Developments in UAV-Photogrammetry

Markus Gerke

Technische Universität Braunschweig, Institut für Geodäsie und Photogrammetrie, Braunschweig/Germany · m.gerke@tu-bs.de

Abstract: Nowadays, unmanned aerial vehicles (UAVs) are used in many disciplines to take images of the surrounding. This might be large areas, captured in vertical or oblique imagery, or single objects like buildings, photographed from arbitrary perspectives. In conjunction with state-of-the-art software overlapping images can be turned into dense matching-based point clouds, surface models or ortho-rectified images, just to name a few. In this overview paper image acquisition by UAV is compared to conventional platforms, and the developments within photogrammetry and computer vision which were enabler of the technology we use today are sketched. Some examples show different applications of UAV-photogrammetry, also in the context of landscape or vegetation mapping.

Keywords: Remote sensing, photogrammetry, computer vision, unmanned aerial vehicle

1 Introduction

Remote sensing and photogrammetry are main sources in the creation of geospatial information. Besides others, they are also used widely in landscape mapping, c. f. (CURETON 2017, ENGLUND et al. 2017). The so-called UAV-photogrammetry underwent a tremendous development within the last years. The term UAV is an abbreviation for Unmanned Aerial Vehicle and quite generally refers to remotely controlled airborne platforms, other abbreviations and terms are also common (NEX & REMONDINO 2014). UAV-photogrammetry can partly replace data acquisition from conventional platforms, but also enables various new possibilities for many domains, both in research and practice (REMONDINO et al. 2012, COLOMINA & MOLINA 2014, GONZÁLEZ-JORGE et al. 2017). UAVs bring users a huge flexibility in terms of available sensor equipment and both, time and flight planning. The combination of affordable UAVs as versatile flying platforms for various sensors and photogrammetric software makes it possible for landscape architects to create a geo-referenced 3D site model at much lower cost and faster than through conventional topographical survey methods. Data collected with UAVs is not limited to topography though but can also be used to assess flood risk, site vegetation or to document landscape construction works.

Software for the device operation, but also for data processing is mature and mostly automatic, but in order to derive an end-product of high quality, still fundamental concepts need to be applied. In addition, a problem is still the under developed and unfavourable regulatory situation in many countries, c. f. (STÖCKER et al. 2017A) and restricted flight times – and therefore restricted coverage.

This paper attempts to analyse the success of UAV-based remote sensing, focussing on passive sensing, i. e. laserscanning-based approaches are not considered. In addition, technological details of the UAV-platforms are not within the scope of this paper. The next section compares conventional and UAV-based platforms concentrating on mapping applications. In section 3 the key developments in photogrammetry and computer vision in terms of image calibration and orientation, which made the efficient use of UAV images possible, are sketched. Section 4 reviews some applications, before the paper is concluded.

Journal of Digital Landscape Architecture, 3-2018, pp. 262-272. © Wichmann Verlag, VDE VERLAG GMBH · Berlin · Offenbach. ISBN 978-3-87907-642-0, ISSN 2367-4253, e-ISSN 2511-624X, doi:10.14627/537642028.

2 UAVs in Comparison to Conventional Platforms

UAV-photogrammetry fills the gap between remote sensing/conventional and close-range photogrammetry. This statement does not only relate to scale in the sense of ground resolution, but also to the possibility to realize almost arbitrary camera location and perspective. This is still the case today, despite the fact that oblique airborne photogrammetry from conventional platforms gained attention within the last decade, including regular acquisition by municipalities (REMONDINO & GERKE 2015). However, here we need to distinguish between multi-rotor and fixed-wing UAV-platforms. While the mentioned flexibility holds for the former category, we face more restrictions when it comes to the latter, where the configuration is comparable to the conventional airborne case. On the other hand, due to limited flight time caused by physical restrictions, the possible area coverage by fixed-wing is a multitude of that of multi-rotors. In this respect the recent development of hybrid UAV, which are fixed-wing planes with the possibility of hybrid vertical take-off and landing (VTOL) should be mentioned (OZDEMIR et al. 2014).

Fig. 1: Conventional oblique images (left column) and UAV images (right column) from the city of Dortmund/Germany. The black rectangles in the left images indicate the area covered by the UAV flight (ISPRS 2018).

Figure 1 shows typical oblique images from a conventional airborne platform and a UAV, taken from the ISPRS/EuroSDR benchmark for multi-platform photogrammetry (NEX et al.

