



Lidar and Photo: differences and integrated processing

Norbert Pfeifer, Wien

Abstract

The differences between Lidar and photo observations are discussed on a sensor level. This highlights the similar and complimentary aspects of both data acquisition methods. A method for the integrated orientation of photo and Lidar observations is presented and its effectiveness is shown. It is argued that integrated acquisition and processing will become a standard for topographic data acquisition. The article is based on the research and experience of the photogrammetry group at Technische Universität Wien.

Keywords: Lidar, Camera, Laser Scanning, Photogrammetry, Complementarity

Kurzfassung

Für die Erfassung topographischer Information über größere Bereiche stehen praktisch zwei Messkonzepte zur Verfügung: Lidar (Light Detection And Ranging), auch unter dem Namen Laserscanning bekannt, misst direkt 3D und ist die jüngere Technologie und die 3D-Rekonstruktion aus Photographien, die auf bereits 150 Jahre Erfahrung zurückgreift. Beide Technologien entwickeln sich rasch weiter. Anhand der Gemeinsamkeiten und der Unterschiede der beiden Messkonzepte, untersucht auf dem Sensor-Niveau, wird gezeigt, wie sehr sich diese beiden Methoden ergänzen. Eine gemeinsame Prozessierung kann potentiell genauere, zuverlässigere und vollständigere Modelle unserer Umgebung liefern, die noch dazu effizienter erstellt werden können. Eine solche integrierte Verarbeitung ist aber nur für wenige Aufgaben entlang der Prozessierungskette von der Datenaufnahme bis zum 3D-Modell realisiert. Ein Ansatz zur gemeinsamen Orientierung wurde bereits vorgeschlagen und praktisch eingesetzt. Dieser Artikel soll die Komplementarität der beiden Sensoren stärker herausarbeiten und dazu beitragen die integrierte Aufnahme und Prozessierung von Lidar- und Photo-Aufnahmen als Standard etablieren.

Schlüsselwörter: Laserscanning, Kamera, Lidar, Photogrammetrie, Komplementarität

1. Introduction

Technische Universität Wien (TU Wien) has 8 faculties, among it the faculty of „Mathematics and Geoinformation“. This faculty has four institutes, three in mathematics and the “Department of Geodesy and Geoinformation”. With more than one hundred employees it is, also considering international standards, a large group in this domain, reaching from engineering and advanced geodesy, via remote sensing, geophysics and photogrammetry to geoinformation and cartography.

The photogrammetry research unit has built up a reputation in airborne laser scanning research, including full waveform analysis (Wagner et al., 2006; Schwarz et al., 2019), sensor orientation (Kager et al., 2004; Glira et al., 2016), and terrain modelling (Kraus et al., 1998) as well as vegetation modeling (Hollaus et al., 2006) and hydrography (Mandlbürger et al., 2015). The scope of research is, however, wider and reaches from the sensor to the application. The photogrammetric aspect in this research is that i) the method of data coll-

ection is concentrated in the optical part of the electromagnetic spectrum (using Photons), thus the visible and near infrared light. It means further that geometrical optics (line of sight, lenses) is in most cases sufficient to explain the imaging process. A further photogrammetric aspect in the research is, that ii) the sensors are imaging, which means that the entire scene is recorded in a profile- or area-wise image (referring to graphic recording), as it is typically acquired by a camera or a laser scanner. Thus no interpretation of the scene is performed during data acquisition, e. g. by measuring “only” specific object points like corners. Finally, the aim is iii) a metric exploitation of those measurements, in order to build georeferenced models of our environment from those images. Applications play an important role in the research work of the group. This steers the basic research into directions relevant for society, public administration, and economy.

The most important sensors in topographic photogrammetry are cameras and pulsed laser scanners. In this article the research work of the

photogrammetry group w.r.t. the mutual differences of those two data acquisitions methods and the current state of joint processing are presented¹⁾. Such a combined exploitation of photographic images and Lidar (light detection and ranging) sensors is not an aim per se, but has to support more efficient, reliable, complete, and accurate extraction of 3D information.

The aim of this article is to study the differences between photo and Lidar observations on a fundamental, sensor-oriented level and present one domain, i. e. orientation, in which a successful integrated processing has been established. This builds the basis to formulate expectations for future developments.

2. Photos and Lidar

Photos are taken by cameras and have been studied in photogrammetry since roughly 150 years (Albertz, 2007). They operate by recording reflected sunlight, which is focused by the lens system onto a matrix arrangement of light sensitive elements (pixels) in the focal plane. Scanning Lidar is comparatively newer and is applied since roughly 30 years in photogrammetry (Kilian et al., 1996). Lidar operates by emitting a short laser pulse in a specific direction, which then travels through the atmosphere, is scattered back at objects, and the (small fraction of the) signal that travels back to the sensor is detected to record the time lapse between emission and detection. With the known speed of light this is converted

1) As this article presents the research work from one group, the list of references is somewhat imbalanced.

to a range measurement. Both sensors, cameras and Lidars, are typically used to acquire area-wide data from either static or mobile (e.g. airborne) platforms. Lidar and photos, through dense image matching (e.g. Hirschmüller et al., 2008), can both be used to provide a point cloud describing the recorded object surfaces.

2.1 Comparison

Considering that both sensors use either visible light or the near infrared, also different objects (building, streets, terrain, etc.) are measured in a similar way. The density of measurements scales with flying height above ground for both sensors, and also does the swath width. However, there is also a list of differences, which originates in the physical measurement principles and has effect on measured objects. Understanding those differences and exploiting them is part of the fundamental research in the photogrammetry group of TU Wien. It has impact on applications, e.g. when mapping vegetation, water bodies or urban regions, but also influences the possible deployment of these sensors. Practical differences between Lidar and photo point clouds are presented, e.g., in Mandlbürger et al. (2017), Ressler et al. (2016) and Otepka et al. (2013), see also Figure 1.

