COMBINING AIRBORNE OBLIQUE CAMERA AND LIDAR SENSORS: INVESTIGATION AND NEW PERSPECTIVES

I. Toschi [1], F. Remondino [1], R. Rothe [2], K. Klimek [2]

[1] 3D Optical Metrology (3DOM) unit, Bruno Kessler Foundation (FBK), Trento, Italy, (toschi, remondino)@fbk.eu
[2] Leica Geosystems AG, Geospatial Content Solutions, Heerbrugg, Switzerland, (rene.rothe, kristin.klimek)@leica-geosystems.com

Commission I, WG I/2

KEY WORDS: oblique photogrammetry, airborne laser scanning, quality evaluation, dense image matching, data fusion

ABSTRACT:

Hybrid sensor solutions, that feature active laser and passive image sensors on the same platform, are rapidly entering the airborne market of topographic and urban mapping, offering new opportunities for an improved quality of geo-spatial products. In this perspective, a concurrent acquisition of LiDAR data and oblique imagery, seems to have all the potential to lead the airborne (urban) mapping sector a step forward. This contribution focuses on the first commercial example of such an integrated, all-in-one mapping solution, namely the Leica CityMapper hybrid sensor. By analysing two CityMapper datasets acquired over the city of Heilbronn (Germany) and Bordeaux (France), the paper investigates potential and challenges, w.r.t. (i) number and distribution of tie points between nadir and oblique images, (ii) strategy for image aerial triangulation (AT) and accuracy achievable w.r.t ground truth data, (iii) local noise level and completeness of dense image matching (DIM) point clouds w.r.t LiDAR data. Solutions for an integrated processing of the concurrently acquired ranging and imaging data are proposed, that open new opportunities for exploiting the real potential of both data sources.

1. INTRODUCTION

Multi-sensor data integration is the new trend in terrestrial and aerial 3D mapping applications, as demonstrated by the quickly growing market sector of hybrid mapping systems, that feature complementary ranging and imaging sensors on the same (mobile) platform. Particularly, these hybrid solutions are rapidly entering the airborne market of topographic and urban mapping, offering new opportunities for an improved quality of geo-spatial products. Indeed, since some years, both airborne LiDAR (light detection and ranging) and digital photogrammetry have become state-of-the-art techniques to obtain 2.5D/3D value-added spatial information products, and a joint exploitation of their specific advantages is nowadays the promising next step.

From the one hand, airborne LiDAR provides for high reliability of height information and multi-target capability, thus enabling the penetration of vegetation for bare ground acquisition and modelling. These strengths have made ALS (airborne laser scanning) the state-of-the-art mapping method for deriving digital terrain model and 2.5D point clouds, as well as for the automatic generation of building models (Shan and Toth, 2010; Vosselman and Maas, 2010; Awrangjeb and Fraser, 2014). On the other hand, advances in photogrammetry and computer vision, particularly those related to the developments of innovative dense image matching (DIM) algorithms (Haala and Rothermel, 2012; Remondino et al., 2014), have increased automation in image-based 3D reconstruction of scenes, with the goal of generating high spatial resolution digital surface models. Indeed, given a suitable redundancy and a good geometric configuration of image rays, photogrammetric point clouds can today feature a spatial resolution equal to the GSD (ground sample distance) of the original imagery, and an optimal vertical accuracy below the GSD level. Furthermore, the overall dense matching quality improves when oblique images are included in the block (Zhang et al., 2018). Indeed, the rapidly maturing sector of airborne oblique photogrammetry (Remondino and Gerke, 2015), supported by the development of increasingly more sophisticated sensor and algorithmic solutions, has moved forward the geometric processing towards "true" 3D space, providing for a more complete and precise information extraction in urban scenarios (Haala and Rothermel, 2015; Remondino et al., 2016). Particularly, DIM from multi-view aerial blocks can be an effective solution to overcome the problem of viewpoint restrictions and to generate an high resolution 3D input for the automatic building reconstruction (Toschi et al., 2017a).

Investigating the potentials of each method and the advantages of integrating their complementary information, still represents an open research topic, as demonstrated by the high volume of papers published in this domain. Particularly, previous works were mainly aimed to (i) provide a general comparison between airborne LiDAR and DIM, to fully understand their specific properties (Leberl et al., 2010; Ressl et al., 2016; Tian et al., 2017); (ii) evaluate the quality of DIM point clouds, using LiDAR data as reference (Nex et al., 2015; Maltezos et al., 2016; Zhang et al., 2018); (iii) develop approaches for the registration of optical imagery and LiDAR point clouds (Rönnholm and Haggrén, 2012; Parmehr et al., 2014); (iv) present methods for multi-sensor data fusion within various application fields, e.g. true orthophoto generation, classification, building reconstruction and forest inventory, as reviewed by Zhang and Lin (2017).

1.1 Leica CityMapper hybrid sensor

In this perspective, a simultaneous multi-sensor data acquisition from the same platform can further support the (automatic) fusion of the derived complementary information. Particularly, combining LiDAR and multi-camera systems (including both nadir and oblique looking cameras) on the same airborne platform, seems to have all the potential to lead the airborne mapping sector a step forward, especially for 3D city modelling and urban mapping applications.

The first commercial example of such an integrated, all-in-one mapping solution is already available in the market, namely the Leica CityMapper hybrid airborne sensor (Figure 1). Designed for the mapping of urban areas, Leica CityMapper combines:

- a Hyperion LIDAR unit (theoretical ranging accuracy < 2 cm at 1,000 m AGL, 60 m/s aircraft speed), featuring a

linear mode LiDAR sensor with circular pattern, pulse repetition rate up to 700 kHz, capable of handling up to 15 returns and full waveform recording;

- a multi-camera system, featuring one nadir looking camera head (including a RGB CCD size 10,320 x 7,752 pixels – 5.2 μm pixel size and a NIR CCD size 3,336 x 4,500 pixels - 12 μm pixel size, 83 mm focal length) and four 45°-tilted camera heads (RGB CCD size 10,320 x 7,752 pixels - 5.2 μm pixel size, 156 mm focal length).

