

Quality assessment of high-resolution UAV imagery and products

MICHAEL CRAMER¹, SHUHANG ZHANG¹, HENRY MEIßNER² & RALF REULKE²

Abstract: The objective of this paper is to evaluate the performance of high-end and regular UAV-based camera systems. Different factors contribute to the overall accuracy. The evaluation partly relies on the methods, which are part of the new upcoming German standard DIN 18740-8 “Photogrammetric products – Part 8: Requirements for image quality (quality of optical remote sensing data)”. The image data quality in general is quantitatively evaluated at different processing levels e.g. uncorrected, corrected original image, influence of debayering, orthoimage processing, image restoration, etc. This requires pre-processing of the image data to produce a comparable data quality, the acquisition, provision and processing of reference data and additional information. This analysis includes the spatial resolution. Furthermore, the geometric camera stability and the influence of different image block constellations directly influences the overall 3D object point quality. This is typically evaluated from test sites. For empirical testing UAV-based images from the DJI Phantom 4 series with proprietary in-built cameras are compared to drone images taken with Phase One iXM 100 MPix camera.

1 Introduction

Photogrammetry provides methods to obtain geometric and thematic information from data captured by different sensors i.e. cameras, radar and laser systems. The quality of the sensor data itself plays an important role for the quality of derived products. Defined quality requirements for the estimation of quality of remote sensing data are necessary. This is why, according to the German Standard series DIN 18740, a new standard for the estimation of optical remote sensing data is close to be published soon. This will be available as *DIN 18740-8, Photogrammetric Products – Part 8: Image quality requirements (quality of optical remote sensing data)*. Parameters describing image quality are of relevance in different application scenarios like

- sensor and mission design,
- comparison of sensors (using defined parameters and numbers),
- image quality estimation including sensor resolution,
- algorithm development,
- in-flight / in-orbit behaviour of the instrument,
- traceability to standardised parameters and instruments.

It is complex to find objective definitions for the parameters of image quality. This is why there is no general standard for describing image quality. In general, the requested quality of image data depends on the problem to be solved (e.g. point detection and co-registration, object recognition and tracking, classification in remote sensing, images related to the human visual system) and

¹ Universität Stuttgart, Institut für Photogrammetrie (ifp), Geschwister-Scholl-Str. 24D, D-70174 Stuttgart, E-Mail: [Michael.Cramer, Shuhang.Zhang]@ifp.uni-stuttgart.de

² Deutsches Zentrum für Luft- und Raumfahrt (DLR), Institut für Optische Sensorsysteme, Rutherfordstraße 2, D-12489 Berlin-Adlershof, E-Mail: [Henry.Meissner, Ralf.Reulke]@dlr.de

cannot be defined in the same way for every imaging system in general. It depends on factors like image content (structure), sensor performance, accuracy (geometric, radiometric, and spectral) and several other factors. The image quality can be determined at different processing levels. This requires pre-processing of the image data to produce a comparable data quality, the acquisition, provision and processing of reference data and additional information. With that, also the influence of image processing algorithms is evaluated. In order to objectively evaluate the quality of the sensor, representative image data is necessary. This is why defined artificial test patterns or suitable natural structures and features are imaged. For resolving power estimation sharp natural edges or artificial resolution charts like bar pattern or Siemens stars are typically in use. The image quality is maybe the most important part in any photogrammetric imaging and processing chain as it directly influences the tie point extraction and multi-stereo matching. This now can be evaluated following the definitions in the new upcoming standard DIN 18740-8.

The quality of 3D object points, as one main result in geometric reconstruction, is closely dependent on the individual image block geometry. This image acquisition fixes the overlap between neighbouring images and therefore is of importance for the final overall geometric accuracy. Any derived product, like dense point cloud and orthophotomosaic relies on this image block constellation and georeferencing. This also includes the quality of in-situ calibration to determine the internal sensor geometry containing any additional (optional) boresight parameters. This is not covered by the new standard, as it is, similar to the individual requests on image quality as mentioned above, application dependent. Each application will have its specific scenario and with that it is a problem to find standardized procedures to evaluate the performance of georeferencing and derived products. Within this context the availability and accessibility of official evaluation and calibration test sites is often mentioned and for sure, the use of highly accurate test sites is widely accepted and several campaigns already have been done to estimate individual sensors performance from test fields, i.e. the German DGPF Camera Test (CRAMER 2010). Still, this is typically only done for very few samples of a sensor series, in scientific context quite often only one single specific sensor is very deeply analysed. With that it is more or less a sensor type, i.e. sensor series than a specific individual sensor characterization. Even when flown in the same test site with same flight configuration, slightly different environmental conditions or different operators' expertise may already influence the processing result. This all does not consider the fact, that due to operational issues multiple defined evaluation test sites have to be installed, distributed all over the country to allow efficient overfly while systems are in air and on their way to operation. Due to such complexity, standardized test sites and test site evaluations are not yet considered in German standards, even though the US Geological Survey (USGS) some years ago started to put effort into such initiative and some freely accessible test sites have been installed and maintained in the US by now (USGS 2020). Three of these sites are designed for geometric evaluation of satellite and aerial sensors. The quality check is based on the evaluation of the final product, the orthophoto namely – this is different to the current methodology in DIN 18740-8, which focuses on the (original) image data. As mentioned in the objectives from USGS, “accurate and standardized high resolution orthophotos (reference image) will be used to determine the relative accuracy of other high resolution aerial and satellite images. Co-registration analysis will be performed between this image and other search images. The results of the analysis will provide insights into the accuracy

of the imaging system.” Another forth test site is especially prepared for the geometric in-situ calibration of satellites and aerial sensors.