Photos record simultaneously the entire field of view, which depends on the focal length and the opening angle, respectively, of the lens system. The simultaneous recording also means, that the incoming radiation is recorded for all pixels simultaneously, which is a large advantage when it comes to georeferencing. For modern photogram-

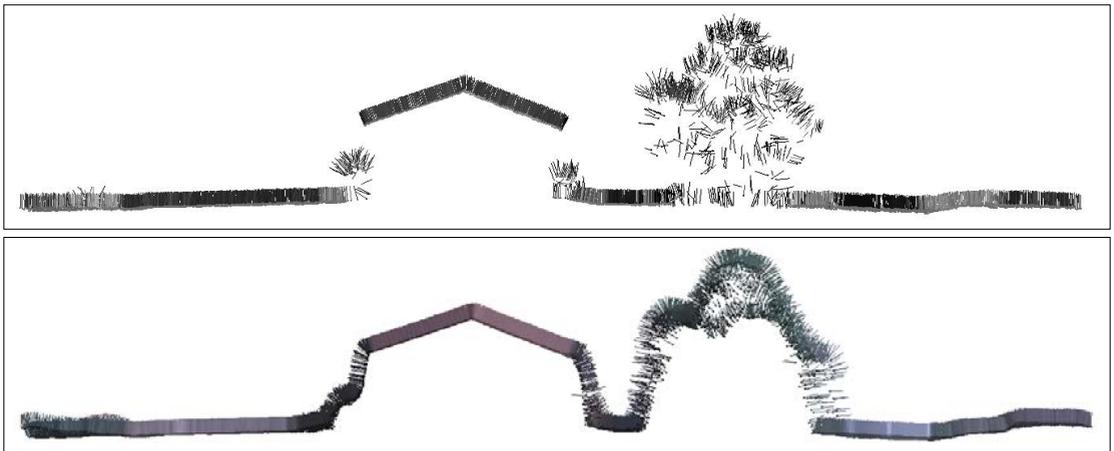


Fig. 1: Lidar (top) and photo (bottom) point clouds showing a profile with open ground, a house and tall vegetation. Additionally to the points, estimated normal vectors are shown. The Lidar point cloud shows as grey value the amplitude of the detected signal, whereas the photo point cloud has the color of the corresponding images. (Figure taken from Otepka et al., 2013)

metric cameras this means that 100 million pixels of one image – or for the newest cameras rather a few 100 million pixels – have the same exterior orientation. Acquiring overlapping photos leads to overdetermination, and by bundle block adjustment the exterior orientation can be estimated using only a few ground control points. There are some limits, e. g. motion blur, which can partly be compensated, or special mechanisms, like, e. g., the rolling shutters. The latter are not advisable for moving platforms or capturing non-static scenes. Professional aerial photogrammetric cameras have a fixed focal length (and a fixed focus), which means that their field of view is constant.

Lidar, on the other hand, is a **sequential measurement** technology. The measurements are performed one after another. For mobile platforms this means, that each measurement has its own exterior orientation, and thus direct georeferencing using, e. g., GNSS and INS (Global Navigation Satellite System and Inertial Navigation System) are mandatory. In earlier commercial systems only one pulse was in the air traveling from sensor to object and back. This limited flying height or the pulse repetition frequency and the effective scan rate, respectively. Newer systems are not limited in that sense and can have multiple (e. g., 5 and more) pulses simultaneously in the air. Additionally, multiple laser range finders can be mounted in one scanner, using the same or different beam deflection devices. In this way the scan rate is further increased. Additionally this starts (at a very low level) introducing within-strip overdetermination, because the same location is measured more than once. The palmer scanners (using a nutating mirror or a rotating wedge prism for beam deflection), effectively producing a circular scanning pattern on the ground, naturally provides this. Other methods to increase the number of beams operate by splitting each emitted beam into beamlets, using, e. g., diffractive optical elements (DOE). All these beamlets are released simultaneously, therefore having the same exterior orientation. Such DOEs are planned, e. g., for NASA's Lidar swath imaging space mission LIST (Yu et al., 2010), but is currently also employed in topographic scanning (Degnan et al., 2016). Those methods increase the number of observations with one exterior orientation.

A physical limit to the **resolution** is given by diffraction. For photos this means, that an ideal point is imaged onto a circle, which is in the order of the unit-less aperture number interpreted in μm .

Thus, pixel size is accordingly in the order of 5 μm or somewhat smaller. Opening angles for individual pixels can, in tele configuration, be below 0.02 mrad. For wide angle cameras these values naturally go up. The ground sampling distance (GSD) is obtained by the multiplication of the pixel size with the image scale. Typically the entire sensor area is sensitive to light, e. g. with the help of micro lenses that focus the entire light of one pixel in the focal plane onto a smaller, photo sensitive region. Thus, the ground is covered contiguously with pixels, which are spaced at the GSD.

For Lidar the beam width (or beam divergence) is typically higher, in the order of 0.2 mrad. It is in the order of λ/D , with λ the wavelength and D the aperture diameter of the emitter. For cameras the limiting aperture is given by the lens system. The area illuminated with one laser shot of an airborne laser scanner, the footprint, is thus typically larger than the GSD of the panchromatic image acquired by a professional photogrammetric camera flown at the same height (see also Figure 5). A smaller opening angle of the laser beam would increase the resolution, but focus more energy into a smaller region, which can become problematic w.r.t. eye safety. Also, repetition frequency, flying speed, and footprint size should be chosen to map the entire ground (contiguously) in airborne operation. In terrestrial laser scanner systems correlated sampling (i. e. overlapping footprints) is possible. In that case the "GSD" (i. e. the linear point spacing) is smaller than the footprint (see, e. g., Milenković et al., 2018). The footprint can also be tailored to applications, e. g., forming an elliptic footprint with its longer axis orthogonal to the flying direction, in order to increase the chance of mapping linear elements in flying directions (e. g. power lines).