As the world's first example of hybrid mapping solution integrating oblique and nadir imaging as well as LiDAR sensor into one system, it offers the unique basis for testing the geometric performance of a concurrent multi-data acquisition and the development of innovative processing and mapping strategies.

Figure 1. The Leica CityMapper hybrid sensor

1.2 Paper objectives

This contribution aims to investigate the potentials and challenges of combining oblique photogrammetry and LiDAR in the same airborne acquisition system. By processing and analysing two datasets acquired in Europe with the Leica CityMapper hybrid sensor (Section 2), we want to assess the geometric quality of the acquired oblique imagery for photogrammetric purposes, with special focus on (i) number and distribution of tie points between nadir and oblique images, (ii) strategy for image aerial triangulation (AT) and accuracy achievable w.r.t ground truth data, and (iii) local noise level and completeness of DIM point clouds w.r.t LiDAR data.

In addition, we discuss the need for an integrated (automatic) processing of the ranging and imaging data in order to improve their co-registration and exploit the full potential of both data sources, starting from the analysis of the existing workflow solutions (Section 3). While potential synergies of jointly processing LiDAR data and nadir imagery are proposed by Mandlburger et al. (2017), here we propose and test new

perspectives for fully integrating the processing of oblique images and LiDAR data and cope with sensor-specific issues. These include (i) the use of LiDAR points to improve the matching between different camera views and support image AT and geo-referencing, and (ii) the combined exploitation of oriented oblique images and LiDAR data for efficient visualization and measurement purposes.

2. CASE STUDIES AND INPUT DATA

In order to test the existing data processing workflows and propose alternative strategies to further improve them, two datasets acquired with the Leica CityMapper system are selected, namely the data captured over the city of Heilbronn (Germany) and Bordeaux (France).

In case of Heilbronn, located on the Neckar River in southwest Germany, the flight plan was designed using an average nadir GSD of 12 cm, and along-across overlaps of 80% and 60%, respectively. At 1,900 m AGL, it resulted in 14 strips and a total of 3,050 images (Figure 2, left), i.e. 610 images for each camera. The area coverage of the images acquired during a single exposure (including one nadir and four oblique images) is showed in Figure 2, centre. The flight was executed in September 2017 with favourable weather conditions, and covered a total area of ca. 7 km x 7 km. Ground truth data was provided in the form of 125 control points, surveyed with RTK GNSS with a mean 3D accuracy of 5 cm. The concurrent LiDAR acquisition was planned to have an average point density of ca. 7 points/m^2 (last return); a sample of the derived point cloud is shown in Figure 2, right. Thanks to the availability of well-distributed reference points, this dataset was selected as case-study for the quality investigations of existing and improved processing strategies (Section 4), after having identified a smaller area of interest (AOI), that includes a total of 460 images. The AOI covers an area of ca. 3.5 km x 3.5 km over the city centre of Heilbronn (Figure 2, left).

The CityMapper data over the city of Bordeaux, located on the Garonne River in southwestern France, included 55 stripes and a total of 27,705 images, with a mean nadir GSD of 5 cm. The block was flown at 900 m AGL, with along-across overlaps of 80% and 60%, respectively. Concurrently, a LiDAR point cloud featuring a mean point density of ca. 10 points/m^2 (last return) was acquired. As no ground truth is available, the dataset was adopted to test the combined use of oriented oblique images and LiDAR-derived digital terrain model, as support for the extraction of vector information and precise measurements for mapping purposes (Section 5).

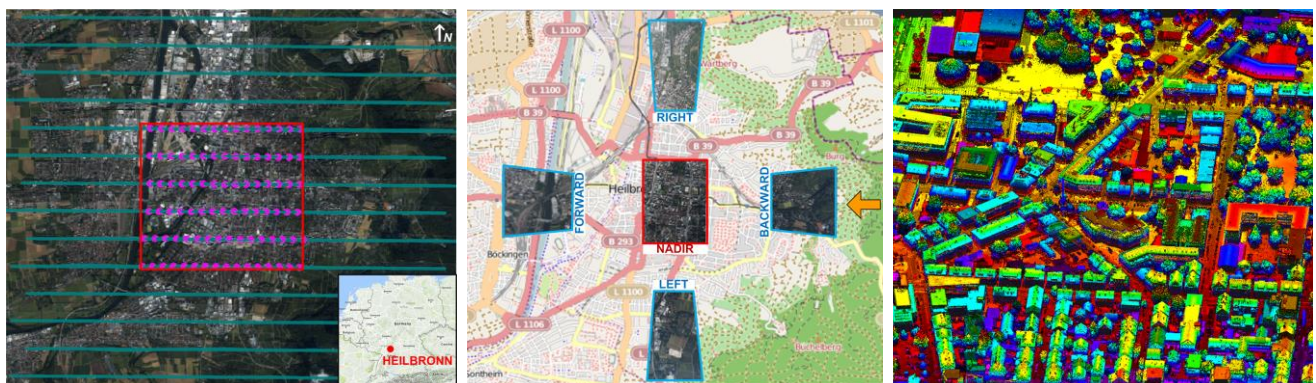The main specifications of both datasets are listed in Table 1.