Within the following sections, the evaluation of empirical UAV-based images will be done in two parts. The first one is following the concept as proposed in the DIN 18740-8 standard. A Siemens star pattern is captured in several overflights from which the spatial resolution is derived in a defined way. The software used here will become official part of the future standard. In a second step, the overall 3D object point performance is obtained from check point analysis. Highly precise reference points have been signalized in a local test field. In order to compare different level of UAV sensors, two different systems are selected. There is a large variety of operational camera UAVs in the market, many of them especially designed for survey and photogrammetric mapping applications. Additional RTK-GNSS functionality for the precise trajectory implementation and measurement is common for the professional systems and can help to obtain the geo-data in a more effective way. The DJI Phantom 4 series with its proprietary in-built cameras represents a very popular drone system increasingly used for 3D mapping. In addition to this, larger professional mapping drone platforms are able to lift up heavier sensors, like latest generation Phase One iXM cameras providing image images up to 150 MPix. Their performance will be characterized and compared in the following.

2 Sensors and test data

The DJI Phantom 4 and Phantom 4 RTK are all consumer-level quadcopters manufactured by DJI (DJI 2019b; DJI 2019c). The Phantom 4 model is designed for daily aerial image and video capture, while the Phantom 4 RTK is designed for aerial mapping. Compared with professional mapping drones, the Phantom quadcopters are lighter in weight and cheaper in price; both systems are all within 1.4 kg. The compact design and lightweight advantages of Phantom systems allow for one-man operation, which makes them more cost-effective. Both systems have ca. 25 minutes flying time in normal cases. The Phantom 4 is equipped with a consumer-grade GNSS receiver, inertial measurement unit (IMU), and barometer for performing autonomous flights. It can achieve accuracy up to ± 1.5 m vertical and ± 0.5 m horizontal positioning accuracy (DJI 2019c). The Phantom 4 RTK has a multi-frequency multi-system high-precision RTK GNSS that can achieve positioning accuracy up to 1 cm+1 ppm horizontal, 1.5 cm+1 ppm vertical (DJI 2019b).

Both DJI Phantom systems have integrated gimbal cameras with 3-axis stabilization that can compensate minor vibration or instability of the drone during mapping flights. For a detailed comparison, parameters of the cameras are listed in Table 1. Both cameras are proprietary and have fixed focal length, which leads to better stability of interiors and therefore more ‘metric-like’ (CRAMER et al. 2017). The camera of Phantom 4 RTK shares similar specifications compared with that of Phantom 4, but the increase in sensor dimensions and resolution is significant, which leads to higher resolving ability. In addition, the Phantom 4 RTK camera has a mechanical shutter that avoids the rolling shutter effect to further increase the imaging quality. The ground controlling software of these two systems differs. Although Phantom 4 is not designed for aerial mapping, it has an official flight planning software called DJI GS Pro (DJI 2019a). It can plan nadir/oblique-imaging missions according to given ground sampling distance (GSD) and overlap automatically. The Phantom 4 RTK has a built-in software DJI GS RTK, it is similar to DJI GS Pro but with

additional RTK management features. Additionally, the factory-calibrated interiors and measured exteriors are stored in the XMP data of images, they can be imported to data processing software like Agisoft Metashape or Pix4DMapper directly.

Different to the DJI integrated system solution, the Phase One cameras in principle can be integrated in any type of professional drone with sufficient maximum take-off mass (MTOM) and corresponding interfaces. As the growing UAV mapping market is foreseen as one of the most important field of application for this camera, the iXM-RS already provides fully integrated solutions for the integration with existing drone platforms like the DJI M600 Pro aerial platform. The camera offers different optional lenses with different focal length and characteristics. The selected 35 mm lens is a fixed focus lens, focussed to infinity. The corresponding depth of field starts at 30 m minimal distance with maximum opening of aperture. This has to be considered when setting the flying height above ground. However, these fix focus lenses can also be set to a fixed distance, non-equal to infinity, which is equivalent to the mean flying height above ground for example. This must be adjusted at the factory and it is not possible for the user to later change this setting. In addition Phase One also offers so-called ‘motorized’ lenses, where the focus can be changed and a precise mechanism allows to repeat defined focus positions with a repeatability better than 6 μm (WIESER 2018). The iXM100-RS uses back-side-illuminated CMOS sensors, which should allow for higher dynamic range, improved sensitivity and less noise despite smaller pixels compared to the Phase One iXU predecessor models. Details can be found in Tab. 1 and PHASE ONE (2020).

