IMPROVED TOPOGRAPHIC MODELS VIA CONCURRENT AIRBORNE LIDAR AND DENSE IMAGE MATCHING

G. Mandlburger [a, b*], K. Wenzel [c], A. Spitzer [d], N. Haala [b], P. Glira [a, e], N. Pfeifer [a]

[a] TU Vienna, Department of Geodesy and Geoinformation, Vienna, Austria -
(gottfried.mandlburger, philipp.glira, norbert.pfeifer)@geo.tuwien.ac.at
[b] University of Stuttgart, Institute for Photogrammetry, Stuttgart, Germany -
(gottfried.mandlburger, norbert.haala)@ifp.uni-stuttgart.de
[c] nFrames GmbH, Stuttgart, Germany - konrad.wenzel@nframes.com
[d] *RIEGL* Laser Measurement Systems, Horn, Austria - aspitzer@riegl.com
[e] Siemens AG, Corporate Technology (CT), Vienna, Austria, philipp.glira@siemens.com

Commission II, WG II/1

ABSTRACT:

Modern airborne sensors integrate laser scanners and digital cameras for capturing topographic data at high spatial resolution. The capability of penetrating vegetation through small openings in the foliage and the high ranging precision in the cm range have made airborne LiDAR the prime terrain acquisition technique. In the recent years dense image matching evolved rapidly and outperforms laser scanning meanwhile in terms of the achievable spatial resolution of the derived surface models. In our contribution we analyze the inherent properties and review the typical processing chains of both acquisition techniques. In addition, we present potential synergies of jointly processing image and laser data with emphasis on sensor orientation and point cloud fusion for digital surface model derivation. Test data were concurrently acquired with the *RIEGL* LMS-Q1560 sensor over the city of Melk, Austria, in January 2016 and served as basis for testing innovative processing strategies. We demonstrate that (i) systematic effects in the resulting scanned and matched 3D point clouds can be minimized based on a hybrid orientation procedure, (ii) systematic differences of the individual point clouds are observable at penetrable, vegetated surfaces due to the different measurement principles, and (iii) improved digital surface models can be derived combining the higher density of the matching point cloud and the higher reliability of LiDAR point clouds, especially in the narrow alleys and courtyards of the study site, a medieval city.

1. INTRODUCTION

Airborne LiDAR (Light Detection And Ranging), also referred to as Airborne Laser Scanning (ALS), has become the state-of-the-art data acquisition method for topographic mapping due its capability to provide dense and reliable height information and to penetrate vegetation through small gaps in the foliage. The acceptance of the technique is underlined by a number of facts: (i) national mapping agencies all over the world have started, have already completed, or are updating their country-wide data acquisition, (ii) approximately 20k articles have been published since the turn of the century according to ScienceDirect when searched for the keywords "lidar" and "topography", and (iii) several text books on the topic have been published (Shan and Toth, 2008; Vosselman and Maas, 2010).

Recent advances in digital photogrammetry and computer vision have brought passive imagery back into the focus for area-wide terrain elevation mapping. Especially the introduction of sophisticated Dense Image Matching (DIM) techniques (Remondino et al., 2014; Haala and Rothermel, 2012; Hirschmuller, 2008) providing elevations for each image pixel has increased the achievable point density and led to surface models with a point spacing equal to the ground sampling distance (GSD) of the images (typically 5-20 cm).

Therefore, several advanced 3D mapping sensors feature laser

scanners and (RGB) camera systems on the same airborne platform. One obvious advantage of such a hybrid sensor system is that the (monochromatic) laser echoes can be enriched by color information from synchronously acquired images. Beyond that, the combination of active laser and passive image sensors may open new opportunities for improving topographic data w.r.t. completeness, density, geo-referencing quality, precision, reliability, (point cloud) classification, object detection, surface reconstruction, and processing speed. For exploiting the full potential of the respective data sources it is necessary to fully understand their properties.

Whereas general comparisons of airborne LiDAR and DIM for terrain modeling and forestry applications were already published (Ressl et al., 2016; Vastaranta et al., 2013), this contribution specifically focuses on the potentials and challenges of simultaneous data acquisition from the same platform. This rules out that differences in the products are due to changes of the scene (construction activities, phenological state of vegetation, etc.) or changes in environmental factors (atmosphere, sun position, etc.).

Concurrent data acquisition with laser and imaging sensors asks for integrated data processing which will only then result in optimal final products (point clouds, surface and terrain models, meshes, orthophotos) if the specific characteristics are considered in the processing pipelines. Thorough orientation of overlapping LiDAR strips and images, respectively, is a precondition for high-quality surface reconstruction. Sophisticated flight block orientation strategies and software implementations exist for both