Color can be recorded in photos using multiple (lens) cones, and therefore multiple cameras, thus providing multiple photos in different parts of the electromagnetic spectrum, e. g., blue, green, red, and near-infrared. These cameras do typically have lower resolutions than their pan-chromatic counter parts, and pan-sharpening is used to fuse the higher resolution pan-images to the lower resolution color images. An alternative, found more often in consumer cameras and in professional "mid-format" cameras used for photogrammetry, are color filter arrays as the Bayer pattern. Each light sensitive pixel in the focal plan is covered with either a red, a green, or a blue filter. To obtain the "full" resolution, the colors are interpolated

from the respective recordings to obtain red, green, and blue values for each pixel. It is noted that both approaches lead to a lower resolution of the color information, often by factors between 2 and 4 (linear), in comparison to the pan-chromatic recordings. The grey or color values are often provided as digital numbers, i. e. not in a physical unit like radiance at the sensor. If multiple objects are within the instantaneous field of view of one pixel, they all contribute in an integral way to the sensor reading.

In contrast, a Lidar measures in a very narrow spectral band (around the wavelength of the Laser), opposed to the (broad) color bands of a camera. Assuming that only one target is within the instantaneous field of view of the laser beam, the amount of diffuse reflection towards the sensor, and the loss due to spreading of the reflected signal into the entire (half-)space, influences the amount of recorded energy. As the amount of emitted energy can be or is often known for laser scanners, also the amount of received energy can be expressed in physical units. Because of the active system and the very narrow spectral band of the detector, radiometric calibration is straightforward for Lidar observations in comparison to photos, for which the solar illumination and the composition of the atmosphere play a more important role. Currently, the number of laser scanners operating at multiple wavelengths is increasing. While commercially available ALS sensors feature two or three wavelengths, experimental terrestrial systems have eight and more channels (Hakala et al, 2012).

Photographic imaging is singular, which means that the process of focusing the incoming radiation onto the focal plane reduces the 3D object space to a 2D image. This mapping cannot be inverted, and thus – at least – one additional photo

is required to reconstruct a scene in 3D from photos. Matching of homologous points is necessary, which requires computing resources. In a practical investigation Tran et al. (2018) have shown, that computation time grows with decreasing resolution, specifically it grows with current standard approaches stronger than linear with the number of points that can be reconstructed. Furthermore, the matching process, although featuring over-determination, can also lead to matching of non-homologous points, and wrong points in object space. Additionally, matching requires that the same texture is recorded (and recognized) from different exposure positions, which are furthermore illuminated by the sun. Thus at least three rays have to hit a point, two directly from the camera projection centers and one from the sun, at least indirectly. Therefore, reconstruction of the ground below forest cover is in general not possible from photos (see also Figure 1). This demand for at least two viewpoints in order to reconstruct a 3D point also means that narrow alleys are more difficult to reconstruct, because a correspondingly dense strip layout in the flight plan is necessary to reconstruct especially those alleys that are parallel to the flying direction. Finally, this singularity also means that multiple objects along the instantaneous field of view of one pixel are recorded at the same position.

Laser scanners use a polar measurement technique, recording range and direction to the measured targets simultaneously. Obviously, this brings an advantage in computation time. The measurement to a single object point is, however, not controlled. This becomes obvious in multi-path situations, caused by specular reflecting surfaces (e.g. glass façades), which lead to an additional mirrored image of the mapped objects (see Figure 2).

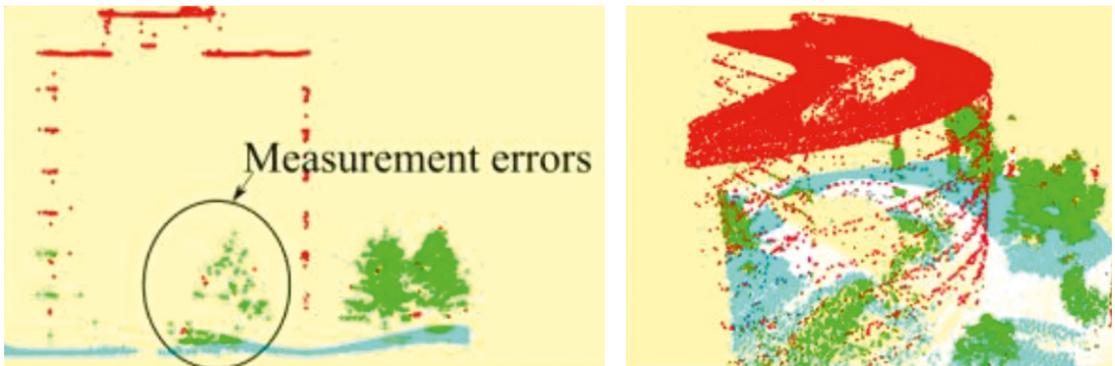


Fig. 2: Due to multi-path, the polar measurement method generates a mirrored image of a real object. The building glass façade is acting as mirror. Class labels (building, ground, vegetation) were derived automatically (Tran et al., 2017).

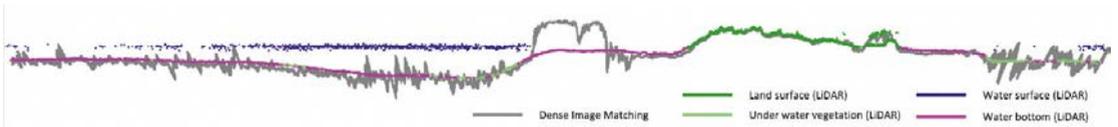


Fig. 3: A comparison of Lidar measurements and a point cloud reconstructed from photos. The left part of the profile is acquired over water surface, and the Lidar sensor provides both, water surface and water bottom. As the water is clear, dense image matching through water can be performed, also providing an estimate of the bottom surface, but not the water top surface. It is noted that the noise of the image matching point cloud is larger than of the bathymetric Lidar point cloud. Over open land, both datasets show the same surface. (Figure taken from Mandlburger, 2018)

The ranging capability of Lidar means that range-resolved measurements are possible. Under the assumption that a short signal of laser light (often <10ns) is used to scan (and therefore illuminate) the object surface, targets that are further apart (along the beam direction) can be discriminated. Each target causes an individual echo. This **multi-target capability** is especially advantageous over tall vegetation, because it allows measuring the crown (first echo), vegetation elements below (intermediate echoes), and the ground (in an ideal but not untypical situation the last echo, see also ground points below vegetation in Figure 1). The polar measurement technique is also advantageous in urban canyons or under semi-transparent objects (tree canopy), as only one beam has to reach the lower surfaces to record a position. A further example, in which this is exploited, is the bathymetric measurement with Lidar. Using a suitable wavelength (green light, e. g. 532nm)

allows to record a range measurement to the water surface, but also a further measurement to the bottom of the water body (see Figure 3). Also laser scanning can deliver wrong points in object space. They originate from multiple targets along the laser beam, which are spaced so close together, that the returning signals overlap strongly and only one point in a middle (and therefore wrong) distance is recorded.