2015, ISPRS 2018). The average ground sampling distance (GSD) for the airborne data varies from 8 to 12cm for oblique images, and it is between 1 and 3cm for UAV. While in the former case a relatively large area is captured, the latter are suitable for single object-related studies. This aspect of area coverage vs. ground resolution and flexibility is also covered in Table 1, where different platforms are compared.

Table 1: Platforms for photogrammetry – comparison under different aspects

Aspect	Conventional airborne	Close-range	UAV
Capture geometry – object visibility	Nadir (vertical), oblique – land cover, terrain, roofs, facades (restricted)	Terrestrial horizontal and oblique views – facades/vertical faces	Full flexibility: nadir, oblique, horizontal, upwards – more complete building coverage
Area coverage	1-n km^2	Single objects	Single objects to approx. 15 ha (multi-rotor), 45 ha (fixed-wing)
Typical ground resolution per pixel (RGB camera)	> = 5 cm	in mm-range	in cm-range or less
Multitemporal acquisition	Depending on weather, season, budget	As often as needed if outer constraints allow	As often as needed if outer constraints allow, possible regulatory restrictions
Optical sensors	RGB, Multispectral, Thermal, Hyper-spectral	RGB, NIR, Thermal	RGB, Multispectral, Thermal, Hyperspectral

While direct sensor orientation has found its way into daily practice for conventional platforms, light-weight technology for highly accurate GNSS and attitude sensing only recently is available for UAV. Especially fixed-wing platforms come with quite matured solutions (GERKE & PRZYBILLA 2016, STÖCKER et al. 2017B). If such devices are not available, the classical indirect sensor orientation, employing well distributed ground control points (GCP), is still necessary in order to derive products of high quality and reliability.

3 Developments in Photogrammetry – some Background

The last decade has seen a rapid development of image calibration and orientation approaches. Especially the number of software packages freely or commercially available which are able to reconstruct the 3D scene geometry from just the images, including proper image calibration and orientation, and dense point matching is increasing.

Major developments were driven by the computer vision (CV) community and there is a constant discussion between the "photogrammetrists" and CV community about the "right" approach. An interesting view on the different opinions is summarized in a paper which is more than 15 years old (FÖRSTNER 2002), and by now many items mentioned there regarding closer cooperation between the fields are achieved.

When reading literature related to this topic, one might be confused by the different terms used. When photogrammetrists talk about "calibration" they normally refer to the computation of the camera interior (IO) parameters principal distance, principal point coordinate or lens distortions. The term "image orientation" is then used when it comes to find the exterior parameters (EO: location and pose of camera). In CV-terms often the term "calibration" is used for both mentioned steps, as the calibration of internal parameters is considered as an integrated part of the entire workflow, which is sometimes also referred to as "structure-from-motion", SfM, see below.

One major task in photogrammetry is to reconstruct the image geometry with the ultimate aim to derive further information from image measurements: If we know both, the IO and the EO of the image, we are able to relate any point in the 3D object space to a point in the image plane. In a stereo or multiple view setup, 3D coordinates of all the points which are visible in at least 2 images can be computed utilizing the rays intersection concept. Thus, all products derived within a photogrammetric workflow, like height/terrain models, ortho-rectified images, topographic maps, 2D/3D (GIS) information, etc. do rely on proper image orientation information.

The focus of research in the computer vision domain has been more on the questions on how to derive the scene geometry from uncalibrated cameras and even without any pre-knowledge of the unknown parameters. In retrospect there were some cornerstones which gave a boost to the development of all the fully automatic image orientation/calibration software which is available today.

Robust estimation of the fundamental (F) matrix: The F-matrix models the coplanarity constraint and implicitly represents what the photogrammetrists call "relative orientation" between two images, but including the modelling of interior camera parameters. In the early 1990s researchers found algorithms on how to compute the F-Matrix from point correspondences, see e. g. (FAUGERAS 1992), but also in close-range photogrammetry similar approaches for calibrated cameras have been developed, cf. (HINSKEN 1987). If embedded into a RANSAC sampling approach, this can be done robustly, i. e. even with a substantial number of blunders (HARTLEY & ZISSERMAN 2003).

