

HYBRID GEOREFERENCING, ENHANCEMENT AND CLASSIFICATION OF ULTRA-HIGH RESOLUTION UAV LIDAR AND IMAGE POINT CLOUDS FOR MONITORING APPLICATIONS

N. Haala^{1*}, M. Kölle¹, M. Cramer¹, D. Laupheimer¹, G. Mandlbürger² & P. Glira³

¹ Institute for Photogrammetry, University of Stuttgart, Germany – (norbert.haala, michael.koelle, michael.cramer, dominik.laupheimer)@ifp.uni-stuttgart.de

² TU Wien, Department of Geodesy and Geoinformation, Wien, Austria – Gottfried.Mandlbürger@geo.tuwien.ac.at

³ AIT Austrian Institute of Technology - Philipp.Glira@ait.ac.at

KEY WORDS: UAV-based LiDAR, Dense Image Matching, Hybrid Adjustment, Classification, Deformation Monitoring

ABSTRACT:

This paper presents a study on the potential of ultra-high accurate UAV-based 3D data capture by combining both imagery and LiDAR data. Our work is motivated by a project aiming at the monitoring of subsidence in an area of mixed use. Thus, it covers built-up regions in a village with a ship lock as the main object of interest as well as regions of agricultural use. In order to monitor potential subsidence in the order of 10 mm/year, we aim at sub-centimeter accuracies of the respective 3D point clouds. We show that hybrid georeferencing helps to increase the accuracy of the adjusted LiDAR point cloud by integrating results from photogrammetric block adjustment to improve the time-dependent trajectory corrections. As our main contribution, we demonstrate that joint orientation of laser scans and images in a hybrid adjustment framework significantly improves the relative and absolute height accuracies. By these means, accuracies corresponding to the GSD of the integrated imagery can be achieved. Image data can also help to enhance the LiDAR point clouds. As an example, integrating results from Multi-View Stereo potentially increases the point density from airborne LiDAR. Furthermore, image texture can support 3D point cloud classification. This semantic segmentation discussed in the final part of the paper is a prerequisite for further enhancement and analysis of the captured point cloud.

1. INTRODUCTION

The quality of area-covering 3D point clouds as captured by aerial and mobile mapping platforms still experiences a considerable boost due to the ongoing advancements in LiDAR technology and Multi-View-Stereo-Matching (MVS). One main advantage of MVS is that the resulting geometric accuracy directly corresponds to the Ground Sampling Distance (GSD) and thus the scale of the evaluated imagery. In contrast, the potential to measure multiple responses of the reflected signal using LiDAR sensors is advantageous, especially to collect data both on and below vegetation using airborne data acquisition. Airborne LiDAR and MVS were originally developed as competing approaches. Meanwhile considerable research efforts focus on systems for joint collection and evaluation of LiDAR and image data to further improve the accuracy, density and reliability of the generated point clouds. Furthermore, full use of the geometric information provided from these data sources additionally requires a semantic analysis of the respective point clouds. Also triggered by the astonishing improvements in the field of pattern recognition and machine learning, automatic interpretation of area-wide point clouds is moving to a mature state (Hackel et al., 2016). Resulting approaches can analyze airborne 3D point clouds for object detection and classification or can be used for further refinement of the point clouds for object-dependent filtering and smoothing (Bláha et al., 2017). Within our paper, we discuss and demonstrate the potential of combining LiDAR and image processing at different steps of the photogrammetric processing chain while aiming at the generation of ultra-high accurate 3D point clouds using a UAV platform. One main purpose of this research is the deformation monitoring of a ship lock and its surrounding area as depicted in Figure 1. The area of interest stretches over 570 m (east-west) × 780 m

(north-south) and includes the ship lock facilities, the river and the riparian area as well as vegetated areas, farmland, and residential areas of the surrounding village.

In that test area subsidence of about 1-10 mm/a relative to the stable surroundings have been observed over the past few years (Kauther & Schulze, 2015). For area-covering monitoring of such changes, 3D point clouds at mm-accuracy have to be provided twice a year. Up to now, such accuracy demands presume terrestrial data collection using geodetic instruments, such as level instruments, total stations and differential GNSS. However, these engineering geodesy techniques are usually limited to specific parts of built structures or natural objects due to economic reasons. In contrast, area-covering 3D measurement calls for the use of airborne platforms. Photogrammetric data collection at mm-scale requires image acquisition at a similar resolution, which typically presumes the use of UAVs. If (signalized) ground control points are available with sufficient accuracy and distribution, in principle integrated georeferencing and subsequent dense image matching can provide 3D point clouds in the accuracy of some millimeters. However, our project aims on the measurement of subsidence of terrain surfaces, which are covered to a considerable part by vegetation like trees, bushes and shrubs. This requires the use of LiDAR, due to its ability to penetrate such vegetation, especially if multiple returns are measured and analysed, e.g. using full waveform recording. While originally, UAVs were limited to camera-based systems, meanwhile the use of even high-end LiDAR became state-of-the-art (Cramer et al., 2018). Thus, subsidence measurement can be realized by collecting LiDAR as well as image data from UAV platforms at different epochs.

The original set-up of our test area and preliminary results from joint image and LiDAR flights in March 2018 were already introduced in Cramer et al. (2018). Section 2 of our paper now

* Corresponding author

describes data collection for the subsequent epochs in November 2018 and March 2019. Section 3 then discusses the required highly accurate georeferencing of the captured data, which is the most important prerequisite while aiming at the monitoring of subsidence based on these two epochs. In a nutshell, the flight trajectory of the LiDAR platform are required with sufficient accuracy to compute the respective 3D points from LiDAR range measurement. Therefore, IMU and GNSS measurements provide the required position and attitude of the platform, which allows 3D point accuracies of some centimeters, especially if a suitable calibration of the sensor system is guaranteed. However, since our application aims at sub-centimeter accuracies, further improvement is required. As discussed in Section 3, we apply a hybrid orientation of airborne LiDAR point clouds and aerial images as proposed by Gliira et al. (2019). This integration of aerial imagery not only increases the resulting accuracy of the LiDAR points during georeferencing, it also provides a precise co-registration of both data sources. As demonstrated in Section 4, this allows accurate alignment of LiDAR points to 3D points from dense image matching. By these means, further processing can benefit from the complimentary properties of LiDAR and MVS point clouds. In addition to a joint point cloud filtering, their semantic segmentation is indispensable for many applications. As discussed in Section 5, we use such a classification to eliminate points measured on objects like cars or vegetation. This finally leaves only points measured on plane, stable surfaces like building roofs or terrain, which are suitable for the aspired measurement of subsidence.