Lidar is an active measurement method. It records the backscatter from a signal that was emitted by the very same system. This makes it independent from illumination conditions. The atmosphere must, however, be clear for both, Lidar and photos. This makes data acquisition during night time possible. **Photos** are passive sensors, recording (at least outdoors) the backscattered radiation of the sun. This means, that they are affected by shadows, which leads to a lower precision in matching of points located in shadow areas

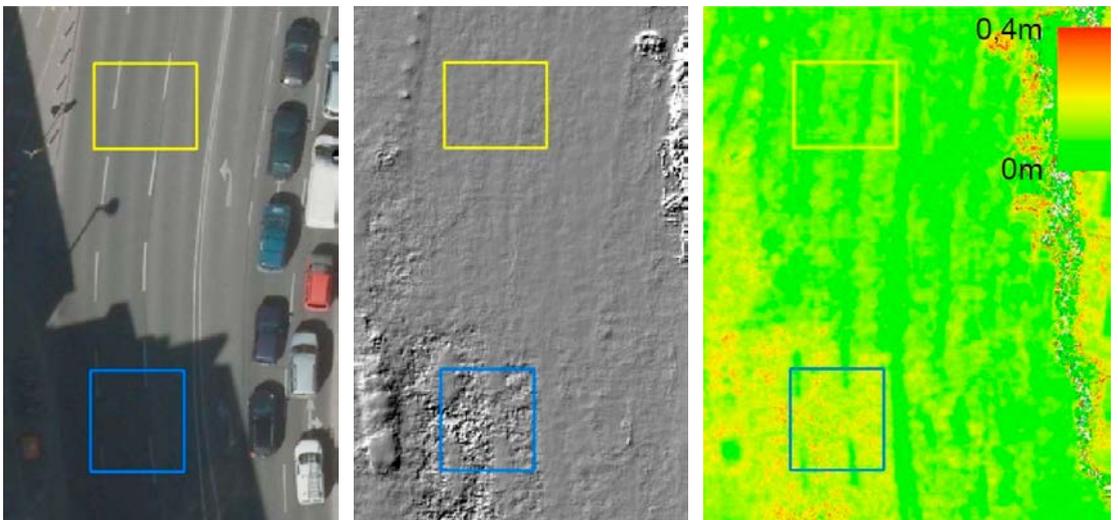


Fig. 4: Shadows in the original image (here orthophoto for comparison), DSM derived robustly from dense image matching, and std.dev. of heights originating from multiple image pairs. The higher std.dev. in the shadow areas is highlighted. Data provided by Stadt Wien, MA41.

(see Figure 4). Recording reflected sun light has the advantage that no power source is needed as it is required in Lidar to generate Laser pulses. This leads to a lower weight for the sensor, which is notable for terrestrial, static systems, and for operation on unmanned, light weight platforms.

Pulsed Lidars emit a short beam. This is in contrast to the sun continuously illuminating a target, which reflects a constant stream of light to the sensor (i. e. the camera). The Lidar signal scattered back to the sensor, on the other hand, is the convolution of the emitted pulse shape (more correctly the system waveform) with the differential target backscatter cross section. This differential cross sections contains information on the location of reflection along the beam as well as the brightness at the wavelength of the laser. There are different possibilities to **detect and process this returning signal**. Recording of the full waveform provides information on the elongation of objects along the laser beam. While strongly inclined surfaces stretch the returning signal. However, it is especially the occurrence of distributed targets with similar but not identical range along the beam axis, which causes a widening of the returning pulse. The echo-widening, typically high over tall leaf-on vegetation for footprints in the order of some dm, can thus be used to detect vegetation. This also applies to low vegetation, which is otherwise hard to discriminate from ground echoes. Recording and analyzing this full-waveform has led to an increased number of (discrete) echoes in vegetation. For denser media, e. g. water, an exponential decay of the returning energy can be expected. This technology allows ranging over

long distances (km) with precision in the order of cm. It requires, however, that the returning signals are sufficiently strong. Detection of weaker signals becomes possible using so-called single photon detection (Degnan et al, 2016). In this case, one photon (or a very small number of photons) is sufficient to detect a returning pulse. This may cause false detections, and a relatively high number of erroneous points, which have to be eliminated in further post-processing. An example of single photon Lidar, an aerial photo, and full waveform Lidar is shown in Figure 5.

While some of the differences between photos and Lidar elaborated above are governed by physical principles, the effects on the recorded measurements are depending on the current state-of-the-art in the sensor technologies. For both, laser scanners and cameras, the last years have provided an ever increasing density of measurements and thus higher productivity. The advantage of cameras lies mainly in the smaller (planimetric) GSD, whereas Lidar provides multiple echoes and higher precision in range, which is basically the vertical direction for airborne acquisition. Some developments in Lidar technology, e. g. bathymetric measurements using a green Lidar, are already established, while multi-wavelength Lidar has not yet found a wide area of application in surveying.

Both technologies can be expected to develop further, and they are complimentary, as shown above. It thus stands to reason to ask, how can the datasets be used together in the best possible way?