Structure from motion/from video: In case we do not observe the scene with a still frame camera, but with a video camera we have the advantage that the search for frame-to-frame correspondences boils down to a so-called feature-tracking problem, see e. g. KLT (LUCAS & KANADE 1981). The assumption is that corresponding features are hardly moving between two frames. Together with some more techniques like the mentioned F-matrix estimation, adding more frames to the solution using resection, some advanced self-calibration technique and finally a bundle adjustment, we are able to reconstruct a scene from a video sequence up to an unknown scale, see e. g. for an overview (POLLEFEYS et al. 2004). As a final result one obtains the EO parameters of the cameras, self-calibration (IO) information and the 3D point coordinates of tie points in object space. Without any additional information, however, this Euclidean reconstruction is only up to scale. This group of techniques where a sequence of images showing a static scene, resulting in individual camera orientations and 3D sparse point clouds is called structure-from-motion (SfM). The disadvantage of video-based SfM is that we have short baselines between images and this results in a quite bad ray intersection geometry and in addition the resolution of video cameras is not good, at least compared to high resolution still frame images. The interested reader can refer to MUSIALSKI et al. (2013) for an overview on the developments in SfM and other techniques.

Local scale invariant features and point descriptors: The video-based SfM works, because tie information is retrievable through feature tracking in video frames. But what is if so-called wide base-line images are involved, i. e. images, where corresponding features are probably not close by in adjacent images, and − in addition − where we do not know which image overlaps with which one? Since we also do not assume to have any pre-knowledge about the scene and image orientation, the traditional area-based matching techniques would result in too many false matches. This is because of all the ambiguities, let alone the computational complexity if the search space cannot be reduced. If it would be possible, however, to find matches in those wide baseline images, we could exploit the much better image geometry.

Thus, a method is needed which is able to "describe" a point in an image, independently from its scale or rotation of the image and independently of any prior information on image orientation. A solution to this problem is found in the Scale Invariant Feature Transform, SIFT (LOWE 1999, 2004). The SIFT operator (and nowadays further developments like e. g. SURF or ASIFT) works on the basis of gradient histograms. The idea is that in a region, which is salient and stable over many scales, the histogram around this point are quite unique and also largely illumination invariant. Ultimately, each point is described by a feature vector which encodes the histograms in a well-defined pattern around the stable region. By finding the closest feature vector within the (stereo) mate image we can identify possible matches for any point in a source image. Those possible matches can then be used in the described workflow, but the robust estimation of the F-matrix is very important to become insensitive against wrong matches.

One remarkable step has been achieved by Noah Snavely who released an open source tool "bundler", which implements the entire workflow, also making use of SIFT, see software and papers on (SNAVELY 2018), or the project "Reconstructing Rome" (AGARWAL et al. 2010).

It seems that the increasing use of UAVs for photogrammetric applications during the last decade gave the final momentum to the further merge of photogrammetry and computer vision approaches on image orientation and calibration. On the one hand some photogrammetric software could not in the beginning solve the problem of image orientation for UAV images because the basic assumption of having good approximate values and a well-ordered set of images may not be satisfied in UAV image-blocks. On the other hand, the so-far used CV methods provided means to orient the images, but statistic measures and – more important – the thorough use of ground control information are usually not applied in those methods.

Current state-of-the-art software products combine advantages of both worlds, photogrammetry and computer vision. They do not require good approximations, because basically they follow a SfM workflow for large baseline images, including self-calibration of off-the-shelf cameras. However, if proper 3D reference points are provided they are used for geo-referencing including some statistic measures, like residuals at check points or correlation between estimated parameters. If only approximate GNSS locations are available those are used to estimate the scale and geo-reference. Although final products like points clouds or ortho mosaics are put into the right coordinate frame in that case, the final accuracy and possible block deformation cannot be quantified (NEX & REMONDINO 2014). Hence, a careful use of ground control points and independent check points is necessary whenever a qualified measurement task is aimed at with the derived data.

4 Examples

In the following some exemplary UAV projects are sketched which show some typical applications. The first one is related to urban mapping: parts of the campus of the Technical University of Braunschweig have been acquired with a DJI Phantom 4Pro. The average image resolution was 1.5cm and besides a regular grid or nadir and oblique images, some individual shots in free-flight mode have been taken. The upper left image in Figure 2 shows a slanting view onto the dense matching point cloud, overlaid with image locations, while the right-hand image shows the same with colour coding the height. In the lower row the digital surface model (DSM) and ortho image is superimposed to the geocoded city map and the right images give close-up views of signalized control points.