2. TEST SET-UP

Figure 1 depicts our test area in Hessigheim, Germany with a maximum site extension of 570 m (east-west) x 780 m (north-south).

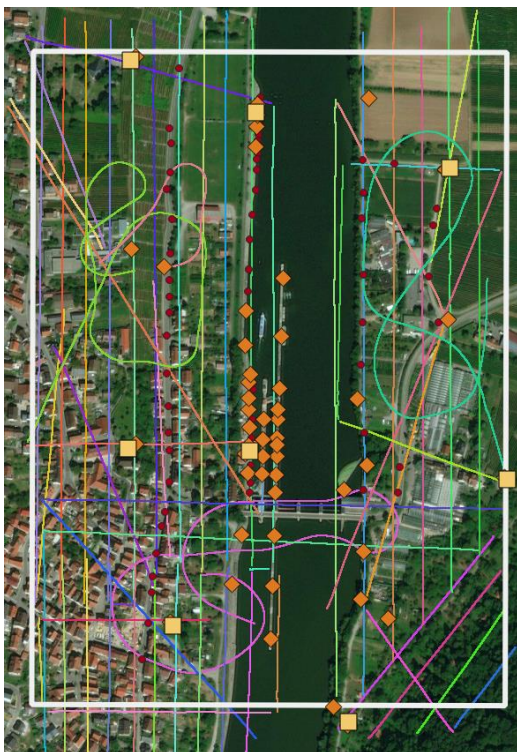


Figure 1. Area of investigation (white box) with captured LiDAR strips (coloured individually), LiDAR control planes (dark yellow squares), checkerboard targets (orange rhombs) and height checkpoints (red dots).

Data from a RIEGL VUX-1LR LiDAR sensor combined with two Sony Alpha 6000 oblique cameras were captured using the RIEGL RiCopter octocopter in November 2018 and March 2019. The LiDAR data acquisition comprises 17 longitudinal (i.e. north-south) strips, 4 diagonal strips to cover the steep wooded slope in the south-eastern corner of the investigation area, and some extra flight lines with diagonal and curved trajectories for further block stabilization (cf. Figure 1). With a flying speed of 8 m/s, a nominal flying altitude of 50 m above ground level, a strip distance of 35 m, a pulse repetition rate of 820 kHz, a scan line rate of 133 Hz, and a used scanner Field-of-View (FoV) of 70°, the resulting mean laser pulse density is 300-400 points/m² per strip and more than 800 points/m² for the entire flight block due to the nominal side overlap of 50%. These flight mission parameters guarantee a laser footprint diameter on the ground of less than 3 cm enabling a high planimetric resolution of 5 cm. The ranging accuracy, reported in the data sheet of the sensor is 10 mm (Riegl 2018). The two Sony Alpha 6000 oblique cameras mounted on the RiCopter platform covered a FoV of 160° at a GSD of 1.5-3 cm.



Figure 2. Ship lock and photogrammetric control point (PCP) on a pillar (left), LiDAR control plane (LCP) (right).

Georeferencing of the acquired UAV LiDAR data is directly accomplished since the trajectory of the platform is measured by the GNSS/IMU system. However, accuracy of this standard approach is limited to up to a few centimeters. Thus, as further discussed in Section 3, we additionally apply signalized ground control information and finally realize a hybrid orientation of LiDAR and image data. Figure 1 shows the distribution of these signals within our test area. Checkerboard signals scattered within the test area serve as photogrammetric control points (PCPs) during Automatic Aerial Triangulation (AAT) of the imagery. These checkerboards are partly mounted on pillars (see Figure 2 left), tripods, and on ground. They feature a diameter of 27 cm which allows automatic measurement of image points. For absolute georeferencing of the LiDAR data, distinct planar geometries with known position and orientation in space need to be provided as well. For our experiments we used LiDAR control planes (LCPs) constructed by two roof-like oriented planes as depicted in the right of Figure 2. Each roof plane features a size of 40 cm x 80 cm. Point coordinates for ground control were provided by geodetic survey, i.e. either GNSS or tachymetric measurement. The accuracy of these reference points is in the range of 1-3 mm. As demonstrated by Cramer et al. (2018), providing such quality for a considerable number of points scattered in a larger test area causes great effort.

3. JOINT GEOREFERENCING OF IMAGE AND LIDAR DATA

For direct georeferencing of the VUX1-LR scanner data, the RIEGL RiCopter platform integrates an APX-20 UAV. In Section 3.1. we show that the estimated trajectory is further improved by a suitable correction model during LiDAR strip

adjustment. Further improvement is possible by additional aerial image information, which in our experiments is provided from the Sony Alpha cameras also mounted on the RIEGL RiCopter. This so-called hybrid orientation is discussed in Section 3.2.

3.1 LiDAR Strip Adjustment

During acquisition of LiDAR data, the measurements of the inertial navigation system and the laser scanner are time-stamped. In post-processing, the IMU and GNSS measurements are combined in a Kalman filter to a trajectory providing the position and attitude of the platform over time. The polar elements, i.e. range and angles, are determined for each laser shot during data acquisition. For full-waveform scanners, the range of each laser echo is calculated either offline in post-processing (Mallet and Bretar 2009) or onboard during data acquisition (Pfennigbauer et al. 2014). Based on i) the trajectory, ii) the scanner's mounting calibration (i.e. position and orientation offset between the estimated trajectory and the scanner's own coordinate system), and iii) the polar measurements of the laser scanner, the 3D coordinates of each detected laser echo can be calculated by direct georeferencing. Any error in the estimated trajectory, the mounting calibration, or the scanner's measurements will cause an offset between point clouds of different flight strips in overlapping areas. Respective discrepancies are detected within standard quality control procedures (Ressl et al., 2008). If the deviations exceed acceptable limits, a typical LiDAR workflow also includes a strip adjustment to minimize the offsets between the strips (Shan & Toth 2018).

