Universität Stuttgart

Institut für Photogrammetrie

Direktor
Prof. Dr. Uwe Sörgel

Universität Stuttgart • Geschwister-Scholl-Str. 24D • 70174 Stuttgart
Institut für Photogrammetrie

Ansprechpartner
Dr. Michael Cramer

Kontakt
Geschwister-Scholl-Str. 24D
70174 Stuttgart
T 0711 685-84118
F 0711 685-83297
E-Mail:
michael.cramer@ifp.uni-stuttgart.de
www.ifp.uni-stuttgart.de

Aktenzeichen

14.06.2023

Doctoral dissertation

M.Sc. Theodora Lendzioch

UAV remote sensing of hydrological processes and fluvial dynamics

Charles University, Faculty of Science
Study Program Physical Geography and Geoecology

Experts opinion by Dr. Michael Cramer

General overview

It is all about water – this could also be a possible title for the work presented here by Ms Lendzioch. As part of her dissertation project, Ms. Lendzioch investigates the influence of water, in the form of ice, rivers and other hydrological processes on our natural environment.

The connections between the cryosphere, fluvial geomorphology and other hydrological processes are discussed. Each of these components can affect others. In order to study the effects on the earth's water system, a deep understanding of all connections is important, especially against the background of the ongoing climate crisis.

An essential basis for understanding these comprehensive ecological relationships is precise and above all up-to-date data on e.g. soil moisture, particle sizes of sediments, water balance in peatland, changes in river topographies. Unmanned, remotely controlled aerial platforms (UAVs) can make a significant contribution here, as they are a suitable means of repeatedly recording small-scale areas with relatively little effort and thus enabling time series analyses to detect changes. With the advancing technical development and, above all, the miniaturisation of sensors, UAVs as multi-sensor platforms can make a significant contribution to the analysis of environmental processes. This is one of the reasons why Ms. Lendzioch's work has mainly focused on UAVs for data collection. The practical investigations carried out - with a focus on five different applications - cover a longer period from around 2014 to 2022. The various flight platforms and sensors used therefore also provide an insight into the technological development in the field of UAVs during this period.

Ms. Lendzioch started collecting data with first-generation drone systems such as the dji Phantom 2 vision+ and customized solutions based on the MikroKopter platform. At the end of the work, the latest technology drones such as the DJI Phantom 4 RTK or the DJI Matrice series, with their respective imaging systems, came into use. The software package agisoft PhotoScan, later modified and renamed Metashape, was mainly used for the (photogrammetric) evaluation of this data, partly in combination with the Trimble software inpho Photogrammetry Suite for the creation of orthophotomosaics and the processing of terrain models and other software tools.

In addition to conventional RGB cameras, (multi-head) multispectral cameras and thermal imaging cameras were also used for data acquisition. Hyperspectral cameras or active systems like LiDAR sensors are not considered in the work due to weight and cost issues and (possibly) higher complexity. To validate the UAV-based results, additional sensors were used to collect reference data on the ground.

The UAV data was captured to solve research questions in five different application areas. The main findings are also presented below:

1. The determination of **snow depth and leaf area index** from UAV data, as discussed in two journal papers I (2016) and III (2019) with Ms. Lendzioch's contributions of 80% each.

   It has been shown, that UAV RGB images are also capable to derived DSM even for snow surfaces. Comparing this snow surface model to the underlying terrain model derived from no-snow UAV flights yields the snow depth, also known as normalized surface, which contains the height of objects (here snow surface) with respect to the terrain. DSM generation of snow surfaces is particularly challenging since the photogrammetric process entirely relies on the matching of corresponding points which depends entirely on the texture of the images. Especially fresh and bright snow has an almost texture-less surface (also dependent on solar radiation), which has a negative effect on the photogrammetric image orientation and the later dense image matching process.

   Determining leaf area index (LAI) from UAV imagery offers a major advantage as no additional equipment is required when such UAV data is available anyway. This is an advance compared to the traditional LAI methods, which are mainly based on additional terrestrial measurements and equipment. Further research to determine the optimal flight settings for improved LAI results is also mentioned.