---

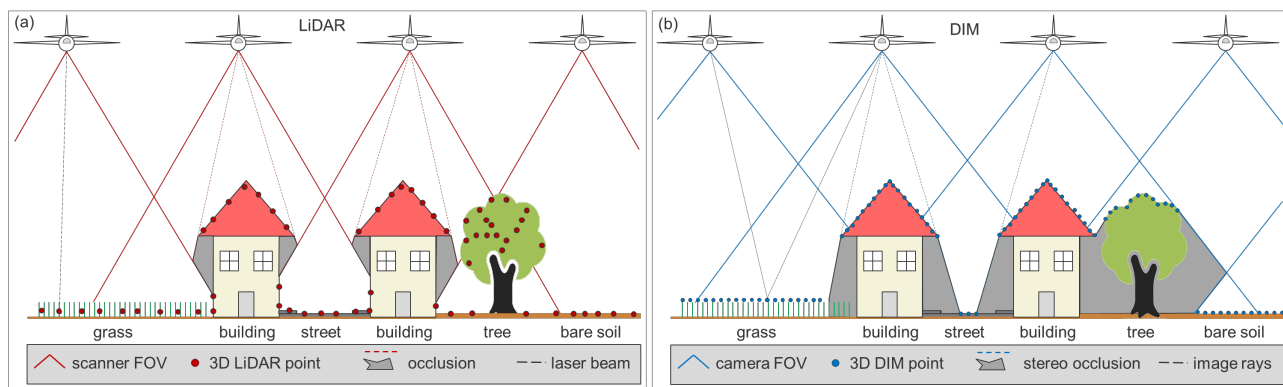*Corresponding author: gottfried.mandlburger@ifp.uni-stuttgart.de

Figure 1. Schematic drawing of data acquisition based on airborne LiDAR (a) and DIM (b)

laser scans (Skaloud and Lichti, 2006; Glira et al., 2015a) and images (Kraus, 2007; Förstner and Wrobel, 2016). However, the simultaneous use of data from LiDAR and image matching implies a second level of relative orientation between scans and images which is important to consider when jointly using the point clouds of both acquisition techniques for deriving products (surface and terrain models, orthophoto mosaics, etc.). Hence, one of the main goals of this paper is to examine the different properties of the point clouds obtained from LiDAR and image matching as only those areas must be used for mutual orientation of scans and images where no sensor related differences exist. In low vegetation, for instance, inherent height differences between the two techniques have been reported by Ressl et al. (2016). Building on that, new approaches for jointly processing scans and images are developed.

This article, therefore, briefly reviews the techniques and their properties in Section 2 and presents the typical individual data processing chains as well as proposals for integrated workflows in Section 4. The setup of an experimental data acquisition with a hybrid LiDAR/imaging sensor is described in Section 3 and the respective results are presented and discussed in Section 5. The major findings and recommendations are finally summarized in Section 6.

2. TECHNIQUES AND PROPERTIES

In this section, the principles of mapping topography with LiDAR and DIM are shortly summarized and the properties of the individual measurement techniques are discussed. This will result in requirements for joint data processing.

2.1 Airborne LiDAR

Airborne LiDAR is an active polar measurement system which uses short laser pulses for time-of-flight range detection. The linear LIDAR technology utilized in *RIEGL* LiDAR systems makes use of echo digitization and either full waveform analysis or online waveform processing (Ullrich and Pfennigbauer, 2016). Systems using this technology provide accurate 3D point clouds as well as additional attributes per point including calibrated amplitude and reflectance readings, but also attributes derived from the shape of the echo waveforms itself. One of the specific strengths of LIDAR technology is its multi-target capability enabling the penetration of vegetation to reveal objects below the canopy and to provide data from the ground for deriving a high resolution digital terrain model (cf. Figure1a).

2.2 Dense Image Matching

Driven by advances in digital camera system technology and algorithms, limits of automatic image based 3D surface reconstruction were pushed to a very high level regarding precision, robustness, and processing speed (Remondino et al., 2014; Haala and Rothermel, 2012). Starting from a set of images and their corresponding camera parameters (i.e. interior and exterior orientation), multi-view stereo (MVS) algorithms first compute socalled disparity maps applying suitable stereo matching algorithms (Hirschmuller, 2008; Yan et al., 2016). As a result, stereo correspondences between pixels across image pairs are established.

A set of images rectified to epipolar geometry, also referred to as stereo vision geometry, constitutes the typical input for dense stereo matching. In this case potential correspondences are located on identical rows across an image pair. Matching each image against multiple proximate images then generates redundant depth observations, which leads to so-called depth maps. Thus, each single pixel of an image block provides one or more corresponding 3D coordinates. The resulting dense point cloud can subsequently be filtered during DSM generation. If suitable redundancy is available from multiple views, state-of-the-art algorithms can reconstruct surface geometry at a resolution and accuracy which corresponds to the resolution of the available imagery.

2.3 Point cloud properties

The main characteristics of LiDAR and DIM are summarized in Table 1 and exemplary consequences for the acquired 3D point clouds are visualized in Figure 2 and Figure 3. Although the data samples are taken from the specific flight block detailed in Section 3, the aim is to highlight general properties of LiDAR and image matching point clouds.

Figure 2 visualizes the different penetration behavior. The depicted sports facility features a soccer field (grass) surrounded by a running track (hard surface) and a beach volleyball court in the northwestern part of the area (sand). In Figure 2b the height differences "LiDAR minus DIM" are plotted. The volleyball court and the running track are impervious for both acquisition techniques and, thus, the height differences are close to zero indicated by the whitish color tones. In contrast to that, the laser signal penetrates the grass layer to a certain extent and the heights are measured somewhere between the top surface and the ground depending on the vegetation density. Multi-view stereo matching, in turn, always returns the topmost surface and does not penetrate the vegetation layer. This can clearly be seen from the red

	Airborne LiDAR	Dense Image Matching
data acquisition	multi-sensor system (GNSS+IMU+scanner+camera)	
light source	active (laser)	passive (solar radiation)
measurement principle	time-of-flight	image ray intersection
measurement rays per point	1 (polar system)	>2 (multi-view stereo)
target detection	multiple targets per pulse	topmost surface
radiometry	mono-spectral (@laser wavelength)	multi-spectral (IR-R-G-B)
typical point spacing	20-50 cm	5-20 cm
preconditions	(diffuse) object reflectance	texture (image contrast)
typical height precision	2-3 cm	0.5-2 x GSD

Table 1. LiDAR/DIM characteristics and properties

color tones in Figure 2b and from the vertical section in Figure 2c and has a concrete implication for mutual orientation of scans and images. Although grass appears smooth in LiDAR and DIM datasets, the sensed surfaces exhibit a systematic bias and, thus, correspondences between images and scans must be avoided for these low vegetation areas.