Since strip-to-strip differences without trajectory correction exceed the desired accuracy by far (see Figure 3a), a LiDAR strip adjustment was carried out applying the method presented in Glira et al. (2016). Within a sophisticated calibration procedure, the six parameters of the mounting calibration (lever arm and boresight misalignment), a global datum shift, as well as trajectory corrections were estimated to minimize the discrepancies (defined as point-to-plane distances within the overlap area of flight strip pairs). For absolute orientation, the LCPs are considered. Two different solutions of the LiDAR strip adjustment have been investigated. The bias correction model (see Figure 3b) considers a constant offset (Δx , Δy , Δz , $\Delta roll$, $\Delta pitch$, Δyaw) applied to the original trajectory solution of each individual strip. With the spline model (Figure 3c), time-dependent corrections are modelled for each of the six above mentioned parameters by cubic spline curves with equidistant time intervals of 8 s. This adds much more flexibility to further minimize the discrepancies between overlapping strips (Glira et al. 2016).

The relative strip height differences after applying both approaches are plotted in Figure 3b and c. Compared to the raw measurements, the bias model improves strip fitting precision considerably. The residual height error decreases to 1.1 cm compared to 5.8 cm measured as robust standard deviation of absolute strip height differences (σ_{MAD}) in smooth and open surface areas. Locally, systematic effects are still perceivable due to insufficient correction of trajectory errors. Applying the more flexible spline correction model reduces σ_{MAD} to 0.4 cm. The absolute orientation of the LiDAR block is improved in the strip adjustment using the LCPs, while the checkerboard targets allow for an independent evaluation. The residual errors at these points are defined as the point-to-plane distances to the best fitting plane estimated from the neighbouring points of the LiDAR strips. Featuring a standard deviation (std) of the residual deviations of 1.8 cm, the spline model performs worse compared to the bias model with a std of 1.0 cm. This clearly indicates that the spline model is to be preferred regarding relative adjustment, but can

only be applied when sufficient stabilizing reference points are available to avoid global block deformation. However, since measuring ground control data is costly and therefore not economic, Glira et al. (2019) developed the idea of hybrid adjustment, where the flight trajectory is further stabilized by the integration of LiDAR and image measurements.

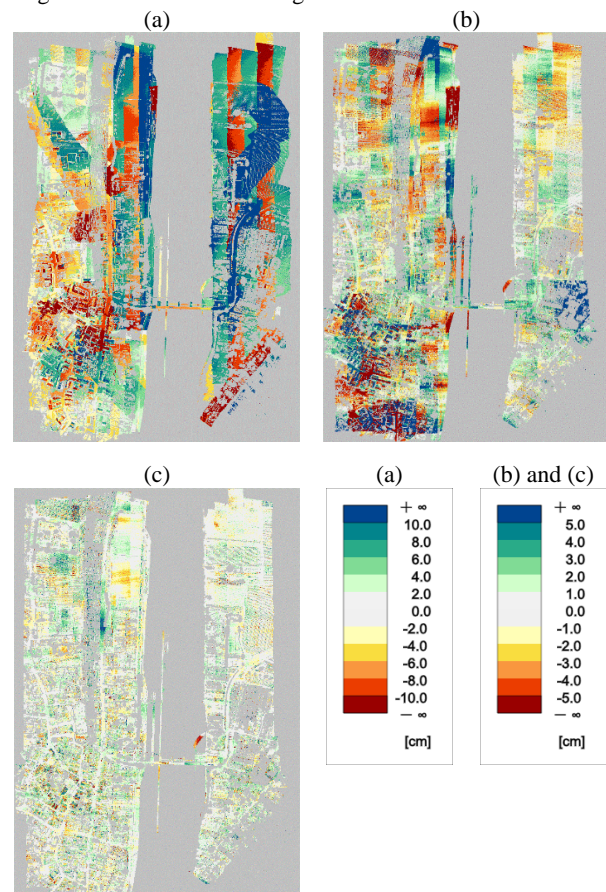


Figure 3. Relative strip differences of (a) raw data and after strip adjustment using trajectory corrections with (b) bias model and (c) spline model.

3.2 Hybrid Orientation of Airborne LiDAR and Aerial Images

The hybrid georeferencing of LiDAR and aerial images is an extension of the traditional LiDAR strip adjustment with additional observations from the bundle adjustment of image blocks (Glira et al. 2019). Usually, bundle block adjustment estimates the respective camera parameters from corresponding pixel coordinates of overlapping images, while the object coordinates of these tie points are a by-product. Within the hybrid orientation approach, tie points' object coordinates are used to establish correspondences between the LiDAR and the image block, which are minimized within a global adjustment procedure. In this respect, hybrid orientation is an extension of LiDAR strip adjustments. In addition to the respective calibration parameters, both approaches estimate parameters of a spline model for time-dependent corrections of the flight trajectory. However, while the flexibility of this model can result in systematic deformations during LiDAR strip adjustment, the integration of tie points generated from stable 2D image frames as oriented during bundle block adjustment reliably avoids such negative effects. This is especially helpful if both sensors are flown on the same platform and thus share the same trajectory. Figure 4 shows that the idea of tightly coupled hybrid adjustment of concurrently captured LiDAR and image data significantly

stabilizes the measured trajectory, exemplary visualized for one longitudinal strip. This strip features a time range of about 90 s corresponding to 720 m at the speed of 8 m/s. It can clearly be observed that low frequency fluctuations can be reduced by introducing the tightly coupled images in the hybrid adjustment. While the estimated spline for the LiDAR-only adjustment shows significant low-fluctuations, in case of hybrid adjustment this effect is considerably reduced by tight coupling to imagery.