2. The determination of **granulometry of flood river deposits**, as discussed in two journal papers II (2017) and V (2023, under review) with Ms. Lendzioch's contributions of 50% and 100% respectively.

UAV RGB imagery captured from flying heights below 10 m provide very high resolution orthophotomosaics, which form the base of the automated extraction of river sediment granulometry. It was suggested to acquire images with GSD around 1.5 mm to reliably determine the relative coarse sediment grain size in the selected study area (D_{50} percentile of mean grain size is between 13-33 mm). For automated particle size classification, traditional segmentation is compared to CNN-based classification based on the ResNet50. This optical granulometry proved to be an efficient method for particle size detection. As with all approaches related to neural networks, the quality of the CNN-based classification depends on the quality of the training data. Further limitations stem from the quality of the original UAV images, most likely due to the very high resolution required for this application.

3. The **sustainability of river restoration** projects, as discussed in the journal paper VI (2023) with Ms. Lendzioch's contributions of 25%.

   Also, in this application, the role of UAV imaging is traditionally to provide accurate 3D models of the riparian zones of various rivers, including orthophoto mosaics. From this, additional hydromorphological parameters are derived to assess the effectiveness and quality of river restoration projects. Again, the traditional products derived from UAV RGB images allowed very efficient monitoring of the river restoration process and comparison of the preliminary, planned and realized scenarios.

4. The reconstruction of **river bed topography using optical bathymetry** – in one journal paper VII (2023, under review) with Ms. Lendzioch's contributions of 15%.

   Optical bathymetry offers a flexible and relatively inexpensive approach to modeling the geometry of shallow water river beds. This approach is also based on the generation of (riverbed) elevation models from RGB imagery captured by drones. The difference to conventional elevation model generation is the influence of water. Due to the change in optical depth from air to water, the perspective model has to be modified by the corresponding refraction index (multi-media photogrammetry). The method proved to be repeatable, the absolute accuracy was not always sufficient but the shape (i.e. relative accuracy) of the river bed profile was generally preserved.

5. The estimation of **water content in a peat bog environment** – in one journal paper IV (2021) with Ms. Lendzioch's contributions of 80%.

   Similar to the data capturing in topic 2, data from multiple epochs were used to estimate the changes in soil moisture and ground water level (GWL) to find possible interconnections to microclimatic variables of the peat bog environment. Besides classical nadir looking RGB imagery, multi-spectral (including NIR) and thermal IR cameras were also used in this survey. A total of 10 different UAV missions were flown in a time interval of 15 months. Various spectral and other indices are derived from the UAV data, from which, after training, soil moisture and GWL were predicted by Random Forest classifiers and compared to the available ground truth data. The results demonstrate the complex relationships between the indices and the GWL and soil moisture variables. In particular, the models should only be transferred to other areas of applicability with great care.

General comments

The submitted thesis highlights the role of UAV based data capturing for hydrological applications. The work implements UAV sensors and platforms to provide comprehensive data sets to efficiently solve the above tasks. The topics covered in this work cover a significant part of the water cycle – from snow, to rivers to ground water level and soil moisture – which leads to a rather high complexity of the work, due to these diverse applications. Ms. Lendzioch has dealt with this complexity thoroughly and very comprehensively proving that she has complete command of the complex technologies. The results obtained are consistent with the work of other authors.

Ms. Lendzioch's work makes a significant contribution to further establishing these UAV technologies as a workhorse for such applications. The work is formally and linguistically correct and is therefore recommended for defence.

The detailed technical comments below should be considered before finalizing the thesis. In addition, I propose to replace the currently under review papers (Paper V and VII) in the Supplementary Secton with their final versions when acceptanced for publication.

Detailed comments

In the following part of the expert report, more detailed comments are to be made on some parts of the work. This commenting corresponds roughly to a standard peer review process for journal articles. Since the reviewer's expertise lies particularly in the field of sensors and photogrammetric processing, the comments here mainly focus on the parts related to UAV data acquisition and processing.