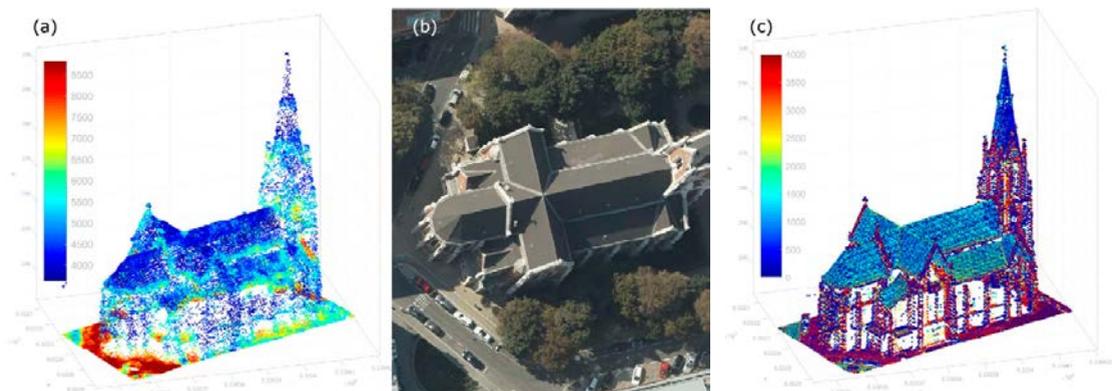


Fig. 5: Left the point cloud from single photon Lidar is shown, in the middle an aerial image, and right the point cloud from full-waveform Lidar. All sensors were flown at the same height. (Figure taken from Mandlbürger et al., 2019)

2.2 Integrated orientation

Fusing Lidar and photo observations can occur on different levels, either in the orientation phase, in the 3D modeling phase, or in the application phase. Obviously, fusing data from both sensors in earlier stages of the workflow will provide higher accuracy and reliability, as the entire information is available in all subsequent stages. The disadvantage is that current workflows or algorithms have to be adapted to accommodate both sources.

The integration of Lidar and photo observations is suggested by integrating strip adjustment and bundle block adjustment. The aim of strip adjustment is to determine improved trajectories (improved w.r.t. the direct georeferencing solution) and to estimate calibration parameters (e.g. misalignment between laser scanner and INS). Subsequently, new 3D object point coordinates are computed. The optimization principle minimizes the distances between overlapping laser strips and the differences to ground control data (points or point clouds). Bundle block adjustment operates in a similar way, estimating exterior orientation, camera calibration parameters, and – in contrast – also the tie point coordinates. The optimization minimizes the offsets between the rays of corresponding points, as well as the offset between ground control points and corresponding ray(s).

The equation relating the unknown exterior orientation, the unknown tie point, and the measured image point to each other is the collinearity.

$$(\mathbf{X}_i - \mathbf{O}_j) = mR_j(\mathbf{x}_{ij} - \mathbf{o}) \quad (1)$$

Here the terms have the following meaning: \mathbf{x}_{ij} are the observed point coordinates in the image plane, and \mathbf{X}_i is the object point; \mathbf{o} is the interior orientation, possibly augmented by the distortion; \mathbf{O}_j and R_j are the exterior orientation of image j ; m is image scale, individual for each point measurement and eliminated by dividing the first two rows of this vector valued equation by the third row. The index i is specific for each point, the index j for each photo.

The direct georeferencing equation of airborne lidar is:

$$\mathbf{X}(t) = \mathbf{G}(t) + R^n(t)R^b(t)(\mathbf{a}^m + R^m \mathbf{x}(t)) \quad (2)$$

Also, this equation relates the measurement of a point in the laser scanner coordinate system $\mathbf{x}(t)$, thus a range and two angle measurements, to the object point $\mathbf{X}(t)$. Here the index pair (i, j) is

replaced by the time of the measurement (t) . The other parameters are: the mounting, consisting of the bore sight (mis)alignment R^m and the lever arm \mathbf{a}^m , which rotate and shift the vector $\mathbf{x}(t)$ into the system of the INS; $R^b(t)$ describes the rotation from this body system to the navigation frame and is provided by the Kalman filter output of the GNSS and INS observations, typically designated as roll, pitch and yaw angle; $R^n(t)$ rotates from this frame, the local horizon or navigation frame, to an earth centered earth fixed system (e.g. WGS84), which depends on the current latitude and longitude; $\mathbf{G}(t)$ finally is the vector denoting the position of the GNSS antenna.

Bringing those equations, photo collinearity and Lidar direct georeferencing, together, requires interchange between the paradigm of observing key points in multiple images \mathbf{x}_{ij} and a continuous stream of point measurements $\mathbf{x}(t)$. The concepts of exterior orientation are comparable: \mathbf{O}_j and $\mathbf{G}(t)$ for the location and R_j and $R^n(t)R^b(t)$ for the angular attitude. Because of direct georeferencing, the mounting parameters \mathbf{a}^m and R^m are necessary, but this would equally apply if integrated georeferencing (i. e. ground control points, tie points, and direct georeferencing) is performed for photos. Similar to the case of photos, calibration parameters may be added to the Lidar observations $\mathbf{x}(t)$, as functions of the observed range and angles.

In the bundle block adjustment, the residuals ν_{ij} are added to each point, and their square sum is minimized in order to determine the unknown parameters. The homologous points do not exist in laser scanning. The solution suggestion by Kager (2004) uses homologous planar patches, which have three unknowns, similar to the tie points of the bundle block. The alternative solution suggested by Glira et al. (2016) replaces the exact correspondence between points by approximate correspondence, as applied in the ICP algorithm (Besl et al., 1992). For relative orientation of 3D point clouds, the formula is:

$$\sum_{i=1}^n \left[\left(T(\mathbf{Y}_i) - \mathbf{X}_{c(\mathbf{Y}_i)} \right)^T \mathbf{n}_{c(\mathbf{Y}_i)} \right]^2 \xrightarrow{T} Min \quad (3)$$

Here, \mathbf{X}_j is an individual point of a fixed point cloud, and \mathbf{n}_j is its normal vector. \mathbf{Y}_i is a point of the second point cloud with n points, which is to be transformed, e.g. by a Euclidean transformation, $T(\mathbf{Y}_i) = \mathbf{Y}_0 + R\mathbf{Y}_i$, in order to fit as good as possible to the first (i. e. the fixed) point cloud. The function $c(\mathbf{Y}_i)$ delivers the index of the points \mathbf{X}_j

which is spatially closest to Y_i . Alternatingly the transformation parameters are determined and the correspondences for the transformed points $T(Y_i)$ are computed. Each transformation brings the points Y_i closer to X_j until an optimum is reached. The optimum is that the square sum of the orthogonal distances d_i (the term in the square brackets above) is minimal.