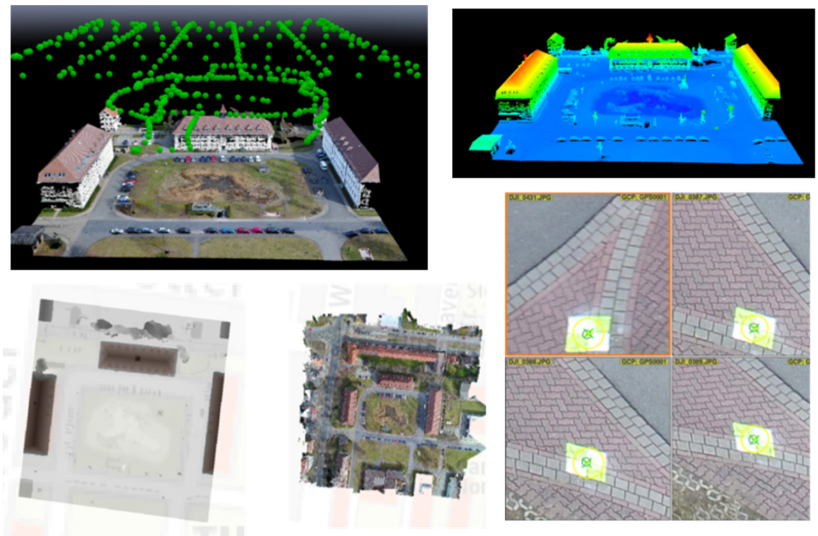

Fig. 2: Data acquisition in an urban context. Upper row: dense image matching point cloud, lower row: DSM and ortho image superimposed to city map, close view to ground control points.

TUMLISAN (2017) employed multi-temporal UAV data to monitor the growth of maize (Figure 3). It was shown that thanks to the very high resolution and accuracy of data the different states of maturity can be identified and quantified. A pre-requisite to compute the shown difference height models is to realize a very good absolute accuracy, hence several 3D control points were spread in the field to enable a precise geo-referencing of each epoch.

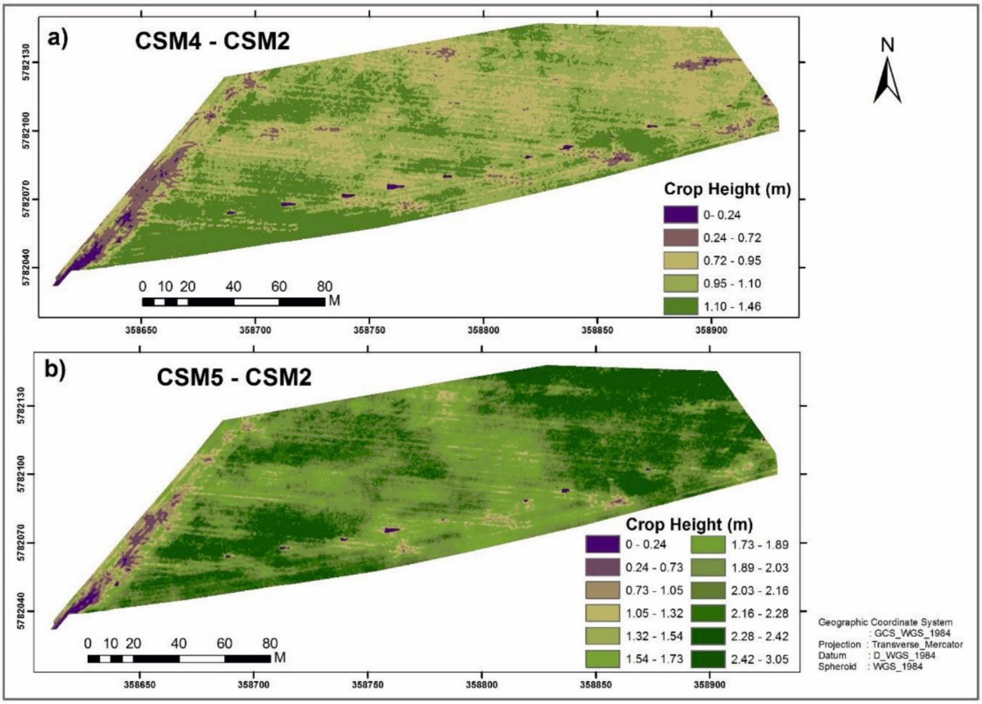

Fig. 3: Difference height models of maize (TUMLISAN 2017)