Figure 2. The Heilbronn CityMapper dataset. Left: the planned flight trajectories (blue lines) and camera positions of the selected image subset in the AOI (pink circles within the red square). Centre: the image footprint of one single exposure (nadir image in red, oblique images in blue). Right: a view of the LiDAR point cloud over the same area (height colour scale).

Dataset	Tot. # images (nadir + oblique)	# Images in the AOI	Image GSD (nadir)	Image overlap (along/across)	LiDAR point density (last return)	Flying height	# Control points
Heilbronn	3,050	460	12 cm	80/60 %	7 points/m^2	1,900 m	125
Bordeaux	27,705	1,205	5 cm	80/60 %	10 points/m^2	850 m	-

Table 1. Main specifications of the two CityMapper datasets.

3. WORKFLOW OF DATA PROCESSING

The data processing was accomplished with the Leica HxMap suite, a unified multi-sensor software solution that integrates the typical processing chains for LiDAR and image data within the same platform (HxMap, 2018). The general workflow is shown as a schematic diagram in Figure 3, and briefly discussed in the following sections, distinguishing between the processing steps for image data (Section 3.1, i.e. yellow steps in Figure 3) and for LiDAR data (Section 3.2, i.e. green steps in Figure 3).

After a first quality control (QC) performed on the raw data, the common pre-processing step is intended to determine the precise trajectory, by post-processing the available on-board Real Time GNSS/IMU data. Based on this refined trajectory, the multi-sensor session is then generated, by applying the sensor calibration parameters, either available within the globally defined sensor repository or estimated with an ad-hoc calibration flight. These include the camera calibration parameters (focal length, principal point, radial symmetric distortion, correction grid and IMU misalignment) and the LiDAR mounting calibration results (position/orientation offsets between trajectory and scanner coordinate system).

3.1 Image processing steps

After the enhancement of image radiometry and the assignment of an improved radiometric profile to the session, the image aerial triangulation step is performed to improve the GNSS/IMU observations. First, image correspondences are identified across the different views, using as input the refined flight trajectory, the calibrated misalignment values and a rough terrain model (digital elevation model, DEM). Particularly, the APM (automatic point matching) step adopts a feature-based method to detect point gradients in different scales and then applying a BRISK-based descriptor (Leutenegger et al., 2011). Keypoints found in the user-selected master images are then transferred to all overlapping images, by exploiting the known external orientations (EO) and the shape of the terrain (see Section 4.1). At present, only a Global DEM (e.g. NASA SRTM or USGS GTopo) can be used to re-project candidate tie points to the overlapping images, while height buffers are added/subtracted to the terrain elevations in order to account for systematic effects in the EO, DEM accuracy and presence of buildings. Finally, a consistency check based on cross-correlation is applied to merge points transferred from different views and filter out outliers.

The automatically extracted homologous points and, when available, the manually measured ground control points (GCPs), are then used to refine image orientations within a bundle block adjustment (BBA). Internal camera parameters, boresight calibration and datum transformation parameters (modelling a similarity transformation between GNSS and photogrammetric data) can be optionally estimated. Although a rigorous triangulation method is available, a simplified approach can be efficiently adopted, which solves the triangulation without matrix inversion until the very last iteration, in order to provide standard deviations and other statistics. Starting from the adjustment results, a dense image matching step based on a

semi-global matching (SGM) approach can be finally applied, to determine pixel-wise disparity information and generate dense point clouds. 2D and 2.5D additional products (i.e. orthophoto and LOD1/2 building models, respectively) can be furthermore generated.

3.2 LiDAR processing steps

A quality control (QC) is initially performed, to evaluate the vertical accuracy of the LiDAR data via patch-based comparisons with (i) control points, (ii) overlapping flight lines, and (iii) forward/backward scans within the same flight line (see Section 4.3). Indeed, offsets and vertical errors may affect the LiDAR data, due to any error in the trajectory solution, the mounting and the sensor calibration. If deviations exceed defined thresholds, first a refined LiDAR misalignment calibration can be estimated. The process exploits planar surfaces tilted in different directions, to estimate the boresight and the system parameters used for point cloud generation. At the moment, the LiDAR calibration method doesn't use GCPs. Furthermore, a LiDAR patch-based matching step can be applied, to correct errors in roll, pitch and Z. Finally, surface/terrain model interpolation and automatic building modelling, can be performed.
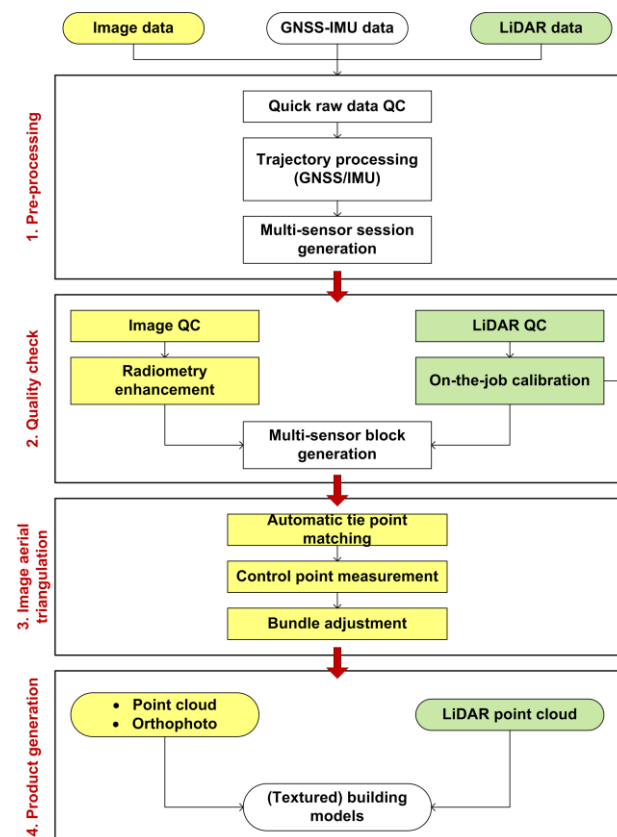

Figure 3. General workflow of multi-sensor data processing (yellow steps refer to image processing, green steps refer to LiDAR processing).