Figure 3a and 3b show that the LiDAR dataset captures more points of the leafless tree and that the construction crane visible in the LiDAR point cloud is missing in the DIM dataset (Figure 3b). The polar measurement principle of LiDAR allows the detection of one or multiple returns along a single laser ray path. Multiple laser returns occur if the laser beam cone hits targets smaller than the laser footprint in different distances along the ray path (e.g. power lines, twigs, branches). As opposed to this, surface reconstruction based on (multi-view) stereo matching requires that the same object is seen from at least two camera po-
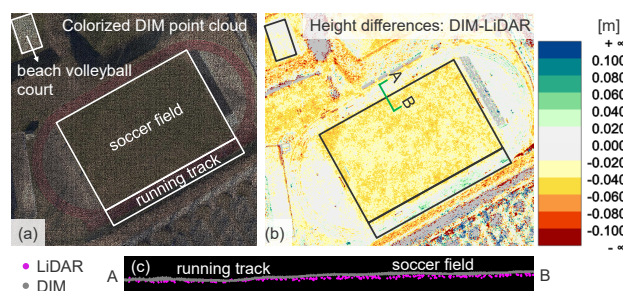

Figure 2. Penetration depths of LiDAR and DIM over different surfaces, (a) RGB-colored DIM point cloud featuring different surface types (white boxes), (b) color coded height differences LiDAR minus DIM, (c) vertical section of profile AB
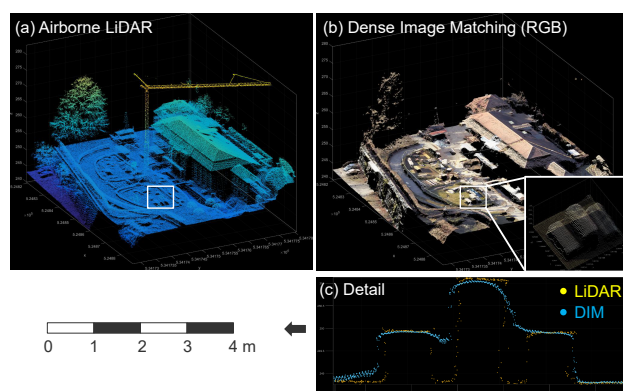

Figure 3. Comparison of 3D point clouds from LiDAR and DIM, (a) LiDAR, (b) DIM (RGB), (c) vertical section of detail (white rectangle) showing a pile of construction material

sitions. Whereas modern digital cameras allow high along-track overlaps (typically $\geq 80\%$) minimizing the occlusion problem, still polar measurements are advantageous whenever the object appearance changes rapidly when seen from different positions (i.e. semi-transparent objects like vegetation, crane bars), when the objects are in motion (vehicles, pedestrians, etc.), or in very narrow canyons. Figure 3c gives an example for the latter. The detail shows three piles of construction material (height: 1-2 m, width: 1 m) with a narrow 50 cm gap between each pile. Whereas the laser beam occasionally reaches the ground between the piles, both stereo-occlusion and cast shadows prevent the generation of DIM ground points. Whereas this is an extreme example, the same situation also occurs in narrow alleys and small courtyards surrounded by high buildings.

Figure 3c also shows artificial ramps around the piles caused by the smoothness constraint within the DIM framework. The heights between the ground level and the pile's top level show a smooth transition in the DIM point cloud compared to the more abrupt height jump in the LiDAR point cloud. Within this transition zone, flying points occur in the DIM dataset which do not represent real targets. The simultaneous use of laser scanners and cameras shows a clear potential in this respect, as the higher reliability of the LiDAR points can be utilized for DIM-based surface reconstruction, either within image matching or later during point cloud fusion and DSM derivation.

Also the LiDAR point cloud shows averaging effects at the boundary of piles as well caused by the finite laser footprint hitting both vertical and horizontal surface parts. The respective laser echo is calculated along the laser beam axis with an averaged distance derived from the broadened echo waveform. Analysis of the backscattered waveform allows the identification of such low quality points as the width of the echo is much longer than the width of the emitted pulse in such situations.

Figures 2c and 3b/c show a clear advantage of DIM over LiDAR, namely the higher point density. To sum it up, advantages of LiDAR w.r.t. higher reliability and less occlusion are compensated by the higher point density of DIM. Together with inherent differences concerning the penetration of semi-transparent objects, these properties need to be considered during data processing to achieve optimal results.

3. DATA ACQUISITION

For testing existing and implementing innovative data processing strategies, an area around the historic town of Melk (located at the Danube River in the eastern part of Austria; N 48 °13 '37 ", E 15 °20 '38 "; WGS 84) was selected. This specific study area

Figure 4. Study area city of Melk (Austria), center: planned flight trajectories (black lines) and camera positions (red circles), upper right: aerial image of the abbey and city of Melk, bottom left: location of study area (red circle) within Austria

was chosen because it contains different types of landscape comprising deciduous and coniferous forests, cropland, water bodies, as well as suburban and urban areas, including the challenging narrow alleys in the historic part of the city of Melk (Figure 4).

