisition.

	Reprojection error [pix]	Control Points RMSE [cm]	Check Points RMSE [cm]
UAV	0.57	1.69	4.88
MMS	0.68	1.34	3.73
UAV + MMS	0.63	1.76	3.48





Figure 12. The resulting textured reality mesh model of the entire project area (coloured) extended with the simplified city model (white).

5.2 Models Comparison

In order to compare complete models, we have excluded the model reconstructed exclusively from MMS images as it only produces models of the streets and facades but not the roofs and courtyards which were not specifically visited and photographed by this system. In the following paragraphs, we distinguish between the comparison in terms of model geometry, model texture and size of the produced file.

5.2.1 Geometry

The combined reconstruction using UAV and MMS imagery has produced models with very high geometric accuracy and level of detail. Figure 13 illustrates two different 3D reconstructions of identical sub-scenes resulting from different data sets. The combination of UAV and street level imagery has proved essential in most cases to provide a complete and detailed 3D reconstruction where the combination is possible.

This shows that the mesh has much more detail in the main street zone where MMS images are used for the reconstruction. This also demonstrates that the model is much less homogeneous than reconstruction using UAV images only. The combination of the different datasets produced much more detail in the areas with high image density (figure 14). For the objective of this study, i.e., modelling for urban design, it is an ideal solution because it optimizes the size of the model without sacrificing completeness. The intervention of the urban project interested in this reconstruction will be in the streets where the survey is performed with MMS. Therefore, we get much more geometric detail that is also needed for close-up views.



Figure 13. Edged faces renderings of different 3D models of identical sub-scenes. Top: reconstruction using only UAV imagery, bottom: reconstruction using UAV and MMS imagery. The darker is the zone the more details it has since the edges are black and the mesh faces are white.



Figure 14. Edged faces renderings of different 3D models of identical sub-scenes at street level. Top: reconstruction using only UAV imagery, bottom: reconstruction using UAV and MMS imagery.

5.2.2 Texture

The vast difference in lighting conditions between images taken by UAV and those acquired at street level produced a high contrast that resulted in a poorly displayed texture after merging. Different blending methods such as 'average' did not help to achieve a realistic result. this problem is evident in the facades of buildings. whereas for trees the texturing has been greatly improved with the combination of texture in addition to the highly detailed geometric reconstruction. Figure 15 compares two views of the UAV model and the UAV+MMS model.



Figure 15. Illustrations of different textured 3D models of identical sub-scenes. Scenes 1a and 2a: UAV only, scenes 1b and 2b: fusion UAV and MMS (from street and sidewalks).

5.2.3 Scene Size

Comparing two models in terms of the number of polygons is dependent on the chosen zone and the concentration of the images used in the reconstruction in this zone. So, it can vary a lot from one part of the mesh to another. As an example, for comparison, an area was identified where a survey was carried out with MMS and UAV even though the MMS is limited to two roads (figure 16). For this portion of the model, the mesh based on UAV imagery contains approximately 1.3 million polygons and the mesh based on the UAV MMS combination contains 10.5 million polygons. This shows that including MMS images drastically increases the number of polygons as can be seen visually by the fine details in figures 13 and 14.



Figure 16. Size comparison example: the same mesh portion of different reconstructions 260x170m. Up: UAV-based mesh: 1.3 million polygons / 650 K vertices. Bottom: UAV+MMS-based mesh: 10.5 million polygons / 5.3 million vertices.

6. CONCLUSION & OUTLOOK

We have successfully reconstructed a georeferenced 3D model from UAV and MMS images with a geometric accuracy of a few cm. with this combination it is possible to cover the area at different angles and thus eliminate the limiting factors of these two sets of images separately in the case of reconstruction at urban level. To ensure a successful alignment, a requirement is to achieve a sufficient overlap between the aerial and ground images as well. This represents a challenge in the planning of the survey campaign. The oblique images were crucial for the process. The flight had to be planned well in a way to avoid extreme heights where the facades should not be clearly visible occupying an important part of the image. A challenge which remains is the texturing of the model. Texture generation was successful; however, the quality of the texture is very questionable in terms of colour. Mixing images at very different exposures caused this problem. The UAV images are much better as the exposure is constantly good without back light problems, but they are obviously not enough to texture the parts covered by trees and other urban objects. Based on the results, the next investigations we intend to carry out are those of the texturing methods and the optimisation of the model in terms of size and quality of detail.