The principle of the joint strip and bundle block adjustment is to simultaneously minimize the distances between overlapping strips, the residuals of the tie point observations, and the distances of the tie points to the strips. Furthermore, deviations between photo/Lidar measurements to the control points or control point clouds, respectively, are minimized. The residuals of the tie points are formulated in image space, whereas the strip-to-strip differences with their observed value zero ($d(t, t')$ in Figure 6), are formulated according to the ICP principle in object space. The transformation T is not the Euclidean transformation, but rather the direct georeferencing equation of Lidar. Its unknowns are the mounting parameters and corrections for the rotation and the GNSS antenna phase center. Likewise, the tie points from the bundle block are formulated to have a distance of zero to the Lidar strips (d_i in Figure 6).

More complex models linking both, photos and Lidar, to a common trajectory, and improving a given trajectory from direct georeferencing with

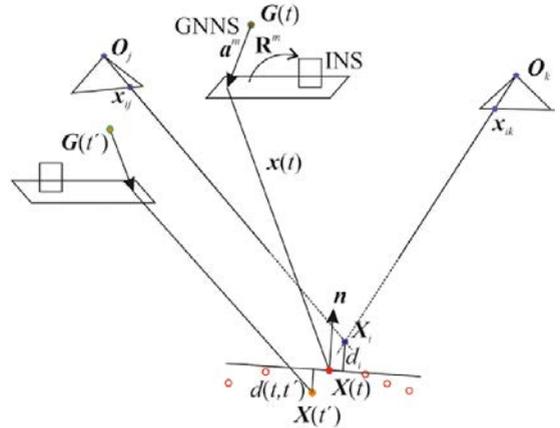


Fig. 6: Integrated orientation of Lidar and photo observations. The normal vector n in $X(t)$, solid red point, is estimated from its neighbors measured within the same Lidar strip (red circles). The Lidar point $X(t')$ is measured typically in another strip. The point X_i is measured in at least two, but preferably more, photos. Airborne scanning Lidar requires direct georeferencing (GNSS, INS) and the corresponding mounting parameters.

time dependent correction functions (e. g. splines) are presented in Glira et al. (2019).

An integrated orientation of Lidar and photo observations triggers the question for the homologous elements. As shown in Figure 7, only “hard” surfaces should be used, whereas surfaces covered with grass feature already a height differences between Lidar and photo point cloud.

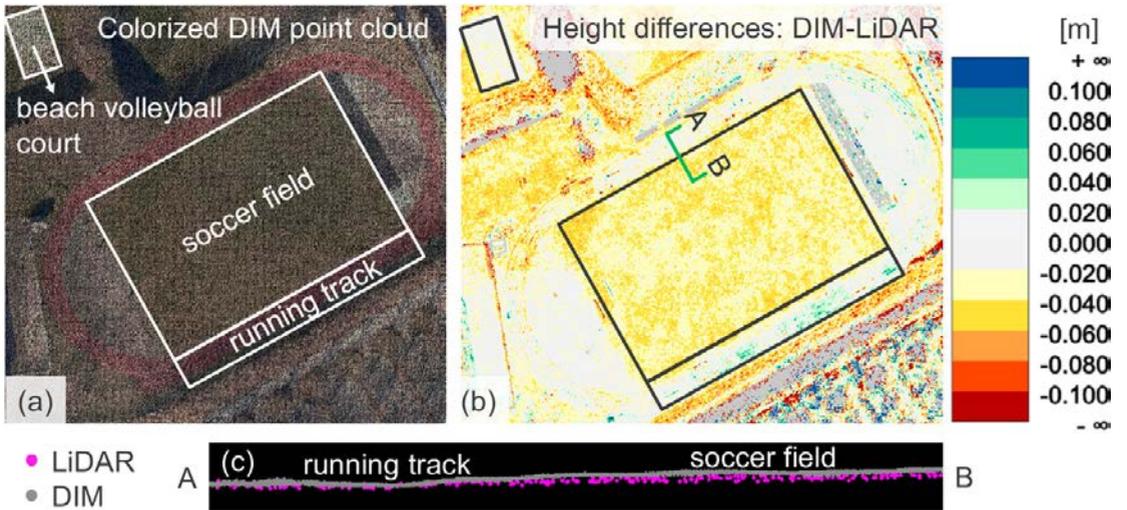


Fig. 7: Difference between points clouds from dense image matching and from Lidar. Data was acquired from the same platform. Left the situation is shown, right the height difference, which is zero up to the single centimeter for solid surfaces (running track) and above 2cm for the grass surface (soccer field). (Figure taken from Mandlbürger et al., 2017)

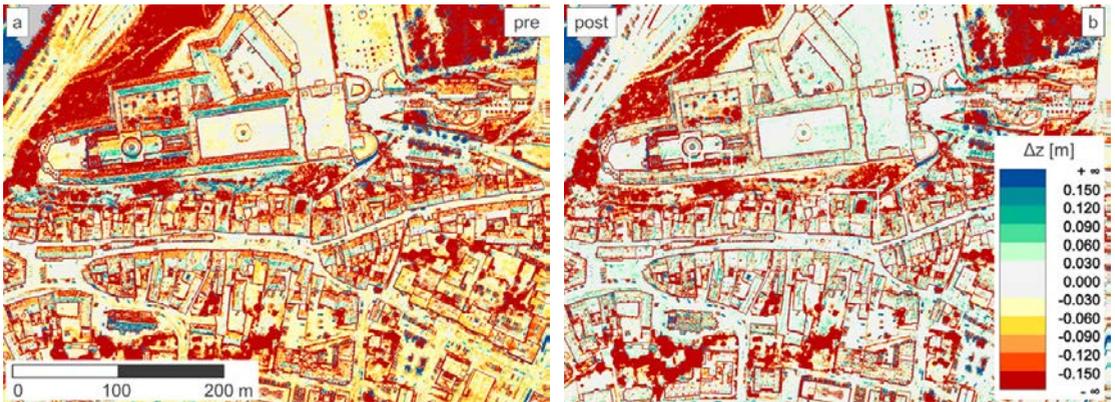


Fig. 8: Differences between dense image matching and Lidar point cloud before (left) and after (right) integrated orientation. For solid surfaces the improvement from above 6 cm to below 3 cm can be seen. For inclined surfaces (roofs) the improvement is bigger, because the alignment is also improved in the horizontal component. (Figure taken from Mandlbürger et al., 2017)