In January 2016 (leave-off season), sample data of 9 parallel flight lines were acquired from 600 m above ground level (AGL) with the *RIEGL* LMS-Q1560 compact, dual channel full waveform laser scanner system mounted on a Diamond DA42 light aircraft. Three additional cross strips were captured to enable optimum flight block adjustment. The system features a fully integrated aerial medium format camera and a high-grade INS/GNSS system.

The two laser channels of the LMS-Q1560 have a field of view (FOV) of 58 ° and are rotated around the vertical axis of the system providing a ± 8 ° forward/backward looking capability at the border of the scan strips. The sample data were acquired with a laser pulse repetition rate of 2 x 400 kHz at a flying speed of 110 knots. This resulted in an average point density of 14 points/m^2.

The aerial images were captured with a PhaseOne iXU-R 180, a 80 MPix RGB CCD camera equipped with forward motion compensation. Image acquisition was carried out with a 50 mm lens providing a FOV of 56.2 ° and a nominal image and side overlap of 80%. At 600 m AGL the resulting GSD is 6.2 cm.

4. DATA PROCESSING

Typical processing chains for LiDAR and DIM are summarized in Figure 5 and briefly discussed in Sections 4.1 and 4.2. Integrated data processing options are presented in Section 4.3.

4.1 LiDAR workflow

During data acquisition, the measurements of the inertial navigation system and the laser scanner are time-stamped based on a common time frame. In post-processing the IMU and GNSS measurements are combined to a trajectory providing the position and attitude of the platform over time.

The (line) angle of each laser shot is determined during data acquisition. For full-waveform scanners, the range of each laser echo is calculated offline in post-processing. First, the precise time stamp of each echo is calculated within the full waveform analysis (Ullrich and Pfennigbauer, 2011). For flight missions with ranges exceeding the maximum unambiguous measurement range, more than one laser pulse is in the air at a time. This is referred to as multiple-time-around (MTA). The correct association of each received echo signal to its causative emitted laser pulse is established during the MTA resolution (Rieger, 2014).

Based on the trajectory, the system's mounting calibration (i.e. position/orientation offset between trajectory and scanner coordinate system), and the range and angle measurements of the laser scanner, the 3D coordinates of each detected laser echo can be calculated. Any error in the trajectory solution, the mounting calibration, or the sensor calibration 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 an adjustment to minimize the offsets between the strips (Glira et al., 2015a).

The remaining part of the LiDAR processing pipeline consists of ground-point filtering, break line detection, and surface and terrain model interpolation (Pfeifer and Mandlburger, 2008). For the study at hand, the described workflow was accomplished with the *RIEGL* RiPROCESS software suite and the scientific laser scanning software OPALS (Pfeifer et al., 2014).

4.2 DIM workflow

For the Dense Image Matching the SURE workflow (Rothermel et al., 2012; Wenzel et al., 2013) was utilized, which is based on a dense stereo matching algorithm in a multi-view stereo environment (Figure 5). After image orientation via bundle block adjustment, pixel-wise disparity information is determined for every stereo pair using a variation of the Semi Global Matching (SGM) algorithm (Hirschmuller, 2008) which enables dense stereo through an efficient approximated global optimization while aiming to preserve sharp edges and discontinuities.

Opposed to the original SGM method, a hierarchical strategy is used within the SURE workflow (the tSGM approach) where the disparity search range is reduced successively to reduce complexity. This is particularly beneficial for scenes with large parallax variations. Thereafter, the dense disparity information of each stereo pair is utilized within a multi-stereo forward intersection where for every pixel all available correspondences are used to generate a 3D point. The dense disparity maps enable correspondence linking through the epipolar geometry as well as outlier rejection and noise reduction. Consequently, one 3D point is generated for every pixel featuring sufficient stereo information.

After dense matching, redundant surface observations from different images are exploited in object space for further outlier rejection and filtering either in 3D or in 2.5D. For the study at hand, the 2.5D approach was used where the points are filtered in a raster grid in order to produce a Digital Surface Model (DSM). Within every raster grid cell, the highest points are clustered followed by a median based filtering. Due to the availability of the airborne LiDAR point cloud, a minimum number of consistent observations constraint was established in order to maintain only the high quality DIM measurements. Tests were performed with different parameters (1, 2, 4 and 8 consistent observations per cell).
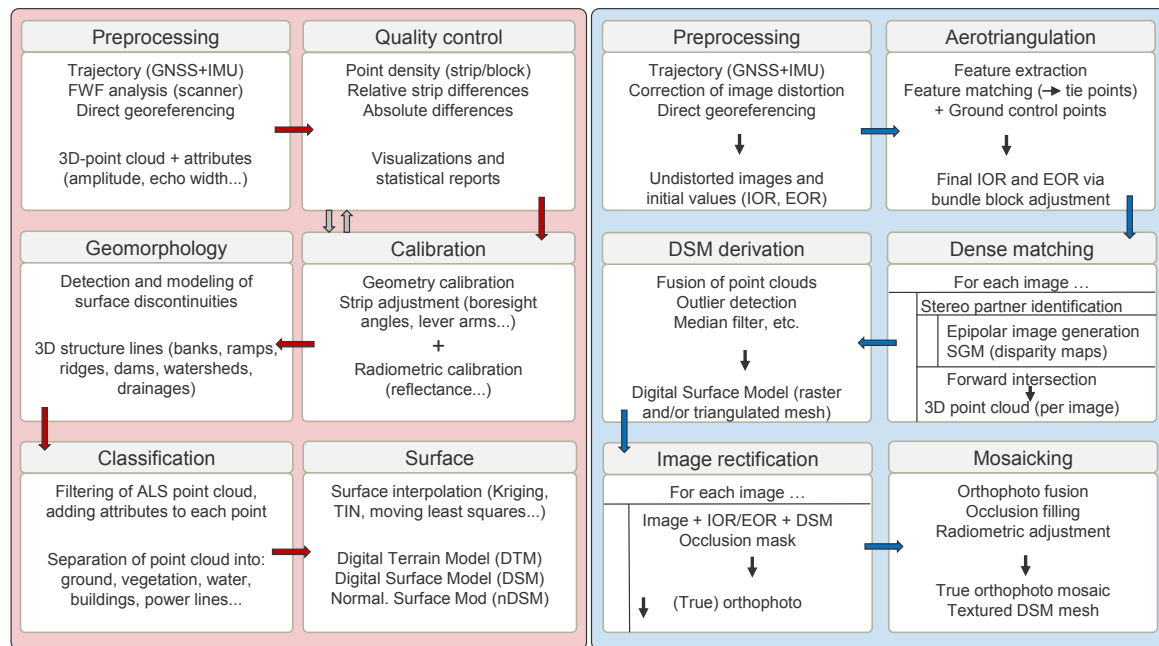
Figure 5. Schematic diagram of typical LiDAR and DIM processing chains