REFERENCES

Berrett, B.E., Vernon, C.A., Beckstrand, H., Pollei, M., Markert, K., Franke, K.W., Hedengren, J.D., 2021. Large-Scale Reality Modeling of a University Campus Using Combined UAV and Terrestrial Photogrammetry for Historical Preservation and Practical Use. Drones 5, 136. https://doi.org/10.3390/drones5040136

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. https://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. https://doi.org/10.5194/isprs-annals-V-1-2020-333-2020

DJI, 2022. Phantom 4 Pro [WWW Document]. URL https://www.dji.com/ch/phantom-4-pro

Haala, N., Cramer, M., Weimer, F., Trittler, M., 2011. Performance Test on UAV-Based Photogrammetric Data Collection. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci. XXXVIII-1/C22, 7–12. https://doi.org/10.5194/isprsarchives-XXXVIII-1-C22-7-2011

Haala, N., Rothermel, M., Cavegn, S., 2015. Extracting 3D urban models from oblique aerial images, in: 2015 Joint Urban Remote Sensing Event (JURSE). Presented at the 2015 Joint Urban Remote Sensing Event (JURSE), pp. 1–4. https://doi.org/10.1109/JURSE.2015.7120479

Kang, J., Lee, S., Yeon, S., Park, S., 2019. Improving 3D Mesh Quality Using Multi-Directional UAV Images. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci. XLII-2/W13, 419–423. https://doi.org/10.5194/isprs-archives-XLII-2-W13-419-2019

Li, M., Nan, L., Smith, N., Wonka, P., 2016. Reconstructing building mass models from UAV images. Computers & Graphics 54, 84–93. https://doi.org/10.1016/j.cag.2015.07.004

Musialski, P., Wonka, P., Aliaga, D.G., Wimmer, M., van Gool, L., Purgathofer, W., 2013. A Survey of Urban Reconstruction: A Survey of Urban Reconstruction. Computer Graphics Forum 32, 146–177. https://doi.org/10.1111/cgf.12077

Nebiker, St., Kamer, K., Eugster, H., 2013. Integrierte Georeferenzierung von luft- und bodengestützten Stereobilddaten im Straßenbereich – Strategien und Ergebnisse, in: Band 22. Presented at the Wissenschaftlich-Technische Jahrestagung der DGPF – Dreiländertagung DGPF, E. Seyfert, Freiburg.

Nocerino, E., Stathopoulou, E.K., Rigon, S., Remondino, F., 2020. Surface Reconstruction Assessment in Photogrammetric Applications. Sensors 20, 5863. https://doi.org/10.3390/s20205863

Schmohl, S., Tutzauer, P., Haala, N., 2020. Stuttgart City Walk: A Case Study on Visualizing Textured DSM Meshes for the General Public Using Virtual Reality. PFG 88, 147–154. https://doi.org/10.1007/s41064-020-00106-z

Toschi, I., Ramos, M.M., Nocerino, E., Menna, F., Remondino, F., Moe, K., Poli, D., Legat, K., Fassi, F., 2017. Oblique Photogrammetry Supporting 3D Urban Reconstruction of Complex Scenarios. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci. XLII-1/W1, 519–526. https://doi.org/10.5194/isprs-archives-XLII-1-W1-519-2017

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. ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci. V-4–2021, 225–232. https://doi.org/10.5194/isprs-annals-V-4-2021-225-2021

Wu, B., Xie, L., Hu, H., Zhu, Q., Yau, E., 2018. Integration of aerial oblique imagery and terrestrial imagery for optimized 3D modeling in urban areas. ISPRS Journal of Photogrammetry and Remote Sensing 139, 119–132. https://doi.org/10.1016/j.isprsjprs.2018.03.004