4. CRITICAL INVESTIGATIONS

Using the Heilbronn oblique dataset, we have critically investigated (i) tie point matching across different viewing directions, (ii) bundle block adjustment accuracy within three different AT strategies and (iii) dense image matching quality w.r.t LiDAR data.

4.1 Automatic tie point matching

Compared to standard APM in nadir-only image blocks, tie points extraction is more challenging when different viewing directions should be matched (Gerke et al., 2016; Moe et al., 2016). Indeed, although the lateral parts of objects (e.g. building facades and footprints) get visible and could be theoretically matched, issues like occlusions, large perspective distortions and symmetrical ambiguities should be dealt with. Therefore, APM based on keypoints detectors/descriptors that only consider the local grey value distribution, is likely to fail when matching images with large direction difference.

This issue may be partially addressed if, given the keypoints extracted on selected (nadir and oblique) master images, the search for correspondences in the overlapping images is guided from object space (Figure 4). For a given keypoint p_1 in the oriented master image I_1, is it possible to define the 3D ray C_1p_1 (in red) on which the correspondent position in object space, P, should lie. Therefore, the intersection of this ray with the DEM representing the elevations of the terrain, allows to determine the 3D position of P in objects space. However, errors in image EO and DEM, together with the presence of un-modelled objects in the elevations source (e.g. buildings and trees), may lead to incorrect point positioning. To account for this, a buffer ΔZ is introduced, that force the point position to lie within a certain range along the ray C_1p_1. When points with height values within the interval $(Z \pm \Delta Z)$ are back-projected on the overlapping image I_2 (rays in blue), the search for the homologous point p_2 is spatially limited and can be more efficiently accomplished by cross-correlation based algorithms.
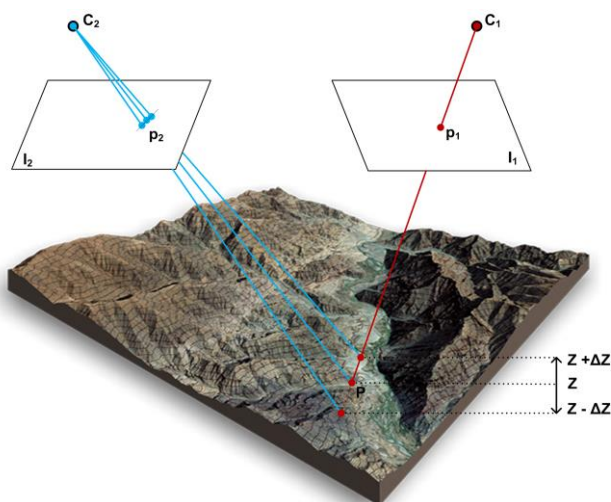

Figure 4. The transfer of keypoints between overlapping images, guided by a-priori knowledge on object space.

This approach, implemented in HxMap with the support of a Global DEM as a-priori knowledge of the terrain shape, leads to a good distribution of tie points in object space (Figure 5, left), and high redundancy of image rays. Indeed, almost 50% of tie points are matched in 6 to 10 images, and almost 40% in 3 to 5

images (Table 2). Furthermore, points matched in more than 6 images are well distributed in the overall block extent (Figure 5, right). This result shows the good performance of the APM approach implemented in HxMap, if compared to the SIFT-based correspondences detection performed by fully automated structure-from-motion (SfM) software. Indeed, by processing the same image block in a typical SfM-based tool, although a higher number of tie points is matched (ca. 8 mil. vs 8,201), their multiplicity in terms of image rays is significantly lower, with more than half homologous points (56.3 %) matched in only two images.

However, an unbalanced distribution of tie points is notable if we analyse the number of matches across the different cardinal views each camera is looking at (Table 3).

# Rays	# Points
2	328 (4.0%)
3-5	3,193 (38.9%)
6-10	3,824 (46.6%)
11-15	752 (9.2%)
16-20	85 (1.0%)
21+	19 (0.2%)
	Tot. 8,201 (100%)

Table 2. The number (and percentages in brackets) of tie points, in relation to the number of images they are matched in.
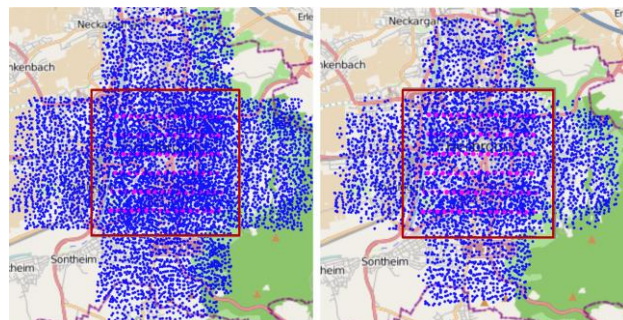

Figure 5. The distribution of tie points in object space: all points matched (left) vs. points matched in more than 6 images (right).

	Down	North	East	South	West
Down	2,132 (26.0 %)	447 (5.5 %)	769 (9.4 %)	685 (8.4 %)	715 (8.7 %)
North		1,902 (23.2 %)	47 (0.6 %)	0 (0.0 %)	177 (2.2 %)
East			1,929 (23.5 %)	142 (1.7 %)	0 (0.0 %)
South				2,044 (24.9 %)	112 (1.4 %)
West					1,896 (23.1 %)

Table 3. The number (and percentages in brackets) of tie points across the different cardinal views.