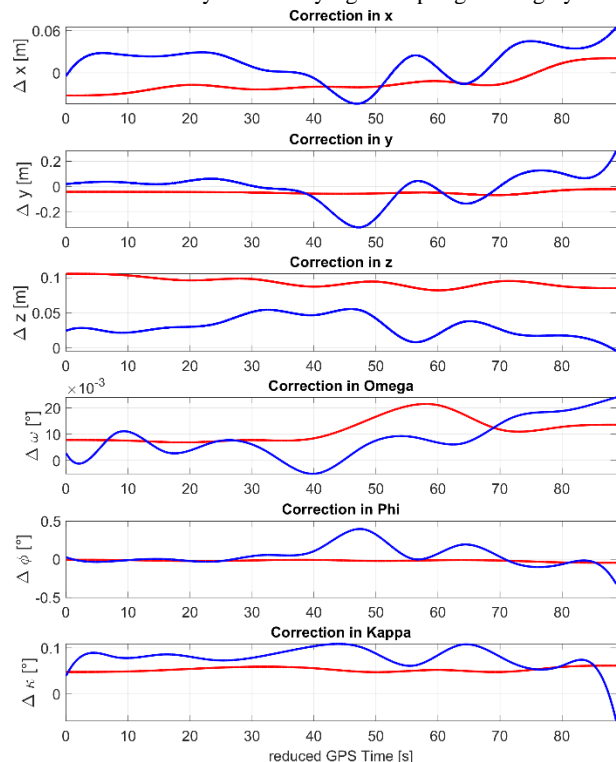


Figure 4. Comparison of adjusted trajectory for one exemplary longitudinal strip when using spline trajectory correction by LiDAR strip adjustment (blue) and hybrid orientation (red). (GPS time reduced to first measurement).

The accuracy of tiepoints, used as observations during hybrid orientation depends on the quality of the respective AAT, which is influenced by the geometry of the image block and the used

camera. As discussed by Glira et al. (2019) and Mandlburger et al. (2017), problems in using correspondences between points from image matching and LiDAR can occur from their different properties, which are further discussed in Section 4. To avoid mismatches we used tie points observed in at least three images and rejected correspondences with a distance larger than $\pm 3\sigma$. In our current experiments, we apply two oblique looking Sony Alpha 6000 cameras mounted at the RIEGL RiCopter platform. In the original concept of the platform, these cameras were intended to just provide RGB color values for the respective LiDAR points. Bundle block adjustment results in differences at independent check points between 5.2 cm (max.) and 1.2 cm (min.). The mean RMS is 2.5 cm which is within the range of the GSD of 1.5-3 cm. While this 3D object point quality does not meet the project requirements of millimeter accuracy, the cameras are tightly coupled to the laser scanner. Both sensors share a common trajectory and, in this case, the images directly support the estimation of the trajectory correction.

We demonstrate the benefits of hybrid adjustment compared to LiDAR strip adjustment for the subset depicted in Figure 5. In both cases, the relative height differences of the respective LiDAR strips reaches an accuracy σ_{MAD} of 4 mm. However, clear distinctions between both approaches become visible from the color-coded elevation differences of LiDAR point cloud to signalized targets. The red circles indicate control points, whereas the remaining signals were used as check points.

Method	Ground Control	Ground Check	min	max	mean	std
LiDAR-only	3 LCPs	42 PCPs	-1.3	2.7	0.9	0.9
		2 LCPs	-2.8	4.7	1.0	3.4
Hybrid	9 PCPs	33 PCPs	-0.8	0.6	-0.2	0.3
		5 LCPs	-1.0	1.5	0.1	0.8

Table 1: Elevation differences of LiDAR point cloud to signalized targets. All measures are given in cm.

The top row of Table 1 as well as Figure 5(a) give the results for the LiDAR strip adjustment. Overall, 5 LCPs were available, 3 LCPs were used as ground control (red circles), the remaining 2 LCPs (blue circles) served as check points similar to the 42 PCPs. To compute the standard deviation of 0.9 cm, Point-to-plane distances between the LiDAR point cloud and the horizontal photogrammetric targets were used. Thus, horizontal errors in the