The resulting product is a 2.5D DSM represented either by a height image raster or a point cloud containing the averaged color information of the original imagery with up to 4 bands. Due to the properties of the tSGM approach, sharp edges can be preserved while being able to provide depth information in radiometrically challenging areas - such as in the presence of image noise. The height precision depends on both the geometric configuration (image scale/GSD, intersection angle, redundancy) and the radiometric information (texture quality). As a property of this passive approach, a sufficient signal-to-noise ratio is required to provide reliable surface information.

The DSM is generated in tiles, followed by an optional interpolation across multiple tiles. The SURE workflow furthermore provides capabilities for the production of True Orthophotos and triangulated surfaces (meshes) which are based on the extracted geometry. For the True Orthophotos, the locally involved images are rectified and merged according to the DSM. Meshes can either be derived in 3D space or from the DSM, as used within this study for visualization purposes. For DSM generation and the production of True Orthophoto or meshes, also additional or modified data can be incorporated into the processing pipeline. Within this study, the LiDAR point clouds were introduced as described in Section 4.3.2.

4.3 Hybrid data processing

4.3.1 Integrated sensor orientation Strip adjustment (SA) of LiDAR strips and bundle block adjustment (BBA) of aerial images are usually performed independently. As a consequence, systematic discrepancies between the DIM and LiDAR point clouds can be observed in general. To minimize these discrepancies, LiDAR data and image data have to be oriented and calibrated in a single hybrid adjustment, in which SA and BBA are integrated in a rigorous way.

The strip adjustment of LiDAR strips, which serves as a foundation for the hybrid adjustment proposed herein, was previously published in Glira et al. (2015a), Glira et al. (2015b), and Glira et al. (2016). Summarizing, this strip adjustment method . . .

- establishes correspondences within the overlapping parts of the LiDAR strips in an iterative manner as in the well-known Iterative Closest Point (ICP) algorithm
- takes the original scanner and trajectory measurements as mandatory input data
- takes ground-truth data, e.g. ground control points, as optional input
- performs an on-the-job calibration of the entire LiDAR multi-sensor system
- corrects the flight trajectory errors individually for each strip

As can be seen in Figure 6, for the strip adjustment correspondences are established between (a) overlapping LiDAR strips (LAS-to-LAS) and (b) LiDAR strips and ground-truth data (LAS-to-GTD). The cost function minimizes the sum of squared (point-to-plane) distances between all these correspondences simultaneously. The unknowns are the mounting calibration, the flight trajectory errors, and calibration parameters of the laser scanner.

Building on the above SA framework, BBA of aerial images was integrated in a hybrid adjustment. For this, additional correspondences are established between (c) the tie points of over-
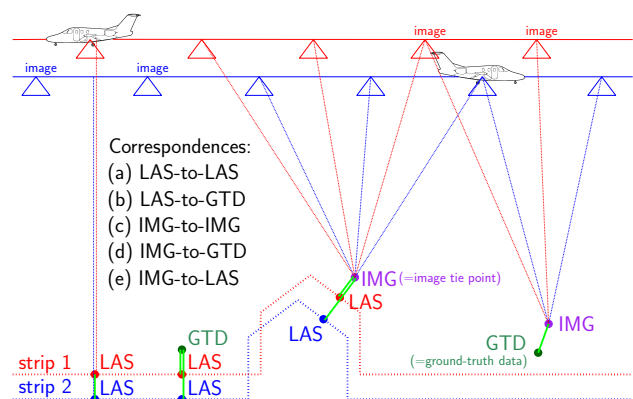

Figure 6. Schematic diagram for integrated orientation of LiDAR scans and images.

lapping images (IMG-to-IMG), (d) image tie points and ground-truth data, e.g. GCPs (IMG-to-GTD), and (e) image tie points and LiDAR strips (IMG-to-LAS). Whilst the correspondence categories (a) and (b) are commonly used in SA, and (c) and (d) are commonly used in BBA, the correspondences of category (e) are the important connecting link between the LiDAR strips and the aerial images. By minimizing the squared point-to-plane distances (point from images, plane from LiDAR strips) for each of these correspondences, the relative orientation between the LiDAR strips and the aerial images is optimized in the hybrid adjustment. Additional unknowns in the hybrid adjustment are, thus, the exterior and inner orientations of the photos and the coordinates of the image tie points.