Figure 6. Example of a tie point matched in five different cameras, observing the scene from different looking direction (e.g. left-right, LE-RI), or from the same looking direction (e.g. forward-backward, FW-BW).

On the one hand, images pointing to the same direction gets almost the same number of correspondences (yellow and green cells in Table 3), i.e. in the range of 23-26% of total matches. On the other hand, nadir-to-oblique image pairs get fewer correspondences (orange cells, 5-9% of total matches), and even less points are matched between images enclosing very different viewing direction (red cells, 0-2% of total matches). This demonstrates that matching is successful between different cameras if they observe the scene from the same direction (like the backward and forward image in Figure 6), whereas images that look the same scene from very different direction can be potentially matched (like the left and right image in Figure 6), but this is generally more challenging.

In order to increase the accuracy of matching across images pointing to the same cardinal direction, and support the matching across different looking directions, an improved version of the APM approach is currently under development. This relies on the LiDAR derived digital surface model as height input source for the geometry-constrained point transfer (Figure 4). The idea behind this approach is to exploit the concurrently acquired height data to provide for a better representation of the scene 3D geometry, including buildings and other ground objects. Besides supporting the transfer of points across different viewing directions, this will have a beneficial effect on the processing time, by limiting the required buffer ΔZ and, consequently, the search area in the overlapping images.

4.2 Aerial triangulation

Although initially georeferenced with the same refined trajectory, the further adjustments of LiDAR strips and of the bundle of image rays (BBA) are then performed separately in HxMap. As a result, a bias between the LiDAR and photogrammetric point clouds can be detected. Moreover, image aerial triangulation still requires field-surveyed control points as ground (soft) constraints, in order to support the adjustment and compensate for the remaining systematic errors between GNSS and image data. Conversely, GCPs are not necessary for LiDAR calibration and strip adjustment.

In order to reduce the bias between the point clouds and avoid the need for field-surveys ground control points within the AT, an alternative strategy is here investigated.

It relies on the use of LiDAR-derived 3D points as ground control information: after the mounting calibration of LiDAR unit, and the point matching to minimize errors in roll, pitch and Z, few and well distributed 3D points were extracted from the LiDAR point cloud and included as observed unknowns in the BBA. Different configurations of GCPs were tested, to find the best compromise between the number of points and the resulting precision and accuracy of AT results.

To compare standard and alternative BBA approaches, three AT strategies were tested, namely:

- direct georeferencing, using EO parameters derived from the refined trajectory (including boresight calibration);
- BBA, using field-surveyed points as GCPs;
- BBA, using LiDAR-derived points as GCPs.

In all tests, internal and additional parameters (estimated in simultaneous bundle adjustment, with sigma naught of 2 μm), were kept fixed as no significant improvements in AT results were observed if these parameters were refined in a self-calibrating BBA. It seems that the available camera calibration values are stable, and possible deficits are compensated by a good image overlap (80/60%) and a good initial estimation of camera EO, as also shown in Rupnik et al. (2015). Datum transformation parameters (translations in X, Y and Z) and IMU misalignments (to compensate for remaining errors in the boresight calibration) were instead estimated within the BBA. Both precision in image and object space, and accuracy against ground truth were evaluated and compared among the tests. Particularly, the latter was estimated by computing the root mean square errors (RMSE) on 44 check points (CPs), GNSS-measured with an accuracy of 5 cm. The selected distribution of the LiDAR-derived (yellow circles), field-surveyed (red circles) GCPs, as well as the position of CPs (blue triangles), are shown in Figure 7, left. Results are summarized in Table 4, which reports RMS and maximum values of (i) GCPs and CPs residuals and (ii) standard deviations of tie points.

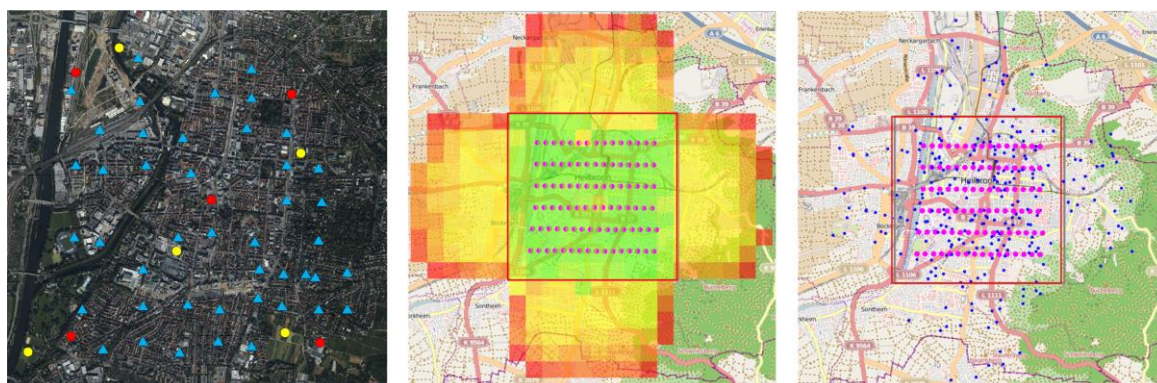|		Direct georeferencing		BBA with GNSS points		BBA with LiDAR points	
		RMS [m]	Max. [m]	RMS [m]	Max. [m]	RMS [m]	Max. [m]
GCPs (5)	X	-	-	0.03	0.05	0.12	0.19
	Y	-	-	0.03	0.06	0.10	0.17
	Z	-	-	0.03	0.05	0.10	0.15
CPs (44)	X	0.06	0.14	0.03	0.11	0.08	0.14
	Y	0.07	0.15	0.04	0.09	0.08	0.15
	Z	0.62	1.42	0.07	0.13	0.10	0.25
Tie points (8,201)	X	-	-	0.10	0.57	0.12	0.65
	Y	-	-	0.11	0.49	0.12	0.56
	Z	-	-	0.14	0.74	0.16	0.81

Table 4. Statistics on ground points.