Section 2.1, page 13:
"Moreover, these sensors are only sometimes consumer-friendly and require significant resources to ensure calibration accuracy and sufficient user training to extract usable and scientific information" – in this context, the complexity of LiDAR compared to image-based photogrammetry is explained. I find it a little unfair to "blame" LiDAR in this way as all metric sensors require some level of calibration and in particular user training is required for all types of data processing. Although newer SfM programs like agisoft Metashape give the impression that photogrammetric processing is very simple, the details of (imaging) sensor georeferencing are still complex and sufficient user experience is always mandatory.

Section 2.1, Table 1, page 13:
This tables covers a wide range of different sensors used in photogrammetry.

- Why was only the Pleiades system chosen for satellite imagery? There are many others, such as the Worldview satellites now operated by Maxar Technologies. What about the European Sentinel satellites?
- The vertical accuracy of ALS is specified as 0.1 - 5 m. That's quite a wide range. In our photogrammetric world, we should normally

expect accuracy in the decimeter range (or better) for nationwide airborne mapping missions. With sophisticated scanners (such as the Riegl VUX-1) installed on drones, it may even be possible to achieve (sub)cm accuracy, as has been documented for technical infrastructure montioring applications.

- In the table, UAV-based photogrammetry (it's better to write indirect georeferencing/bundle adjustment here) performs worse than direct georeferencing. This is at least unlikely as direct georeferencing should normally be a little less accurate, especially when compared to bundle adjustments with sufficient ground control. In addition, direct georeferencing places higher demands on correct overall system calibration, which includes image sensors and additional GNSS/inertial navigation components.

Section 2.3, pages 15 / 16:
This section deals with the problem of image blur for the first time. Motion blur due to camera/platform movement (or vibration) is a problem for all airborne image sensors, especially when dealing with very high resolutions. Sophisticated motion compensation methods are established in traditional mapping cameras that are not typically transferrable to drone-based platforms other than the use of stabilized mounts/gimbals (mostly for rotary-wing UAVs). In principle, both UAV platforms, i.e. fixed-wing and rotorcraft, are affected by this motion blur.
I would also like to add another issue here related to image blur. In many of the applications presented in this work, high-resolution images are required, i. e. 1.5 mm for optical granulometry. Quite often, GSD (ground sampling distance) is taken to be equivalent to (geometric) image resolution, which is not entirely correct. The GSD is found by simply projecting the image pixel from the focal plane onto the ground (only considering the image scale), while the Ground Resolving Distance (GRD) also includes the modulation transfer function. The GRD describes the sensor systems' ability, to detect closely adjacent object structures separately. I would suggest adding this information to be more specific and explain the difference between the two measures.

Section 2.4, page 16:
"... including a Global Navigation System (GNSS) antenna for controlling navigation, flight stabilization, and location data collection." – the antenna is just one part of the GNSS system. Skip antenna here.

"Ground Control Points (GCPs) with known coordinates may be needed for geodetic positioning if a drone lacks an onboard navigation system." – GCPs are always required, unless you rely solely on the direct georeferencing. Even with precise navigation data you should use a (limited) number of GCPs for integrated sensor orientation, i.e. to introduce trajectory information as weighted observations in addition to some GCPs to get the best overall georeferencing result. This is a widely used approach in georeferencing.

Section 2.4, page 17:
"These cameras typically have three values for individual RGB pixels and simultaneously capture data from three channels using colour filters." – you should introduce the term Bayer-filter or Bayer-pattern here. Note, that each pixel captures only one of the three colour bands. Therefore, interpolation of colours is necessary to get full RGB for each pixel.

"Larger sensors generally result in higher-quality images, while shorter focal lengths are associated with levels of geometric distortion and require geometric correction (e.g., fisheye lens)." –
Large sensors could be mis-understood. This means larger pixel sizes on the sensor, which may result in a better signal-to-noise ratio and therefore better radiometric quality.
Regarding lens distortion: In order to achieve the desired accuracy, always lens distortion corrections are needed. This doesn't just apply to fisheye lenses. Shorter focal lengths (i.e. wide-angle lenses) will usually be preferred for 3D mapping as they provide better (vertical) accuracy of 3D object points, but tend to have more distortion.

"Flying at low velocities (e.g., 4ms-1) can minimize this effect, as demonstrated by Vautherin et al. (2016)." – The main intention of the Vautherin article cited here was to introduce a modified sensor model, that models the distortion caused by rolling shutter effects. This allows the movements of the camera to be modeled during the recording. The method is similar to the orientation of line scanning sensors.