4.3.2 Point cloud fusion The fusion of point clouds from LiDAR and DIM follows the general strategy as described in section 4.2. The point clouds of both sources are utilized simultaneously in a 2.5D grid filtering step where the highest points are clustered and filtered using a median based approach. For the DIM point cloud, a DSM point cloud (DIM-DSM) in the identical raster scheme and GSD is used while preserving only the points with highest quality by setting the minimum number of consistent observations per cell to 8. Wherever the DIM results do not meet this rigorous criterion, remaining gaps are complemented by LiDAR points. Besides this constraint, a filter for gross errors (disparity outliers defined by certain area and distance to the surface) is used during the extraction of the DSM from the raw DIM point cloud. The resulting DIM-DSM and LiDAR point clouds are then used in a repeated mutual raster filtering step with the minimum observations per cell set to 1. Further filters, such as the speckle filter, are deactivated in order to maintain all points introduced to the fusion process.

Subsequently, the interpolation across multiple tiles is carried out using the path based interpolation proposed by Hirschmuller (2008) which propagates height information from the local height level. This heuristic approach prevents the mixing of roof and ground information for data gaps on the ground, since points are only interpolated from other ground points. This is particularly useful for urban environments where typically roof information is available, while ground surfaces suffer from occlusions and low radiometric information due to shadows.

Even though a focus of this study is the geometric benefit of the fused point cloud, further products benefiting from the improved geometry such as a True Orthophoto and a textured DSM Mesh are produced for comparison. In summary, the following steps are carried out for the generation and fusion of point clouds. Whereas step 1 is carried out with the OPALS software (Pfeifer et al., 2014), all subsequent steps are based on the SURE workflow (Wenzel et al., 2013):

1. Integrated aerial triangulation and LiDAR strip adjustment
2. DIM point cloud generation (raw 3D points)
3. DIM Digital Surface Model extraction (DIM-DSM, intermediate product)
4. Final Digital Surface Model extraction (DIM-DSM + LiDAR)
5. DIM-LiDAR True Orthophoto and DSM mesh extraction

5. RESULTS AND DISCUSSIONS

The processing strategies outlined in Section 4.3 were applied to the city of Melk flight block and the respective results are presented and discussed in the following section.

Initially, the orientation of LiDAR strips and images was carried out independently. The LiDAR flight block was processed with the OPALS software (Pfeifer et al., 2014) based on the approach of Glira et al. (2015a) and image aerial triangulation was performed with Trimble Match-AT. For the LiDAR block, the precision was estimated by analyzing the height differences of all overlapping strip pairs within smooth areas. The deviations exhibited a robust standard deviation σ_{MAD}=1.2 cm. The image orientation was checked at 9800 image tie points and showed a standard deviation of 0.1 pixel (GSD: 6 cm). In contrast to this very good precision of the individual data sources, comparison of the LiDAR-DSM and the DIM-DSM showed unexpected systematic deviations at impervious surfaces (Figure 7a) with a bias (median) of -4.3 cm and a dispersion σ_{MAD}=4.1 cm. Whereas this represents the situation after orientation of scans and images separately, a substantial improvement could be achieved with the integrated orientation approach described in Section 4.3.1. As can be seen from the dominant whitish colors in Figure 7b the relative discrepancies between the LiDAR and the DIM point cloud could effectively be reduced (bias: 0.2 cm, dispersion: σ_{MAD}=3.3 cm).

The resulting image orientations were subsequently used as starting point for the DIM point cloud generation and the point cloud fusion approach described in Section 4.3.2. Figure 7b still reveals areas with pronounced vertical deviations between the LiDAR and the DIM dataset of more than 10 cm. Most of the large discrepancies occur in vegetated areas and rather stem from the different measurement principle than from sensor or data processing related deficiencies. Apart from this, problematic areas (white boxes marked in Figure 7b) were detected and are discussed below.

The harsh lighting conditions during data acquisition (very bright and low standing winter sun) resulted in cast shadows and caused sub-optimal conditions for DIM surface reconstruction especially in the area of narrow alleys and small inner courtyards. An extreme example is displayed in Figure 8. The colorized DIM point cloud of two building blocks, each with a small inner courtyard, is displayed in Figure 8a. At the transition from roof-level to courtyard-level a large tree (only visible in the sectional view of Figure 8b) causes void areas in the DIM point cloud. Furthermore, despite the high image and strip overlap of 80 %, the low image contrast prohibits proper DIM reconstruction of the balcony located at the first floor level of the inner, eastern wall of the eastern building block while the active, polar LiDAR technique allows proper reconstruction (cf. Figure 8b). Finally, the DIM point cloud shows smoothing effects at the roof ridges. Although, in general, LiDAR is also prone to smoothing, the effect is limited to the laser footprint as already discussed in Section 2.3.