In another work the impact of different flight configurations for UAV data acquisition was tested (NASRULLAH 2016). One conclusion was that RTK-supported GNSS significantly helps to increase the accuracy, oblique images bring an additional stabilisation to block geometry and in case of flat terrain a so-called cross flight pattern supports camera calibration. Cross flight pattern means to cover the area in two main directions of 90° rotation, possibly in two different altitudes. Figure 4 shows in the upper row some charts of vertical accuracy depending on the configuration, in the lower row a dense matching point cloud is printed: left with natural colour, right with colour coding the height. The very good data quality is reflected in the sharp edges of natural objects.

5 Conclusions and Outlook

This paper aimed to show that UAV-photogrammetry is relatively young offspring from photogrammetry and computer vision, but nevertheless it is already well established in many disciplines. Very high resolution georeferenced (ortho) images, point clouds, height models, etc. are the base for a multitude of geospatial information. Methods for dense point matching, height modelling, true ortho projection, etc. are not discussed here, but those techniques underwent an enormous development in the last decade, as well. Unmanned aerial vehicles give researchers and practitioners the freedom to capture current remote sensing data almost at any point in time, i. e. when needed for a certain task. Obstacles like under-developed regulatory frameworks might hinder the further development in some countries. When working with state-of-the-art software to process the images one should take care when it comes to assessing data quality: problems like block deformation and the verification of final (absolute) accuracies is only possible if well-established concepts of photogrammetry are applied.

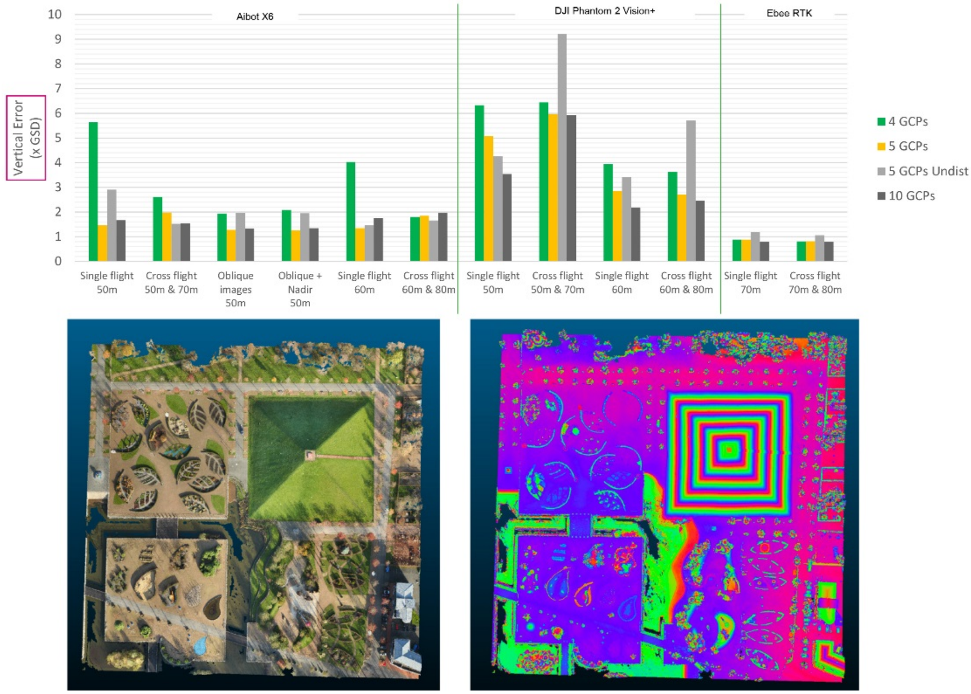

Fig. 4: Test of different configurations, natural-coloured point cloud and colour-coded height, former area of the state garden show 2003 in Gronau/Germany (NASRULLAH 2016)

Development is on-going in many directions. Concerning platforms, the importance of VTOL fixed-wing hybrid systems might increase. This becomes also interesting in conjunction with autonomous charging and "drone-port" systems, see eg. the skyport by SKYSENSE (2018), Figure 5. Here, the approach is to enable a fully autonomous drone flight including control and data handling via internet. With such a drone port and VTOL hybrid systems large areas could be covered, e. g. in the context of regular monitoring applications. Possibly, the biggest threat currently for such systems is the regulative reality in most countries which does only allow autonomous or out of visual line flights in some exceptional cases.