"Orthorectification is a process that can correct these optical distortions and perspective changes resulting from the sensor's viewing angle and position." – One thing is missing here: Orthorectifaction corrects for the distortions caused by sensor tilt (as represented in the exterior orientation of each image) *and* the influence of relief displacement! Camera lens distortion usually is considered beforehand. A precise elevation model is necessary to orthorectify images. In traditional mapping, the classic orthophoto is based on terrain models, while the true ortho relies on the use of surface models.

"MS imaging, the most advanced spectral imaging technique for UAVs..." – what about hyper-spectral sensors? I would see them even more advanced than MS imaging systems. Hyper-spectral sensors are only briefly mentioned in section 5.3 of the thesis as helpful sensors for use in river restoration projects.

Section 2.4, page 18:
"MS sensors are rather expensive and heavy ..." – this is very general. this is quite general. If one looks on sensors like Parrot Sequioa or MicaSense RedEdge they are optimized for drone applications in terms of size, weight and power supply. The prize certainly plays a role.

"... factors such as lens distortion (vignetting)..." – lens distortion and vignetting are two different effects. The distortions affect the straight image rays, vignetting affects the radiometry by darkening the corners of the image.

Section 3.1, page 19:
"The feature-based matching uses collinearity to identify corresponding points in the images based on epipolar geometry, reconstructing multi-stereo pairs." – this should be rephrased. Feature based matching consists of two steps: (a) extraction of significant / distinguishable features in all the images and (b) matching corresponding features between the images. This provides the image tie points that are fundamental subsequent bundle adjustment. Epipolar images derived with AT results are then used for the dense image matching.

Section 3.1, page 20:
"As a result, space resection and the intersection of every tie point are resolved ..." – the bundle adjustment does not separate between space and forward intersection. Exterior orientations and object point coordinates (and additional unknowns) are estimated in just one adjustment step.

"Control information is essential for scaling and orienting the resulting sparse point cloud and photogrammetric block to determine the 3D shape of a surface accurately." – I would suggest rephrasing it here to be more concise. The role of checkpoints is two-fold: (a) they are used to fix the so-called rank deficiency/datum error of our adjusted network. With a minimum of two full GCP (X, Y, Z) and one vertical GCP (Z only) this datum is fixed. This also transfers the local network into our global (mapping) framework. (b) the control points guarantee the geometric accuracy of the block. With a sufficient number of GCPs in a suitable distribution, the object points can be determined with the expected accuracy.

"Once the epipolar geometry of the photogrammetric block has been established, differences are computed for all pixels..." – I would suggest to introduce the term parallax or disparity instead of differences here.

"The georeferenced point cloud, whether sparse or dense, can be exported and interpolated to generate DSMs or DEMs that exclude vegetation or DTMs that represent only the bare ground."- to my knowledge the DEM describes general height information, that is not yet specified for a specific surface. DTM and DSM are then specifications of the DEM. It would also be helpful to add an additional information on how the DTM is typically derived from DSM data.

Section 3.2, page 21:
"The GSD specifies the size of each pixel in the UAV image captured on the ground, and it is influenced by factors such as camera specifications, flying height, and pixel pitch. Higher flying heights result in a larger GSD, while lower flying heights produce a smaller GSD." – it is the camera focal distance, the flying height above ground and the pixel pitch that affect the GSD. If the focal length is modified, the same GSD can be obtained from different flying heights above ground.

"A smaller GSD leads to higher image resolution and increased level of detail, while larger GSD results in lower resolution and reduced level of detail." – please refer to my previous comments on GSD vs. GRD.

"This directly impacts the accuracy and precision of data obtained from the UAV survey..." – I would like to comment on this in a little more detail: A disadvantage of the study is that there is almost no discussion of the theoretical accuracy that can be derived from specific photogrammetric imaging blocks. In order to finally be able to assess the quality of the results, an initial estimate of the expected theoretical accuracy of the object points is required. Frequently one refers here "only" to the results of other groups in order to justify one's own results as meaningful. There are different ways to (pre)estimate the expected quality of the photogrammetric survey. A first simple rule of thumb assumes about 1 pixel GSD for the horizontal and 1.5 – 2 pixel GSD for the vertical accuracy, given good block geometry with strong overlaps, good ray intersection geometry (wide-angle lens) and sufficient control information. The formulas of the so-called stereo normal case could be helpful for different base-to-height ratios. In the most advanced case, the free network adjustment of the corresponding bundle block provides the precision estimate for all unknown parameters.