In all the mentioned cases, the availability of the additional LiDAR point cloud greatly helps to improve the DIM-DSM by substituting low quality DIM points with reliable information from LiDAR. A final example, where the strategy outlined in Section 4.3.2 was successfully applied, is displayed in Figure 9. While reconstructing the ground floor of the narrow alley between the church and the adjacent southern wing of the surrounding abbey complex was not possible with DIM due to stereo occlusion and low radiometric texture, the fusion strategy successfully marked these points as low-quality and, hereby, cleared the way for complementing the DSM with information from the LiDAR point cloud. Thus, by concurrently processing image and scan data, the high spatial resolution of DIM and reliability of LiDAR could be preserved in the final DSM product with a positive impact on follow-up products like the True Orthophotos or

Figure 7. Color coded height differences LiDAR vs. DIM before (a) and after (b) simultaneous orientation of scans and images
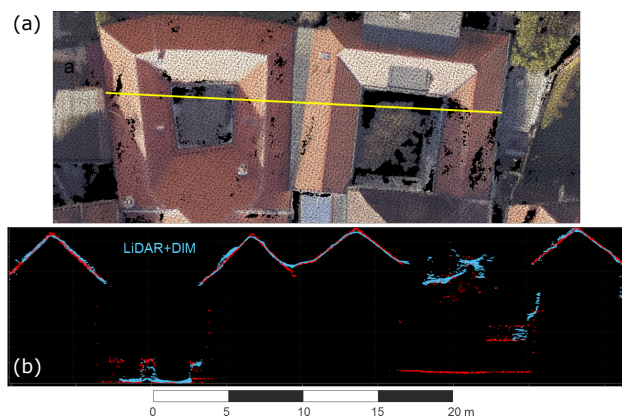
meshes.

Figure 8. Comparison of a complex profile, (a) plan view of RGB colored DIM point cloud, (b) section view, LiDAR points (red), DIM points (blue)
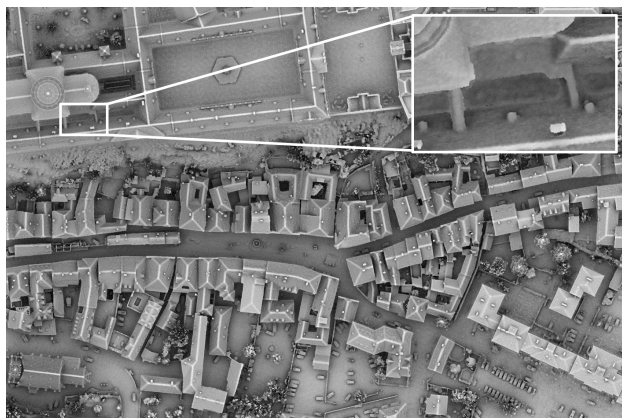

Figure 9. Ambient occlusion shading of hybrid LiDAR-DIM DSM mesh; detail: narrow alley within the abbey with roof points stemming from DIM and ground floor points from LiDAR

6. CONCLUSIONS

An experiment based on the simultaneous acquisition of airborne LiDAR and photos was conducted. Point clouds acquired from either source, i.e. direct georeferencing of the polar LiDAR mea-

surements and multiple ray forward intersection after dense image matching, were compared. Given the higher sampling of the image in comparison to the LiDAR, also the point cloud from dense image matching has a higher density (DIM: 280 points/m^2 vs. LiDAR: 14 points/m^2). Although the image matching point cloud is denser by a factor of 20, the roof ridges appear rounded off more in the DIM point cloud than in the LiDAR point cloud (Figure 8). The simultaneous acquisition allowed investigating differences of the sensing technologies, rather than comparing two epochs of a changing world, acquired with two different methods.

The simultaneous acquisition of airborne LiDAR and photos and subsequent computation of point clouds from either source independently confirmed expected differences for tall vegetation (Figure 3a, b). In leaf-off data acquisition, as was the case in the experiment conducted, the crown is captured well by laser scanning. Image matching provided a comparatively incomplete representation. In contrast, under leaf-on conditions, as shown, e.g., by Ressl et al. (2016), both methods capture the top surface of the crown well, but laser scanning also captures the ground surface due to its multi-target capability.

The simultaneous acquisition allowed also to study subtle differences in vegetation. The grass of a soccer field lead to height differences between airborne LiDAR and image matching, with the image matching result lying 4 cm higher. It is noteworthy, that either surface model appeared equally smooth.

Linear structures above the ground are captured by LiDAR, because of the polar measurement principle (e.g. the crane in 3a). Due to the multi-target capability, also the ground below a beam or a wire (along the measurement line of sight) is captured. The DIM used in our experiment, in contrast, assumes that the surface is a function parameterized over (e.g.) image space. Thus, only one height can be reconstructed along the measurement ray.

The simultaneous acquisition also allowed exemplary investigation of the effective geometric shadow caused by vertical objects not facing towards but away from the flight line. In Figure 9 the advantage of the combined use of airborne LiDAR and DIM was demonstrated by filling gaps of the higher density image matching point cloud with LiDAR data from a narrow alley. The same applies to areas where matching was not successful (tree example of Figure 8). Dark cast shadow areas and materials reflecting highly in the visible domain of the EM spectrum, but poorly at IR

wavelengths, may provide further examples for mutual completion of the measurement technologies, but were not encountered in this experiment.

The simultaneous acquisition may also have disadvantages. The environmental circumstances (low standing sun and very high contrast) caused suboptimal conditions for image matching, but did – in this case – not influence laser scanning. Flying height above ground, flying date (time and season), strip overlap and image overlap, etc., influence the quality of the final products. Optimal conditions for simultaneous acquisition of airborne LiDAR and photos from one platform may provide additional insight in the advantages of either method. A repeat data acquisition of the study area with the same sensor system but under leaf-on conditions is currently being prepared. However, also different typical (i.e. not optimal) acquisition scenarios are required, to better understand the relative contribution of either source.