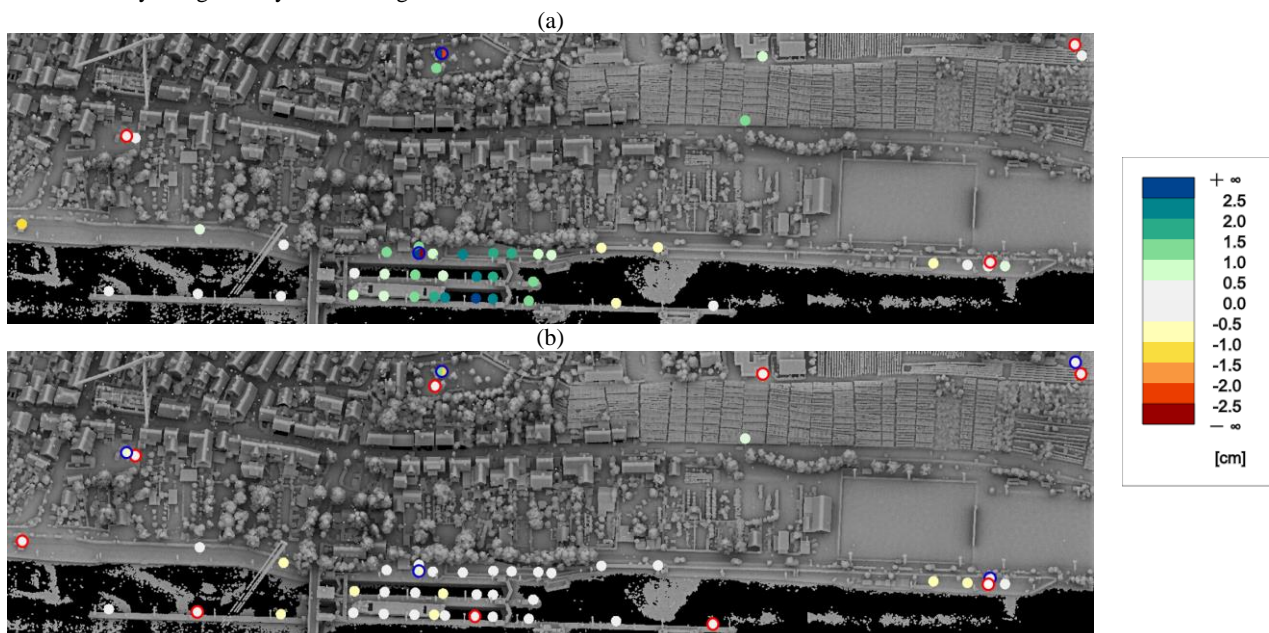


Figure 5. Point-to-plane distances of LiDAR point cloud to signalized targets. *Top*: Results from LiDAR strip adjustment, *Bottom*: Results from hybrid orientation. Control information in terms of LCPs (*top*) and PCP (*bottom*) is highlighted by red circles. LCPs used as check points are indicated by blue circles and denote differences to both roof planes.

adjusted LiDAR point cloud are not taken into account. However, a potential horizontal error influences the elevation difference of the inclined planes at LCPs. This is the reason for the larger standard deviation of 3.4 cm.

The bottom row of Table 1 as well as Figure 5(b) give the results for the hybrid orientation. Again, the red circles indicate points used as ground control. In contrast to LiDAR strip adjustment, which requires LCP as ground control, georeferencing by hybrid adjustment no longer needs dedicated LiDAR control planes. Instead, all control point information is integrated from the standard photogrammetric targets applied during AAT. This dispensation with LCPs considerably reduces the effort for maintaining the test area and thus is of high practical relevance. As indicated, 9 PCPs (red circles) were used as control points. Elevation differences at the remaining 33 PCPs result in a standard deviation of 0.3 cm. Differences to the LiDAR strip adjustment are even more obvious for the 5 LCPs (blue circles) with a standard deviation of 0.8 cm.

3.3 Comparison of Elevation Models from Different Epochs

As discussed by Kauther & Schulze (2015) our test area is subject to potential subsidence in the order of 10 mm/year. Thus, for the time frame between our measurement epochs in November 2018 and March 2019, changes in height in the order of 3 mm are expected, which are still beyond detectability while using our current sensor setup.

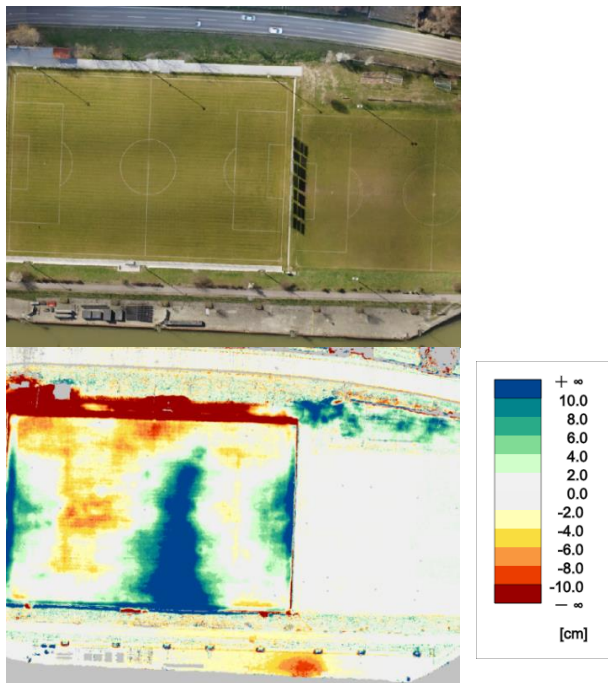


Figure 6. Exemplary elevation differences of LiDAR campaigns 2016 and 2019 (*bottom*) compared to corresponding orthophoto (*top*).

Thus, Figure 6 gives elevation differences between our measure in 2019 and data from an airborne LiDAR campaign from the year 2016. That data set, provided from the State Office for Spatial Information and Land Development Baden-Wuerttemberg (Landesamt für Geoinformation und Landentwicklung Baden-Württemberg, LGL), does not meet our requirements on point density and accuracy. Still, if subtracted from our UAV measure for demonstration purposes, we can see a red spot close to the river, which indicates a subsidence already reported by Krauthar & Schulze (2015). As it is also visible in

the corresponding orthophoto on top of Figure 6, differences also occur at and in the vicinity of the sports field. In that area, maintenance took place between both measures. Precisely, sinks were refilled causing terrain rise, depicted in Figure 6 in dark blue. In between the sports field and the nearby street a new paved path was created, which required ablation of the slope. This deformation is highlighted in dark red. Further differences also occur at vegetated areas. This requires further semantic interpretation of the data as a prerequisite to the analysis of occurring differences. For this purpose, Section 4 first discusses the integration of point clouds from LiDAR with results from Multi-View-Stereo image matching, while Section 5 aims on the semantic segmentation of the collected point clouds.

4. POINT CLOUDS FROM LIDAR AND MULTI-VIEW-STEREO

The integrated orientation of laser scans and images as described in Section 3 ensures optimal alignment of both data sources. This also holds for the 3D point clouds derived from LiDAR or via Multi-View-Stereo Image Matching. For this purpose, commercial software tools e.g. described by Rothermel et al. (2012) provide 3D information basically for each image pixel at considerable quality if sufficient image overlap is available.



Figure 7. Comparison of 3D points from MVS (RGB) and LiDAR measurement (height coded). Yellow line defines profile depicted in Figure 8.