The benefit of integrated processing is demonstrated with one example. Using photos and Lidar data acquired from one platform, at first independent orientation was performed for either dataset: the bundle block adjustment with ground control points for the photos and strip adjustment with ground control patches for the Lidar data. Both datasets exploited direct georeferencing, but for the photos it is “only” an observation stabilizing the bundle, whereas for the Lidar data it is indispensable. Subsequently, a dense image matching was performed to obtain a dense point cloud. The differences between those point clouds are expected to be high over tall vegetation or other rough surfaces, but zero at solid surfaces (street roofs, etc.). Within each single sensor orientation result, no inaccuracy could be detected. However, as shown in Figure 8, over street surfaces larger differences appear. This indicates that at least one source features internal, undetected errors. A joint orientation successfully removes those biases.

2.3 Integrated Point Cloud Processing

While the integrated orientation and calibration of measurements from Lidar and photos is already suggested and operational, there is comparatively little work so far on how to optimally use the point clouds for deriving 3D models or classifying the point cloud.

- Mandlbürger et al. (2017) have suggested to derive better surface models by using the higher reliability of Lidar point clouds and the higher density of photo point clouds.

- The process of ortho photo generation can be speeded up with a concurrently acquired Lidar data. The surface model comes from Lidar and the image content from photos. If the data are acquired simultaneously, changes in the objects are minimized (excluding, e.g., the effect of growth or wind on vegetation). Also the higher resolution of images fits favorably to the observation that the texture of a surface is varying faster than its geometry. This benefit is already exploited today.
- Lidar data can also be used to constrain image matching. Given approximate exterior orientation of images, the search space can be reduced to the epipolar line. If the position of the surface is already known from Lidar, the search space can be further reduced to a short line segment. Ideally, only a small area, depending on the relation between Lidar footprint size and photo GSD, will have to be investigated to pinpoint the location of edges and corners through corresponding image points.
- Certain objects can be acquired better with a Lidar sensor, e.g., power lines or the ground below vegetation. Independent thereof, color and near infrared information from photos provide a valuable input to classify the entire area. The different appearance (see Figure 7) of some objects in the point cloud can support classification.

3. Discussion

As shown in Section 2.1, there are not only many similarities between Lidar and photo observations, but also some differences. For many object classes (vegetation, solid earth, water) the impact of those differences and the similarities are clear. However, a model to quantify those differences has not yet been established. Such a model would predict the (vertical) difference between, e.g., a crown surface over a certain tree species from airborne photo to airborne Lidar data. It is rather the case that for specific experiments the differences in the height estimation are reported (e.g., Ressler et al. 2016).

Both technologies, Lidar and photographic imaging, are developing. Also within each technology different concepts, e.g. full waveform recording vs. single photon counting Lidar, are competing. In imaging an example is the standard nadir imaging concept, augmented by oblique imaging, and moving more and more towards omnidirectional imaging as in backpack solutions. Predicting the trends of the past into the future means, that we will witness a further increase in density of measurements (or point clouds) with more and more directly measured attributes (reflectance in a number of spectral bands or single wavelengths). Those three trends, i) higher density, ii) multi-directionality, iii) directly measured attributes, will allow to replace multiple measurement tasks performed currently with terrestrial devices and especially with tactile measurement by (low flying) airborne sensors. One current limit is the accuracy and the reliability of direct georeferencing. Additionally, legal restrictions (operation of UAVs) will need to change to enable this.

The joint orientation of Lidar and photo observations as described in Section 2.2 is to be considered as one formulation for a (relatively) rigorous joint orientation of Lidar and photo observations. The principle can be applied to terrestrial data as well. It is highly automated by relying only on points, which are restricted to those areas, where both sensors provide comparable results, i.e. within smooth surfaces. While the formulation and implementation of the strip and bundle block adjustment is operational and effective, improved models may be developed in future. These models could, e.g., minimize residuals only in the Lidar observations ($x(t) + v(t)$ rather than $d(t, t')$), and especially models that integrate GNSS and INS observations. The current 2-stage approach (first georeferencing, inevitably producing some

errors, and subsequently sensor calibration to minimize discrepancies) is not optimal and can lead to locally wrong trajectories, which need to be repaired by complicated re-computation. Instead, the overlap and identity at the ground level should support the derivation of the trajectory from the very beginning. However, the joint photo Lidar orientation, together with technological developments of Lidar sensors w.r.t. internal overdetermination, will (or should allow to) decrease requirements on very precise direct georeferencing.

Finally, the joint processing of Lidar and photo data to derive 3D models is not far developed yet. Given the development of sensors, more and more sensor systems acquiring simultaneously high quality photogrammetric images and Lidar data are becoming available. It is expected, that simultaneously acquired Lidar and photo data becomes the standard for topographic acquisition, thus representing a technology push leading to new methods of integrated processing.

The point cloud will play a decisive role, as it offers the possibility to have a common data model for data coming from various sources. Thus, the development of versatile tools for the orientation and processing of point clouds is necessary and they will gain more importance in future. At the photogrammetry research unit of TU Wien the point cloud processing software OPALS is developed (Pfeifer et al., 2014).