On the data processing side, we will see an increasing automation in the semantic analysis: while today we have already a fully automatic processing from images to point clouds etc., as detailed above, the automatic interpretation of data is limited.

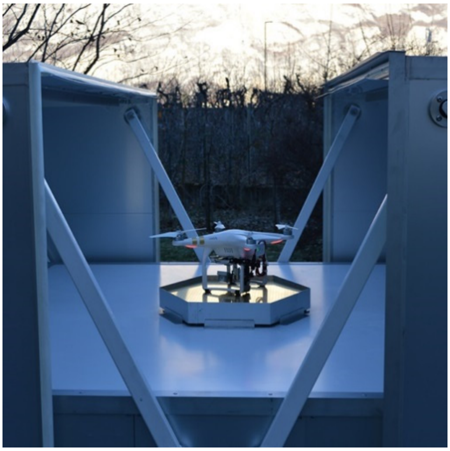

Fig. 5: Skyport and charging pad by SKYSENSE (2018), (c) Skysense

While separate software, like for instance for image classification, is available for a long time already, an integrated approach where the UAV data is used directly is only shortly integrated, e. g. in Pix4dmapper (PIX4D 2018), c. f. Figure 6. This trend to add more functionalities in the direction of data interpretation will probably continue.

Fig. 6: Pix4dMapper: rule-based land cover classification in dense image matching point-clouds (PIX4D 2018)

References

AGARWAL, S., FURUKAWA, Y., SNAVELY, N., CURLESS, B., SEITZ, S. M. & SZELISKI, R. (2010), Reconstructing Rome. Computer, 43 (6), 40-47.

COLOMINA, I. & MOLINA, P. (2014), Unmanned aerial systems for photogrammetry and remote sensing: A review. ISPRS Journal of Photogrammetry and Remote Sensing. https://doi.org/10.1016/j.isprsjprs.2014.02.013.

CURETON, P. (2017), Strategies for Landscape Representation – Digital and Analogue Techniques. Routledge.

ENGLUND, O., BERNDES, G. & CEDERBERG, C. (2017), How to analyse ecosystem services in landscapes – A systematic review. Ecological Indicators, 73, 492-504. https://doi.org/10.1016/j.ecolind.2016.10.009.

FAUGERAS, O. D. (1992), What can be seen in three dimensions with an uncalibrated stereo rig? BT – Computer Vision – ECCV'92: Second European Conference on Computer Vision Santa Margherita Ligure, Italy, May 19-22, 1992 Proceedings. In: SANDINI, G. (Ed.) (Computer V, Vol. 588, pp. 563-578). Springer Berlin Heidelberg. https://doi.org/10.1007/3-540-55426-2_61.

FÖRSTNER, W. (2002), Computer Vision and Photogrammetry – Mutual Questions: Geometry, Statistics and Cognition. Bildteknik/Image Science, Swedish Society for Photogrammetry and Remote Sensing, 151-164.

GERKE, M. & PRZYBILLA, H.-J. (2016), Accuracy analysis of photogrammetric UAV image blocks: Influence of onboard RTK-GNSS and cross flight patterns. Photogrammetrie – Fernerkundung – Geoinformation, 2016 (1), 17-30. https://doi.org/10.1127/pfg/2016/0284.

GONZÁLEZ-JORGE, H., MARTÍNEZ-SÁNCHEZ, J., BUENO, M. & ARIAS, P. (2017), Unmanned Aerial Systems for Civil Applications: A Review. Drones, 1 (1), 2. https://doi.org/10.3390/drones1010002.

HARTLEY, R. & ZISSERMAN, A. (2003), Multiple view geometry in computer vision (Second Ed.). Cambridge: Cambridge University Press.

HINSKEN, L. (1987), Algorithmen zur Beschaffung von Näherungswerten für die Orientierung von beliebig im Raum angeordneten Strahlenbündeln. Deutsche Geodätische Kommission bei der Bayerischen Akademie der Wissenschaften. Reihe C, Nr. 333.

ISPRS (2018), ISPRS /EuroSDR Benchmark for Multi-Platform Photogrammetry. http://www2.isprs.org/commissions/comm1/icwg15b/benchmark_main.html (retrieved February 9, 2018).