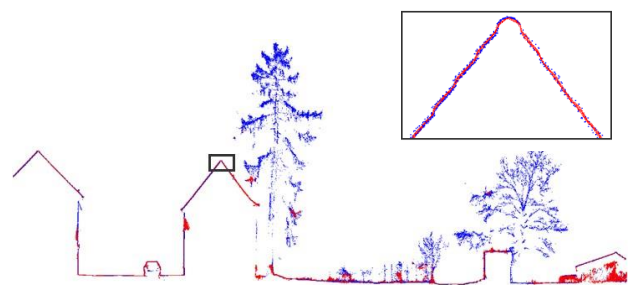


Figure 8. Extracted profile with 3D points from LiDAR (blue) and MVS (red).

Figure 7 and Figure 8 demonstrate the complementary characteristics of LiDAR and MVS for a sample from our test area. Figure 7 depicts the RGB coloured points generated by MVS. For this purpose, the nadir images from the PhaseOne camera captured at a GSD of 3.7 mm and an 80/60 overlap were used. While MVS directly provides the respective RGB colour values for the visualization of these points, the overlaid LiDAR data is colour coded according to the respective elevation. Finally, the yellow line represents the profile used to extract the

points depicted in Figure 8. The discrepancies between the point clouds from MVS (red) and LiDAR (blue) are especially evident at trees. In general, the polar measurement principle of LiDAR allows the detection of one or multiple returns along a single laser ray path. At twigs and branches, but also at power lines, multiple returns are captured for a single laser pulse since the laser beam cone hits targets smaller than the laser footprint in different distances along the ray path. MVS point cloud generation presumes that the same surface patch is seen from at least two camera positions. This can result into problems especially if the object appearance changes rapidly when seen from different positions. This holds true for semi-transparent objects like trees and bushes. Problems can also occur for objects in motion like vehicles, pedestrians, etc., or in very narrow urban canyons due to occlusions. Differences between LiDAR and MVS also occur at grass, which is penetrated by the laser signal to a certain extent. Thus, these heights are measured somewhere between the top surface and the ground depending on the vegetation density, while multi-view stereo matching, in turn, always returns the topmost surface and does not penetrate the vegetation layer (Mandlbürger et al. 2017).

While in the past, LiDAR and image matching were considered as competing techniques, the closer integration of both techniques is the logical next step. As a simple example, colour information can be mapped to the LiDAR point cloud. The top image of Figure 9 shows a section of the LiDAR point cloud coloured by the reflectance of the measured objects. In the bottom image, texture provided from the integrated cameras was added to the meshed point cloud.



Figure 9. LiDAR point cloud with corresponding colour coded reflectance (top) and 3D mesh with corresponding image texture (bottom).

The integrated capture and evaluation of images and LiDAR from a UAV platform can generate 3D point clouds at a very high quality. In principle, a suitable combination of both LiDAR and MVS can further increase the robustness, accuracy and reliability of the resulting 3D point clouds. Since this is beyond scope of

our paper, processing is limited to the LiDAR point cloud for the time being. Still, a number of follow-up applications require further enhancement by adding semantic information. As an example, the aspired monitoring of subsidence requires the delineation of adequate surfaces like bare soil, paved roads, and roof surfaces. In contrast, points at vegetation like trees, shrub or grass as well as points on moving objects like cars have to be discarded. To filter out such points, we aim at a semantic segmentation discussed in the following section.

5. SEMANTIC SEGMENTATION OF POINT CLOUDS

In principle, we could use a binary classification to eliminate 3D points measured at objects not suitable for subsidence analysis. However, as a matter of generality, we aim on a more demanding classification task distinguishing between 11 different object classes, which are inspired by the ISPRS 3D semantic labeling contest (Niemeyer et al., 2014): *Powerline, Low Vegetation, Impervious Surface, Car, Urban Furniture, Roof, Façade, Shrub/Hedge, Tree, Bare Soil* and *Vertical Surface*. In the first step, we compute a feature vector for each point, which describes the characteristic point distribution in the vicinity at different levels of abstraction. Precisely, features are computed within different radii of $r = 1, 2, 3$ and 5 m both for a spherical and unbound cylindrical neighbourhood where r defines the radius of the sphere and the cylinder respectively. To reduce the computational effort for querying neighbours in our high-resolution point clouds, with increasing search radius these points are selected from a gradually more subsampled point cloud as proposed by Hackel et al. (2016).

The most common features for classification of 3D point clouds are based on the structural tensor (Becker et al., 2018). If Eigenvalues and -vectors are extracted from this covariance matrix, characteristics of the point distribution within a local neighborhood can be obtained. Relevant features are then derived by utilizing the ratio of Eigenvalues (Weinmann et al., 2018). Furthermore, estimating Eigenvalues and -vectors of the structural tensor equals fitting a plane in the vicinity of each point. Therefore, features describing the local orientation of this plane and the distance of individual points to this plane can be computed. As a significant feature for delimiting vegetation we observed the robustly determined standard deviation σ_{MAD} of all derived distances to this plane within the respective neighbourhood. Another group of relevant features consider height information. Thus, maximum height difference in a neighbourhood, variance of height differences and height above ground (Chehata et al., 2009) are extracted. In order to add radiometric properties, colour information was mapped from the synchronously acquired oblique images to the respective LiDAR point cloud. Visualizations as e.g. depicted at the bottom of Figure 9 typically apply RGB values. However, these RGB values were transformed to the HSV colour space for classification. To enable a better generalization, these values are additionally averaged within each neighbourhood (Becker et al., 2018). To further benefit from the high resolution imagery, we generate an orthophoto and perform a semantic segmentation using SegNet (Badrinarayanan et al. 2017). The resulting labels of the orthophoto are then mapped to the 3D points to provide a priori labels. Further features are also provided from LiDAR-inherent measures like reflectance and echo ratio. Finally our feature vector consists of 167 elements. We then use these features to train a Random Forest (RF) classifier (Breiman, 2001). Figure 11 exemplarily depicts a comparison of ground truth data (top) and predicted classes (bottom). As it is visible from the confusion matrix depicted in Figure 10. By this, we achieve an Overall Accuracy (OA) of 86.8%. Hyperparameters

were tuned based on a validation dataset, resulting in an ensemble of 300 binary decision trees having a maximum depth of 18. As mentioned before, this classification is motivated by extracting surfaces suitable for deformation monitoring. Thus, confusion between *Impervious Surface*, *Roof* and *Bare Soil* and on the other hand classes inappropriate for monitoring especially having inherent dynamic properties needs to be minimized. In this context, the highest confusion (40%, see Figure 10) of the predicted result can be observed between *Bare Soil* and *Low Vegetation*, which is due to similar height above ground and similar geometrical features.