Figure 7. Analyses of AT results. Left: distribution of control points. Centre: cell-based quality analysis of tie points a-posteriori standard deviations (green to yellow correspond to Std. dev. below 1.5xGSD; orange to red correspond to Std. dev. above 1.5xGSD). Right: distribution of tie points with image residuals above 0.5 pixel.

Direct georeferencing provides for horizontal accuracy values comparable to those achieved by BBA, demonstrating (i) the high quality of navigation data and calibration parameters and (ii) the potential use for quick mapping applications, e.g. in case of rapid mapping purposes (Toschi et al., 2017b). Nonetheless, if one has more accurate applications in mind, a complete adjustment of EO values and image observations is required to improve the resulting accuracy in height. Comparing the two BBA strategies of Table 4, although both RMS and maximum values of residuals increase when LiDAR-derived points are used, the differences are always below the GSD level, for all three types of ground points. This demonstrates the promising performance of the alternative AT strategy, as also confirmed by the graphical analysis shown in Figure 7 (centre): a cell-based analysis reports the (overall) a-posteriori standard deviations of 3D tie points with the highest number of image rays per cell, associated to a quality indicator. With a threshold of 1.5xGSD, this index is then represented by a colour scale, ranging from green to red (good to bad quality, respectively). Within the block extent (red rectangle), where the good ray intersection geometry of oblique images is fully exploited, a good precision of tie points is homogeneously achieved. Furthermore, the number and distribution of tie points with residuals in image space above 0.5 pixel, are shown in Figure 7, right. Comparing it with Figure 5, left, where all extracted tie points are shown, the good precision in image space (residuals below 0.5 pixel) achieved by most of the automatically matched points is evident.

4.3 Dense image matching vs. LiDAR point clouds

To evaluate the quality of DIM starting from the adjustment results delivered by the tested LiDAR-supported aerial triangulation, a SGM algorithm was then applied, adopting the original image resolution. The DIM 3D point cloud, after an automatic filtering, features a mean spatial resolution of 37 points/m^2, and includes information on building facades and footprints, as demonstrated by the close-up view in Figure 8, left. This represents the well-known advantage of dense image matching with oblique datasets (Haala and Rothermel, 2015), that exploits the slanted view of oblique images to overcome the problem of viewpoint restrictions typical of nadir-only block.
In order to take full advantage from the concurrent availability of ranging and image-based 3D information (i.e. LiDAR and DIM point clouds, respectively), it is necessary to investigate the quality of both data sources and their mutual fitting. In this perspective, a smaller AOI of 1 km x 1 km in the city centre of Heilbronn (red square, Figure 8, left) was selected for further investigations. These aimed at: (i) evaluating the precision of

the hybrid sensor system, intended as its capability to provide for mutually fitting 3D geometric information; (ii) assessing the vertical accuracy of both point clouds, w.r.t. ground truth; (iii) analysing the local noise level and completeness of both point clouds.
The vertical agreement between the LiDAR and DIM point clouds was estimated based on a very large number of planar patches of the same size, extracted on common surfaces in both datasets. First, in order to avoid those areas where laser signal can penetrate the vegetation layer (while DIM only returns the topmost surface), a point cloud classification of the two datasets was carried out, by adopting the supervised approach implemented in the Computational Geometry Algorithms Library (CGAL, 2018). After the identification and removal of vegetated areas, such as trees and gardens, the remaining point clouds were then analysed using the patch-based approach available in HxMap Lidar Survey Studio. The tool extracts square planar patches (2 m size) from two (or multiple) point clouds and compares the median Z values of points falling in each patch. To exclude small clusters and noisy segments, only patches with at least 5 points and standard deviations below 12 cm (1 GSD), were included in the comparison. A total of 67,032 square patches were finally generated: results are listed in Table 5, where cells colours correspond to the colour-coded visualization of patches vertical deviations in Figure 8, centre. Almost all patches show a good agreement between the two point clouds, with vertical shifts below 3 cm. The homogeneity level of patch deviations in the AOI demonstrates a good precision of the multi-sensor datasets and the absence of remarkable systematic deviations.
A similar patch-based approach was then applied to check the vertical accuracy of both datasets. As ground truth, 9 GNSS-measured CPs were adopted (blue triangles within the AOI, Figure 8, left). Around each check point, a circular patch (2 m diameter) was extracted, and the median Z of all point falling inside was compared to the reference one. Only patches with at least 5 points and standard deviations below 12 cm (1 GSD), were included in the comparison.

Vertical deviation [m]	# Patches	% of total
≤ 0.03	67,002	99.96
0.03-0.05	16	0.02
0.05-0.10	10	0.01
> 0.10	4	0.01

Table 5. Results of patch-based comparison between LiDAR and DIM point clouds showing their vertical deviations.

Figure 8. Analyses of DIM results. Left: close-up view of the DIM point cloud, with the AOI (red square) and profile section (yellow line). Centre and right: patch-based (colour-coded visualization as in Table 5) and profile comparison between the LiDAR and DIM point clouds.

Deviations computed for LiDAR data show a mean value of 0.03 m and a standard deviation of 0.02 m, whereas DIM point cloud features a mean patch deviation of 0.02 m and a standard deviation of offset of 0.09 m. These results show the comparable overall level of vertical accuracy of both datasets, although the higher reliability of LiDAR is highlighted, intended as the overall consistency of its measurements.