The different properties of the point clouds from laser scanning and image matching are rooted in the different acquisition technologies and need to be considered in a joint orientation of both data sets. Only hard surfaces are suitable tie elements in a combined orientation.

Our investigation also demonstrated the advantage such a simultaneous acquisition can offer for classification. At hard surfaces, height differences between the two point clouds disappear, thus offering a possibility to support classification of surface materials.

ACKNOWLEDGEMENTS

References

Förstner, W. and Wrobel, B. P., 2016. *Photogrammetric Computer Vision: Statistics, Geometry, Orientation and Reconstruction.* Springer International Publishing, Cham, Switzerland, pp. 643–725.

Glira, P., Pfeifer, N. and Mandlburger, 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., Briese, C. and Ressl, C., 2015a. A Correspondence Framework for ALS Strip Adjustments based on Variants of the ICP Algorithm. *PFG Photogrammetrie, Fernerkundung, Geoinformation* 2015(4), pp. 275–289.

Glira, P., Pfeifer, N., Briese, C. and Ressl, C., 2015b. Rigorous strip adjustment of airborne laserscanning data based on the ICP algorithm. In: *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. II-3/W5, pp. 73–80.

Haala, N. and Rothermel, M., 2012. Dense Multi-Stereo Matching for High Quality Digital Elevation Models. *PFG Photogrammetrie, Fernerkundung, Geoinformation* 2012(4), pp. 331–343.

Hirschmuller, H., 2008. Stereo Processing by Semiglobal Matching and Mutual Information. *IEEE Trans. Pattern Anal. Mach. Intell.* 30(2), pp. 328–341.

Kraus, K., 2007. *Photogrammetry – Geometry from Images and Laser Scans.* 2 edn, De Gruyter, Berlin, Germany.

Pfeifer, N. and Mandlburger, G., 2008. Filtering and DTM generation. In: J. Shan and C. Toth (eds), *Topographic Laser Ranging and Scanning: Principles and Processing*, CRC Press, pp. 307–333.

Pfeifer, N., Mandlburger, G., Otepka, J. and Karel, W., 2014. OPALS – a framework for airborne laser scanning data analysis. *Computers, Environment and Urban Systems* 45, pp. 125 – 136.

Remondino, F., Spera, M. G., Nocerino, E., Menna, F. and Nex, F., 2014. State of the art in high density image matching. *The Photogrammetric Record* 29(146), pp. 144–166.

Ressl, C., Brockmann, H., Mandlburger, G. and Pfeifer, N., 2016. Dense Image Matching vs. Airborne Laser Scanning - Comparison of two methods for deriving terrain models. *PFG Photogrammetrie, Fernerkundung, Geoinformation* 2016(2), pp. 57–73.

Ressl, C., Kager, H. and Mandlburger, G., 2008. Quality checking of ALS projects using statistics of strip differences. In: *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. XXXVII. Part B3b, pp. 253–260.

Rieger, P., 2014. Range ambiguity resolution technique applying pulse-position modulation in time-of-flight scanning lidar applications. *Optical Engineering* 53(6), pp. 061614–1 – 061614–12.

Rothermel, M., Wenzel, K., Fritsch, D. and Haala, N., 2012. SURE: Photogrammetric surface reconstruction from imagery. In: *Proceedings of the Low Cost 3D Workshop, Berlin.*

Shan, J. and Toth, C. K. (eds), 2008. *Topographic Laser Ranging and Scanning: Principles and Processing.* CRC Press, Boca Raton, FL.

Skaloud, J. and Lichti, D., 2006. Rigorous approach to boresight self-calibration in airborne laser scanning. *ISPRS Journal of Photogrammetry and Remote Sensing* 61(1), pp. 47–59.

Ullrich, A. and Pfennigbauer, M., 2011. Echo digitization and waveform analysis in airborne and terrestrial laser scanning. In: D. Fritsch (ed.), *Photogrammetric Week '11*, Wichmann/VDE Verlag, Berlin & Offenbach, pp. 217–228.

Ullrich, A. and Pfennigbauer, M., 2016. Linear LIDAR versus Geiger-mode LIDAR: impact on data properties and data quality. In: *Proc. SPIE*, Vol. 9832, pp. 983204–1 – 983204–17.

Vastaranta, M., Wulder, M. A., White, J. C., Pekkarinen, A., Tuominen, S., Ginzler, C., Kankare, V., Holopainen, M., Hyyppä, J. and Hyyppä, H., 2013. Airborne laser scanning and digital stereo imagery measures of forest structure: Comparative results and implications to forest mapping and inventory update. *Canadian Journal of Remote Sensing* 39(5), pp. 382–395.

Vosselman, G. and Maas, H. (eds), 2010. *Airborne and Terrestrial Laser Scanning.* Whittles Publishing, UK.

Wenzel, K., Rothermel, M., Haala, N. and Fritsch, D., 2013. SURE – the ifp Software for Dense Image Matching. In: D. Fritsch (ed.), *Photogrammetric Week '13*, Wichmann/VDE Verlag, Berlin & Offenbach, pp. 59–70.

Yan, L., Fei, L., Chen, C., Ye, Z. and Zhu, R., 2016. A multi-view dense image matching method for high-resolution aerial imagery based on a graph network. *Remote Sensing* 8(10), pp. 18.