Predicted Label	Powerline	Low Veg.	Imp. Surf.	Car	Urb. Furnit.	Roof	Façade	Shrub	Tree	Bare Soil	Vert. Surf.	PA
Powerline	1	0	0	0	0	0	0	0	0	0	0	1
Low Veg.	0	.96	.08	.03	.07	0	.02	.07	0	.40	.02	.85
Imp. Surf.	0	.01	.86	.04	.01	.03	.01	0	0	.15	.05	.90
Car	0	0	0	.51	.01	0	0	0	0	0	0	.88
Urb. Furnit.	0	0	0	.26	.35	0	.03	.01	0	0	.02	.68
Roof	0	0	.02	.02	.08	.92	.05	0	0	0	.01	.87
Façade	0	0	.01	.02	.14	.02	.84	.01	0	0	.16	.82
Shrub	0	.01	0	.07	.17	0	.01	.70	0	0	0	.70
Tree	0	.01	0	.04	.11	.03	.03	.21	.99	0	0	.93
Bare Soil	0	.01	.02	0	0	0	0	0	0	.44	.01	.80
Vert. Surf.	0	0	0	.01	.04	0	.02	0	0	0	.74	.86
PA	1	.96	.86	.51	.35	.92	.84	.70	.99	.44	.74	.87
True Label	Powerline	Low Vegetation	Impervious Surface	Car	Urban Furniture	Roof	Façade	Shrub/Hedge	Tree	Bare Soil	Vertical Surface	UA

Figure 10. Normalized confusion matrix including Producer's Accuracy (PA), User's Accuracy (UA) and Overall Accuracy (bottom right) for our RF prediction.

Integration of colour information did not improve the separation of both classes. The flights had to be carried out in the leave-off season, in our experiments in November and March, respectively. In that periods, vegetation areas are not characterized by their typical green colour, rather both areas of open soil and thin grass are of brownish colour. Furthermore, 15% of the points which actually represent *Bare Soil* are misinterpreted as *Impervious Surface*. This is mainly caused by very similar geometric features, since in our test region, areas of bare soil are primarily represented by very small domestic agricultural fields featuring very smooth surfaces, especially in our flight period in winter. However, both classes are to be used for deformation measurements in the context of binary separation. Therefore this confusion will not impact our monitoring result. As our ultimate goal is to measure vertical subsidence, we further need to exclude façades. Only a very small extent of façades are mistaken as roofs, often caused by roof dormers, which are difficult to delimit.

The visual comparison between ground truth and predicted classes (see Figure 11) verifies these findings and outlines broad agreement. Deviations are particularly noticeable in the upper left area at the construction site, where the classification result differs partly from the correct class *Bare Soil*. Further confusion due to similar geometries and colouring occurs, as expected, for *Façade* vs. *Vertical Surface*, *Shrub/Hedge* vs. *Tree* and *Urban Furniture* vs. *Car*. These misclassifications are not further discussed here,

since all these classes can be merged and excluded for our monitoring purposes.

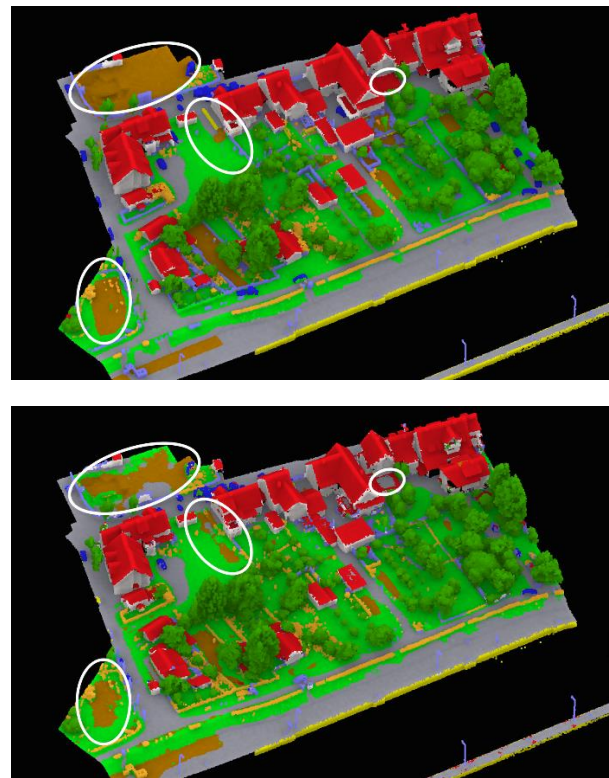


Figure 11. Comparison of ground truth data (top) and predicted classes by RF classifier (bottom). Differences especially occur within white ellipses.

6. CONCLUSION AND FURTHER WORK

This paper presents a workflow for hybrid georeferencing, enhancement and classification of ultra-high resolution UAV LiDAR and image point clouds for monitoring applications. We clearly demonstrated the feasibility of the hybrid orientation of airborne LiDAR and aerial images. By these means the elevation accuracy of UAV-based LiDAR point clouds could be improved to a standard deviation of 0.8 cm. To the best of our knowledge, such accuracies were not feasible so far. We expect a further improvement by replacing the Sony Alpha cameras currently mounted on the RIEGL RiCopter by high quality cameras. A previous flight campaign already demonstrated promising AAT results for high quality nadir images captured by a PhaseOne iXM 100-RS camera with 35mm lens mounted on a CopterSystems CS-SQ8 copter (Cramer et al. 2018). This will improve the GSD of the imagery integrated during hybrid georeferencing from 2cm to approximately 4 mm.