These remarks are deeply investigated by performing plane fitting tests, aimed at analysing the local noise level of the two point clouds. 20 planar patches were manually segmented on the ground level of both datasets, and best-fitting planes were then extracted. The RMS of the point-to-plane distances (after outliers removal) were finally computed, achieving a mean RMS of 0.10 m (DIM) and 0.03 m (LiDAR), standard deviation of RMS equal to 0.05 m (DIM) and 0.02 m (LiDAR). If the previous analyses, performed on the overall AOI level, showed a general agreement between the DIM median Z values of each patch and the corresponding reference (or LiDAR-derived) values, plane fitting tests show the higher noise level that locally affects the DIM data – as expected. Nevertheless, considering the GSD of the images (12 cm), the high quality of the DIM point cloud is still evident.

Finally, the completeness of the two point clouds is examined, with special focus on the reconstruction of building facades. First, 12 vertical patches were manually extracted on facades featuring different spatial orientations (N-S-W-O) and occlusion levels (large vs. narrow streets). As expected, the geometry of ray intersection in oblique image block supports the matching of points on building facades, when large spaces are available (mean number of points equal to 5,351 for DIM against 2,880 for LiDAR). On the contrary, in very narrow canyons surrounded by high buildings, occlusion issues and poor image contrast reduce the completeness of DIM point cloud (mean number of points equal to 1,677), whereas the laser beam occasionally reaches the facades (mean number of points equal to 2,573). The qualitative comparison between the vertical profiles extracted on the two point clouds in a narrow alley (Figure 8, right), further confirms this remark: points on facades are occasionally measured by LiDAR, whereas the amount of details reconstructed by DIM is limited.

5. INTEGRATED SOLUTION FOR MAPPING

The Bordeaux CityMapper dataset was exploited to demonstrate a combined use of oriented oblique images and LiDAR point cloud for mapping purposes. In this perspective, the block of images with their internal and external orientations (refined in HxMap) and the digital terrain model, generated by classifying

the LiDAR point cloud, were given as input to the GeoBLY tool. GeoBLY (GEOmetry extraction tool from aerial OBLique imagery – Toschi and Moe, 2017) is a software solution, developed in the framework of a joint research project between FBK (Italy) and AVT (Austria). The aim of the project was the development of an efficient and flexible solution for the management of large oblique aerial surveys and the extraction of accurate 3D information from them (Moe et al., 2016). Particularly, the tool is intended to meet the need of mapping agencies and other actors involved in mapping applications (e.g. cadastral building update, and other detailed restitution of objects). Indeed, GeoBLY allows the user to (i) visualize a building of interest and rotate around it at 360°; (ii) measure the length, height and area of any feature visible on the façade (i.e. windows, balconies, etc.) or on the roofs (i.e. chimneys, solar panels, etc.); (iii) monitor the status of buildings and infrastructures and identify structural damages. The availability of concurrently acquired oblique imagery and LiDAR point cloud, represents the optimal application field for the GeoBLY tool, that was therefore specifically adapted to manage the CityMapper datasets. Indeed, through various measurements methods, the integration of the two co-registered information sources is here exploited to extract precise vector information, like the digitization of building roofs and footprints (Figure 9).
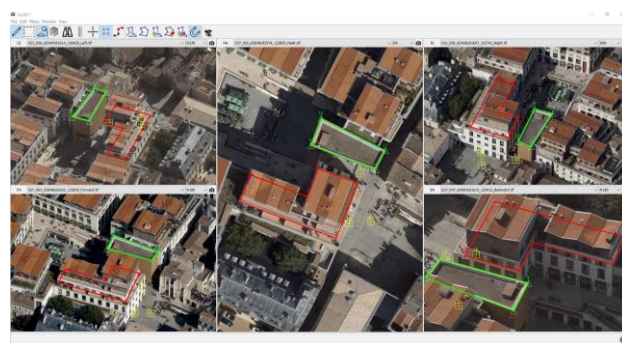

Figure 9. Vector information extraction from oriented oblique images and LiDAR-derived digital terrain model.

6. CONCLUSIONS

By analysing two CityMapper datasets, this contribution investigated the potentials and challenges of combining oblique camera and LiDAR sensors in the same airborne platform. The need for an integrated (automatic) processing of the concurrently acquired ranging and imaging data was evaluated, and first possible solutions were proposed. These include the use of LiDAR points as (i) height source, to guide the matching

between the different camera views; and (ii) control information in the image aerial triangulation, to reduce possible bias between LiDAR and DIM point clouds.

First tests and analyses performed w.r.t automatic point matching (APM), aerial triangulation (AT) and dense image matching (DIM), showed promising results, that encourage future improvements of the existing processing workflows towards the proposed (integrated) direction.

Finally, the combined use of oriented oblique images and LiDAR-derived digital terrain model was demonstrated to be an efficient means supporting the extraction of vector information for mapping purposes. This opens new opportunities for exploiting the real potential of both data sources and their full integration in a hybrid multi-sensor software solution.

ACKNOWLEDGEMENTS

REFERENCES

Awrangjeb, M. and Fraser, C.S., 2014. Automatic segmentation of raw LiDAR data for extraction of building roofs. In: *Remote Sensing*, Vol. 6(5), pp. 3716-3751.

CGAL, 2018. Computational Geometry Algorithms Library, http://www.cgal.org (Last accessed on July 2018).

Gerke, M., Nex, F., Remondino, F., Jacobsen, K., Kremerd, J., Karel, W., Huf, H., and Ostrowski, W., 2016. Orientation of oblique airborne image sets - Experiences from the ISPRS/EuroSDR benchmark on multi-platform photogrammetry. In: *ISPRS Int. Photogram., Remote Sens. & Spatial Inf. Sci.*, Vol. 41(B1), pp. 185-191.