Section 3.2, page 22:
"evenly dispersed GCPs (about five GCPs per hectare for planar georeferencing and 10 GCPs per hectare for vertical georeferencing ..." – this is very general. In photogrammetry the number of control (without using any additional precise trajectory information) is a function of the number of images and flight lines. When we need very high-resolution imagery, the number of images covering one hectare is much higher compared to flights with very low-resolution data.

Section 3.2 / 3.3, page 23:
I would suggest not to include the name of the used software here, to keep it more general, for example "TIFF file format *using PixelWrench2* ... equalization techniques *in Adobe Suite* ... MS images can be accurately incorporated *into the Agisoft software* ..."

"image quality assessment feature that identifies images with a value below 0.5 (sometimes 0.7)" – what do this numbers mean? Any unit?

Section 3.3, page 24:
Please correct: You have used GNNS (p. 24 and p. 25 first line) instead of GNSS twice!

"Once the GCP data is imported, a manual association of points with their locations is necessary." – I would rephrase like follows: The GCPs must be manually identified and measured in all the images. The link between the point and the corresponding reference coordinates is solved by the point numbering.

"If using Agisoft Metashape, it is crucial to uncheck all geotagged photos before this step to prevent the production of a distorted DSM resulting from intervening in high-quality GCP data with less accurate GNSS data." – I never would write something like this. There are RTK drones, that provide cm-accurate GNSS trajectory data. Such data is mandatory for GNSS-supported bundle adjustment, especially when only a (very) limited number of GCPs are available and should definitely be used for the later product generation. Even less accurate GNSS can / should be used: The impact of the observation's quality on the overall bundle adjustment depends on the weighting of the observations. Thus, lower quality can be modelled by assigning lower weights to that specific observation. When there are "distortions" in the derived products it is often due to an incorrect modelling of the GNSS data in the software (i.e. missing GNSS global or strip-wise offsets), often due to inconsistencies between mapping and GNSS coordinate frame and the applied transformations.

Section 3.3, page 25:
"By optimizing camera alignment, georeferencing accuracy can be substantially improved by updating GCPs and camera-estimated coordinates with georeferencing errors" – this is not clearly stated. Optimizing the camera is the Metashape nomenclature. In the general photogrammetric wording, self-calibrating bundle adjustment is performed

here to also improve the estimated camera calibration parameters. Your wording "updating GCPs and coordinates with georeferencing error" is at least unclear.

"To ensure the desired level of accuracy, it is essential to remove GCPs with a high total reprojection error..." – before removing a GCP, one should carefully analyze why that point has a significantly larger error than others. It could be that due to the underlying error propagation in adjustment (so-called redundancy part) erroneous observations may not show the largest residuals. Then a correct point would be eliminated and the erroneous point then still left within the adjustment.

"... high-quality ground control observations are incorporated without significant errors resulting from poorly matched tie points..."- how could erroneous tie points influence GCPs? This is not clearly formulated here.

"... the workflow for MS images remains the same until step 13, but performing reflectance transformation is essential before creating the dense point cloud..." – is creating a DSM from MS imagery really necessary? At least in case additional data is available from higher resolution RGB imagery, I would suggest deriving the DSM from this data and then importing it for the orthomosaic from MS data.

"... to display thermal signatures in DN values ranging, for example, from 0 to 65535..." – you may add that 65535 values correspond to 2 byte (2^{16} bit) of data per pixel.

Section 3.4, page 25:
"... focal length, flight time, GPS, ..." – I think you don't mean the US GNSS system GPS here, rather GPS is used in the sense of GNSS? I would review the entire thesis to consistently use GNSS instead of GPS unless you really want to differentiate between the different GNSS systems.