Similar to the georeferencing of LiDAR and image data, point cloud generation from Multi-View-Stereo image matching and LiDAR were considered as competing techniques with research efforts focussing on the individual improvement of sensors and algorithms. In our future work, we will also aim on the suitable combination of both data sources to further increase robustness, accuracy and reliability of 3D point clouds while aiming at ultra-high accuracy applications from UAV-based data capture. The same holds true for semantic information extraction. Combining very high resolution texture mapped, meshed 3D point clouds from image matching with information from multiple laser returns occurring for example at twigs and branches opens new opportunities for the following point cloud analysis. In addition to the further improvement of point cloud geometry, integrated processing will facilitate the subsequent semantic analysis during

point cloud classification or object detection. Similarly, adaptive filtering and smoothing of meshed 3D points will benefit from integrating knowledge on different semantic classes. In this context, a priori assumptions on the shape of the captured surface patches are frequently applied. Integration of active laser and passive image sensors is thus beneficial for traditional airborne scenarios like topographic data capture.

ACKNOWLEDGMENTS

Considerable parts of the research presented in this paper were funded within a project granted by the German Federal Institute of Hydrology (BfG) in Koblenz.

BIBLIOGRAPHY

Applanix, 2018. APX UAV product specifications. URL: <https://www.applanix.com/products/dg-uavs.htm>.

Badrinarayanan, V., Kendall, A. & Cipolla, R., 2017. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 39(12), pp. 2481-2495.

Becker, C., Rosinskaya, E., Häni, N., d'Angelo, E. & Strecha, C., 2018. Classification of Aerial Photogrammetric 3D Point Clouds. *Photogrammetric Engineering and Remote Sensing* 84, pp. 287-295. 10.14358/PERS.84.5.287.

Blaha, M., Rothermel, M., Oswald, M., Sattler, T., Richard, A., Wegner, J., Pollefeys, M., Schindler, K., 2017. Semantically Informed Multiview Surface Refinement. *2017 IEEE International Conference on Computer Vision (ICCV)*, Venice, 2017, pp. 3839-3847. doi: 10.1109/ICCV.2017.412.

Breiman, L., 2001. Random Forests. *Journal of Machine Learning* 45(1), pp. 5-32.

Chehata, N., Guo, L. & Mallet, C., 2009. Airborne LiDAR feature selection for urban classification using random forests. *Int. Arch. Photogr. Rem. Sens. and Spat. Inf. Sci. XXXVIII-3/W8*, pp. 207-212.

Cramer, M., Haala, N., Laupheimer, D., Mandlbürger, G. & Havel, P., 2018. Ultra-high precision UAV-based LiDAR and Dense Image Matching. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci. XLII-1*, pp. 115-120. DOI: 10.5194/isprs-archives-XLII-1-115-2018.

Glira, P., Pfeifer, N. & Mandlbürger, G., 2016. Rigorous strip adjustment of UAV-based laserscanning data including time dependent correction of trajectory errors. *Photogrammetric Engineering & Remote Sensing* 82(12), pp. 945-954.

Glira, P., Pfeifer, N., Mandlbürger, G., 2019. Hybrid orientation of airborne LiDAR point clouds and aerial images. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci. Volume IV-2/W5*, URL: <https://doi.org/10.5194/isprs-annals-IV-2-W5-567-2019>.

Hackel, T. & Wegner, J. & Schindler, K., 2016. Fast semantic segmentation of 3d point clouds with strongly varying density. *ISPRS Ann. Photogr. Rem. Sens. Spat. Inf. Sci. III-3*, pp. 177-184.

Kauther, R., & Schulze, R., 2015. Detection of subsidence affecting civil engineering structures by using satellite InSAR. *Proceedings of the Ninth Symposium on Field Measurements in Geomechanics*, Australian Centre for Geomechanics, pp. 207-218.

Niemeyer, J., Rottensteiner, F., & Sörgel, U., 2014. Contextual classification of LiDAR data and building object detection in urban areas. *ISPRS Journal of Photogrammetry and Remote Sensing* 87, pp. 152-165.

Mandlbürger, G., Wenzel, K., Spitzer, A., Haala, N., Glira, P., & Pfeifer, N., 2017. Improved topographic models via concurrent airborne LiDAR and dense image matching, *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci. IV-2/W4*, pp. 259-266, URL: <https://doi.org/10.5194/isprs-annals-IV-2-W4-259-2017>.

Riegl, 2019. RIEGL RiCOPTER product specifications. URL: http://www.riegl.com/uploads/tx_pxpriegldownloads/RiCOPTER_sensor-platform_Infosheet_2019-09-02.pdf.

Ressl, C., Kager, H. & Mandlbürger, G., 2008. Quality checking of ALS projects using statistics of strip differences. *International Archives of Photogrammetry and Remote Sensing* 37.B3b, pp. 253-260.

Rothermel, M., Wenzel, K., Fritsch, D. & Haala, N., 2012. *SURE: Photogrammetric Surface Reconstruction From Imagery*. LC3D Workshop, Berlin.

Mallet C. & Bretar, F., 2009. Full-waveform topographic lidar: State-of-the-art. *ISPRS Journal of Photogrammetry and Remote Sensing*, 64(1), pp. 1-16.

Pfennigbauer, M., Wolf, C., Weinkopf, J. & Ullrich, A., 2014. Online waveform processing for demanding target situations. In *Proc. SPIE*, 90800J.

Shan, J. & Toth, C. (eds), 2018. *Topographic Laser Ranging and Scanning: Principles and Processing*. Second Edition, CRC press. URL: <https://doi.org/10.1201/9781315154381>.

Weinmann, M., Blomley, R., Weinmann, M. & Jutzi, B., 2018. Investigations on the Potential of Binary and Multi-class Classification for Object Extraction from Airborne Laser Scanning Point Clouds. *Tagungsband der 38. Wissensch.-Tech. Jahrestagung der DGPF* 27, pp. 408-421.