Haala, N. and Rothermel, M., 2012. Dense multi-stereo matching for high quality digital elevation models. In: *Photogrammetrie-Fernerkundung-Geoinformation*, Vol. 2012 (4), pp.331-343.

Haala, N. and Rothermel, M., 2015. Image-based 3D data capture in urban scenarios. In: *Proc. Photogrammetric Week 2015*, D. Fritsch (Ed.), pp. 119-130.

HxMap, 2018. www.leica-geosystems.com (Last accessed on July 2018).

Leberl, F., Irschara, A., Pock, T., Meixner, P., Gruber, M., Scholz, S. and Wiechert, A., 2010. Point clouds: Lidar versus 3D vision. In: *Photogrammetric Engineering & Remote Sensing (PERS)*, Vol. 76(10), pp. 1123-1134.

Leutenegger, S., Chli, M. and Siegwart, R.Y., 2011. BRISK: Binary robust invariant scalable keypoints. In: *Proc. Int. Conf. Computer Vision 2011*, pp. 2548-2555.

Maltezos, E., Kyrkou, A., and Ioannidis, C., 2016. LiDAR vs dense image matching point clouds in complex urban scenes. In: *Proc. SPIE 9688*, pp. 1–10.

Mandlburger, G., Wenzel, K., Spitzer, A., Haala, N., Glira, P. and Pfeifer, N., 2017. Improved topographic models via concurrent airborne LiDAR and dense image matching. In: *ISPRS Ann. Photogram., Remote Sens. & Spatial Inf. Sci.*, Vol. 4, pp.259-266.

Moe, K., Toschi, I., Poli, D., Lago, F., Schreiner, C., Legat, K. and Remondino, F., 2016. Changing the Production Pipeline - Use of Oblique Aerial Cameras for Mapping Purposes. In: *ISPRS Int. Archives of Photogram., Remote Sens. & Spatial Inf. Sci.*, Vol. XLI-B4, pp. 631-637.

Nex, F., Gerke, M., Remondino, F., Przybilla, H.J., Bäumker, M., and Zurhorst, A., 2015. ISPRS benchmark for multi-platform photogrammetry. In: *ISPRS Ann. Photogram., Remote Sens. & Spatial Inf. Sci.*, Vol. 2 (3), pp. 135–142.

Parmehr, E.G., Fraser, C.S., Zhang, C. and Leach, J., 2014. Automatic registration of optical imagery with 3D LiDAR data using statistical similarity. In: *ISPRS J. of Photogrammetry and Remote Sensing*, Vol. 88, pp. 28-40.

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

Remondino, F. and Gerke, M., 2015. Oblique Aerial Imagery – A Review. In: *Proc. Photogrammetric Week 2015*, D. Fritsch (Ed.), pp. 75-83.

Remondino, F., Toschi, I., Gerke, M., Nex, F., Holland, D., McGill, A., Talaya Lopez, J. and Magarinos, A., 2016. Oblique aerial imagery for NMA - Some best practices. In: *ISPRS Int. Archives of Photogram., Remote Sens. & Spatial Inf. Sci.*, Vol. XLI-B4, pp. 639-645.

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. In: *Photogrammetrie-Fernerkundung-Geoinformation*, Vol. 2016 (2), pp.57-73.

Rönnholm, R., Haggrén, H., 2012. Registration of laser scanning point clouds and aerial images using either artificial or natural tie features. In: *ISPRS Ann. Photogram., Remote Sens. & Spatial Inf. Sci.*, Vol. I-3, pp. 63–68.

Rupnik, E., Nex, F., Toschi, I. and Remondino, F., 2015. Aerial multi-camera systems: accuracy and block triangulation issues. In: *ISPRS J. of Photogrammetry and Remote Sensing*, Vol. 101, pp. 233-246.

Shan, J. and Toth, C.K., 2010. *Topographic laser ranging and scanning: principles and processing*. CRC press.

Tian, J., Schneider, T., Straub, C., Kugler, F. and Reinartz, P., 2017. Exploring digital surface models from nine different sensors for forest monitoring and change detection. In: *Remote Sensing*, Vol. 9(3), pp. 287-312.

Toschi, I., Ramos, M.M., Nocerino, E., Menna, F., Remondino, F., Moe, K., Poli, D., Legat, K. and Fassi, F., 2017a. Oblique photogrammetry supporting 3D urban reconstruction of complex scenarios. In: *ISPRS Int. Archives Photogram., Remote & Sens. Spatial Inf. Sci.*, Vol. 42, pp. 519-526.

Toschi, I., Remondino, F., Kellenberger, T., and Streilein, A., 2017b. A survey of geomatics solutions for the rapid mapping of natural hazards. In: *Photogrammetric Engineering & Remote Sensing (PERS)*, Vol. 83(12), pp. 843-860.

Toschi, I. and Moe, K., 2017. GeoBLY – a tool for accessing oblique datasets. EuroSDR/ISPRS Workshop Oblique 2017 (presentation available at http://3dom.fbk.eu/files/fbk_avt-geobly.pdf)

Vosselman, G. and Maas, H.G., 2010. *Airborne and terrestrial laser scanning*. CRC Press.

Zhang, J. and Lin, X., 2017. Advances in fusion of optical imagery and LiDAR point cloud applied to photogrammetry and remote sensing. In: *Int. J. Image and Data Fusion*, Vol. 8(1), pp. 1-31.

Zhang, Z., Gerke, M., Vosselman, G. and Yang, M.Y., 2018. A patch-based method for the evaluation of dense image matching quality. In: *Int. J. Applied Earth Observation and Geoinformation*, Vol. 70, pp. 25-34.