Acknowledgements

Data used for the figures in this article was acquired by different institutions and made available in research cooperation. Figures were published and the respective references are given. Specifically the contributions of Stadt Wien, Stadtvermessung, of Riegler Laser Measurement Systems GmbH, of nFrames GmbH, and of the DFG project "Bathymetry by fusion of airborne laser scanning and multi-spectral aerial imagery" are gratefully acknowledged.

References

- J Albers, 2007: A Look Back, 140 Years of "Photogrammetry", Some Remarks on the History of Photogrammetry. Photogrammetric Engineering & Remote Sensing, 73(5), 504-506.*
- P Besl, N McKay, 1992: A method for registration of 3-D shapes. IEEE Transactions on Pattern Analysis and Machine Intelligence, 14(2).*
- J Degnan, 2016: Scanning, Multibeam, Single Photon Lidars for Rapid, Large Scale, High Resolution, Topographic and Bathymetric Mapping. Remote Sensing, 8, 958.*
- P Glira, N Pfeifer, G Mandlbauer, 2016: Rigorous Strip adjustment of UAV-based laserscanning data including time-dependent correction of trajectory errors. Photogrammetric Engineering & Remote Sensing 82 (12), 945-954.*

- P Glira, N Pfeifer, G Mandlbürger, 2019:* Hybrid Orientation of Airborne LiDAR Point Clouds and Aerial Images. International Archives of Photogrammetry and Remote Sensing, ISPRS Geospatial Week 2019, Twente, Netherlands (to appear).
- T Hakala, J Suomalainen, S Kaasalainen, Y Chen, 2012:* Full waveform hyperspectral LiDAR for terrestrial laser scanning. Optics express 20 (7), 7119-7127.
- H Hirschmüller, 2008:* Stereo processing by semiglobal matching and mutual information. IEEE Transactions on Pattern Analysis and Machine Intelligence, 30 (2), 328-341.
- M Hollaus, W Wagner, C Eberhöfer, W Karel, 2006:* Accuracy of large-scale canopy heights derived from LiDAR data under operational constraints in a complex alpine environment. ISPRS Journal of Photogrammetry and Remote Sensing 60 (5), 323-338.
- H Kager, 2004:* Discrepancies between overlapping laser scanning strips - simultaneous fitting of aerial laser scanner strips. International Archives of Photogrammetry and Remote Sensing, XXXV, B/1, 555-56.
- J Kilian, N Haala, M English, 1996:* Capture and evaluation of airborne laser scanner data. International Archives of Photogrammetry and Remote Sensing, XXXI, 3, 383-388.
- K Kraus, N Pfeifer, 1998:* Determination of terrain models in wooded areas with airborne laser scanner data. ISPRS Journal of Photogrammetry and Remote Sensing 53 (4), 193-203.
- G Mandlbürger, C Hauer, M Wieser, N Pfeifer, 2015:* Topobathymetric LiDAR for monitoring river morphodynamics and instream habitats—A case study at the Pielach River. Remote Sensing 7 (5), 6160-6195.
- G Mandlbürger, K Wenzel, A Spitzer, N Haala, P Glira, N Pfeifer, 2107:* Improved Topographic Models via Concurrent Airborne Lidar and Dense Image Matching. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, IV-2/W4, 259-266.
- G Mandlbürger, 2018:* A case study on through-water dense image matching. International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences, XLII-2.
- G Mandlbürger, H Lehner, N Pfeifer, 2019:* A Comparison of Single Photon and Full Waveform LiDAR. International Archives of Photogrammetry and Remote Sensing, ISPRS Geospatial Week 2019, Twente, Netherlands (to appear).
- M Milenković, C Ressel, W Karel, G Mandlbürger, N Pfeifer, 2018:* Roughness Spectra Derived from Multi-Scale LiDAR Point Clouds of a Gravel Surface: A Comparison and Sensitivity Analysis. ISPRS International Journal of Geo-Information 7 (2), 69.
- J Otepka, S Ghuffar, C Waldhauser, R Hochreiter, N Pfeifer, 2013:* Georeferenced point clouds: A survey of features and point cloud management. ISPRS International Journal of Geo-Information 2 (4), 1038-1065.
- N Pfeifer, G Mandlbürger, J Otepka, W Karel, 2014:* OPALS—A framework for Airborne Laser Scanning data analysis. Computers, Environment and Urban Systems 45, 125-136.
- C Ressel, H Brockmann, G Mandlbürger, N Pfeifer, 2016:* Dense Image Matching vs. Airborne Laser Scanning—Comparison of two methods for deriving terrain models. Photogrammetrie-Fernerkundung-Geoinformation 2016 (2), 57-73.
- R Schwarz, G Mandlbürger, M Pfennigbauer, N Pfeifer, 2019:* Design and evaluation of a full-wave surface and bottom-detection algorithm for LiDAR bathymetry of very shallow waters. ISPRS Journal of Photogrammetry and Remote Sensing 150, 1-10.
- TGH Tran, C Ressel, N Pfeifer, 2017:* Integrated Change Detection and Classification in Urban Areas Based on Airborne Laser Scanning Point Clouds. Sensors, 18(2), 448.
- TGH Tran, J Otepka, D Wang, N Pfeifer, 2018:* Classification of image matching point clouds over an urban area. International journal of remote sensing 39 (12), 4145-4169.
- W Wagner, A Ullrich, V Ducic, T Melzer, N Studnicka, 2006:* Gaussian decomposition and calibration of a novel small-footprint full-waveform digitising airborne laser scanner. ISPRS Journal of Photogrammetry and Remote Sensing 60 (2), 100-112.
- A Yu, M Krainak, D Harding, J Abshire, X Sun, 2010.* A spaceborne lidar for high-resolution topographic mapping of the Earth's surface. SPIE Newsroom. DOI: 10.1117/2.1201002.002655

Contact

Univ.-Prof. Dipl.-Ing. Dr. Norbert Pfeifer, Vienna University of Technology, Department of Geodesy and Geoinformation, Research Division Photogrammetry, Gußhausstr. 27-29, 1040 Vienna, Austria.

E-Mail: norbert.pfeifer@geo.tuwien.ac.a