Section 3.4, Table 3:
The pixel count for the Phantom 4 Pro is not given correctly. 20 MPix corresponds to 5472x3648 pix.

Section 3.5, page 29:
"On the other hand, the ROK is a smaller stream ..." – I am not a native English speaker, but I would rather use 'river' instead of 'stream' for this type of river. The term 'channel' is also used later. As I understood it, a channel is an (artificial) river, that has been re-shaped by humans (for shipping mostly).

Section 5.1, page 35 (also corresponding Papers I and III):
I would be really interested in getting more detailed information on photogrammetric processing of snowy areas as this is a very demanding application, although you mention that "recent studies have overcome these limitations". I know that photogrammetric processing is not the focus of the work at all - the data obtained from photogrammetry is of more interest - but it is precisely for this type of application that the quality of the DSM is crucial. I was a little confused when I saw the RMSE number given in Paper I, Tables 1 and 2. The RMS of almost 50 cm for 1 cm GSD is well above expectations, especially as this relates to the no-snow scenario.

I also didn't understand why in Paper III only the "medium accuracy" settings were chosen specifically for the alignment of the image block. With medium settings, the images are scaled down by a factor of 2 for each side. The algorithms work with pixel merged from 2 x 2 original pixels. Since the matching quality always refers to a fraction of the pixel size used, you will not use the full potential of the bundle with such settings. I know that reducing to medium speeds up processing significantly, but to get the best overall georeferencing quality you have to work with the pixel size of the original image.
In the same context, in Paper III, the term co-georeferencing is used: "Co-georeferencing has been applied to the georeferencing of both snow and bare ground DSMs based on common GCPs..." - since the GCPs should always be in the same mapping frame, additional co-georeferencing is not required as this is automatically realized by the identical GCPs! I also am not sure if the term co-georeferencing is commonly used. Usually this is known as band-to-band registration in MS sensor campaigns?
I would also like to add that some more detailed discussions about the photogrammetric quality are sometimes missing. I would expect tables detailing the precision and accuracy of each bundle adjustment and a detailed analysis of the numbers reported there. This is not included in Paper III, for example.

Section 5.1, page 36:
"... with increasing flight height, the angle of incidence becomes more oblique, affecting the amount of light reflected by the presence of needles and branches." – perhaps there is a misunderstanding, but the angle of incidence of the rays is only a function of the lens opening angle, which is due to the size of the sensor in the focal plane and the corresponding focal length. Irrespective of the flight altitude above the ground, the imaging rays always leave the lens with the same angle of incidence.

Section 5.2 (also Papers II and V), page 37:
Paper V, section 2.3: Why did you exclude the P2 term of tangential distortion from camera calibration parameter sets?
By the way, the processing here was done with the high-precision options of agisoft Metashape processing.

Section 5.3 (also Paper VI)
I would be interested in the photogrammetric block design for capturing the three almost linear river structures. The photogrammetric survey of such linear features can be challenging, since such strip-wise blocks are less stable due of their strip character. This places additional demands on the distribution of control points to ensure later object point accuracy. Unfortunately, these details of the flight designs are not mentioned in the paper.

Section 5.4 (also Paper VII)
"... with RMSE values ranging from 3.8 to 109.3 cm on site 1 and 2.6 cm to 32.1 cm on site 2." - optical bathymetry or multimedia photogrammetry is certainly one of the most advanced and complicated photogrammetric techniques. The variability of the results obtained in this work is quite high, especially considering that even small changes in the river bed design can have completely different hydrogeological effects. In my experience, errors in the 1 m range should be due to errors in the photogrammetric process chain.

Section 5.5 (also Paper IV)
I would like to reiterate that the following sentences are clear only to those readers who are really familiar with the agisoft Metashape software: "… high accuracy settings and only generic preselection. RGB dense clouds were generated using high settings and aggressive depth filtering … We used very-high settings and aggressive depth filtering for multispectral dense cloud generation." – it really should be necessary to explain, what the settings mean and why the settings are chosen that way.

Stuttgart, June 14, 2023

M. Cramer

Dr. Michael Cramer, AOR

Institut für Photogrammetrie
Universität Stuttgart
Geschwister-Scholl-Str. 24D
D-70174 Stuttgart