LOWE, D. G. (1999), Object recognition from local scale-invariant features. Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on, 2, 1150-1157. https://doi.org/10.1109/ICCV.1999.790410.

LOWE, D. G. (2004), Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision, 60 (2), 91-110. https://doi.org/10.1023/B:VISI.0000029664.99615.94.

LUCAS, B. D. & KANADE, T. (1981), An iterative image registration technique with an application to stereo vision (pp. 674-679). Morgan Kaufmann Publishers Inc., Vancouver, BC, Canada.

MUSIALSKI, P., WONKA, P., ALIAGA, D. G., WIMMER, M., VAN GOOL, L. & PURGATHOFER, W. (2013), A survey of urban reconstruction. Computer Graphics Forum, 32 (6), 146-177. https://doi.org/10.1111/cgf.12077.

NASRULLAH, A. R. (2016), Systematic analysis of unmanned aerial vehicle (UAV) derived product quality. Enschede, University of Twente Faculty of Geo-Information and Earth Observation (ITC), unpublished.

NEX, F., GERKE, M., REMONDINO, F., PRZYBILLA, H.-J., BÄUMKER, M. & ZURHORST, A. (2015), ISPRS benchmark for multi-platform photogrammetry. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, II-3/W4, 135-142. https://doi.org/10.5194/isprsannals-II-3-W4-135-2015.

NEX, F. & REMONDINO, F. (2014), UAV for 3D mapping applications: a review. Applied Geomatics, 6 (1), 1-15. https://doi.org/10.1007/s12518-013-0120-x.

NOCERINO, E., MENNA, F., REMONDINO, F. & SALERI, R. (2013), Accuracy and block deformation analysis in automatic UAV and terrestrial photogrammetry – lesson learnt. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, II-5/W1, 203-208. https://doi.org/10.5194/isprsannals-II-5-W1-203-2013.

OZDEMIR, U., AKTAS, Y. O., VURUSKAN, A., DERELI, Y., TARHAN, A. F., DEMIRBAG, K. & INALHAN, G. (2014), Design of a commercial hybrid VTOL UAV system. Journal of Intelligent and Robotic Systems: Theory and Applications, 74 (1-2), 371-393. https://doi.org/10.1007/s10846-013-9900-0.

PIX4D (2018), Pix4d homepage. https://pix4d.com https://pix4d.com (retrieved February 9, 2018).

POLLEFEYS, M., VAN GOOL, L., VERGAUWEN, M., VERBIEST, F., CORNELIS, K., TOPS, J. & KOCH, R. (2004), Visual modeling with a hand-held camera. Int. Journal of Computer Vision, 59 (3), 207-232. https://doi.org/10.1023/B:VISI.0000025798.50602.3a.

REMONDINO, F., BARAZZETTI, L., NEX, F., SCAIONI, M. & SARAZZI, D. (2012), UAV photogrammetry for mapping and 3d modeling – current status and future perspectives. ISPRS – International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XXXVIII-1/, 25-31. https://doi.org/10.5194/isprsarchives-XXXVIII-1-C22-25-2011.

REMONDINO, F. & GERKE, M. (2015), Oblique Aerial Imagery – A Review. Photogrammetric Week 2015, 75-83.

SKYSENSE (2018), Skysense homepage. http://www.skysense.co (retrieved February 9, 2018).

SNAVELY, N. (2018), Bundler – Structure from Motion (SfM) for unordered image collections. http://www.cs.cornell.edu/~snavely/bundler/ (retrieved February 9, 2018).

STÖCKER, C., BENNETT, R., NEX, F., GERKE, M. & ZEVENBERGEN, J. (2017), Review of the current state of UAV regulations. Remote Sensing. https://doi.org/10.3390/rs9050459.

STÖCKER, C., NEX, F., KOEVA, M. & GERKE, M. (2017), Quality assessment of combined IMU/GNSS data for direct georeferencing in the context of UAV-based mapping. ISPRS – International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLII-2/W6, 355-361. https://doi.org/10.5194/isprs-archives-XLII-2-W6-355-2017.

TUMLISAN, G. Y. (2017), Monitoring growth development and yield estimation of maize using very high-resolution UAV-images in Gronau, Germany. Enschede, University of Twente Faculty of Geo-Information and Earth Observation (ITC), unpublished.