Tab. 1: Camera parameters of DJI Phantom 4 / 4 RTK and Phase One iXM100-RS

Camera	Phantom 4	Phantom 4 RTK	iXM100-RS
Sensor format [mm] / [pix]	6.17 × 4.55 / 4000 × 3000	13.2 × 8.8 / 5472 × 3648	43.9 × 32.9 / 11664 × 8750
Pixel Count	12 Mpix	20 Mpix	100 Mpix
Pixel pitch [μm]	1.56	2.41	3.76
Focal length [mm] / 35mm equiv. focal length [mm]	3.6 / 20	8.8 / 24	35 / 28 (other lenses available)
Field-of-View (FOV)	94°	84°	76°
ISO range	100-1600	100-12800	50-6400
Shutter speed [s]	8 - 1/8000	8 - 1/2000	Up to 1/2500
Shutter type	Rolling	Mechanical	Mechanical
Focus	fixed at ∞	AF/MF	Fixed at ∞ (for selected lens)
Max f-stop	1:2.8	1:2.8	1:5.6
Image format	DNG (raw), JPG	JPG only	IIQ (raw)

For data acquisition, the camera was integrated in the CS-SQ8 multi-copter drone provided by CopterSystems. The gimbal mount was especially designed and optimized for this iXM camera. In order to get precise GNSS/inertial trajectory information, the Applanix APX-15el UAV board was additionally integrated and time synchronized with the camera exposures. This unit contains a multi-frequency multi-system high-precision RTK GNSS and a dual IMU in order to provide support for the movements of the mount with respect to the copter airframe. The post-processed positioning accuracy is around 2-5 cm, the roll and pitch angle accuracy around 0.025° and the heading accuracy is about 0.08° according to the manufacturers' specifications.

3 Quality evaluation

3.1 Geometrical resolution

As motivated above, the quality is analysed in two different ways, namely from original images and from derived products. For estimating the resolution of the original images, a software is used, which will become publicly available as part of the upcoming DIN 18740-8 standard. This software uses the modulation transfer function (MTF) which is the spatial frequency response of an imaging system to a given illumination and therefore estimates the effective image resolution or resolving power, which then can be transformed into object space to get so-called ground resolving distance (GRD) or true ground sampling distance (tGSD). Using the a priori knowledge of the original scene – in this case the well-known Siemens star target – contrast transfer function (CTF), MTF and point spread function (PSF) are approximated by a Gaussian shape function. The mathematical foundations and further details can be seen from MEIBNER et al. (2018, 2019). The parameter σ (standard deviation) of the PSF (assuming a Gaussian-shape function) is one criterion for spatial resolution. It directly refers to image space and can be seen as objective measure to compare different camera performances. Another criterion is the spatial frequency where the MTF reaches a certain (minimal-)value (i.e. 10%, MTF10). The reciprocal of that frequency is the approximation for the size of the smallest line per pixel. This MTF10 value will be used for quality estimation here. As a side note, values for MTF at 10% modulation contrast are obtained as piecewise linear interpolation between adjacent measurements. Alternatively, a polynomial- or Gaussian-fit may be applied but it can be shown that both fit-approaches use all measurements as input and if these are falsified (e.g. due to inhomogeneous target illumination or sharpening artefacts) approximated MTF10 values may vary substantially.

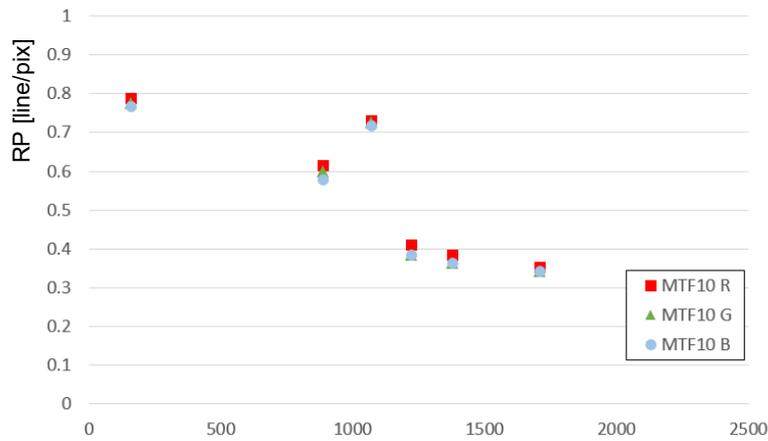
The Siemens star images have been taken during regular mapping flights. The target was fixed on ground while the drone was flying the planned test. Placing the target in a multiple-overlap area, due to several overlapping flight lines, the star will be imaged multiple times from different camera stations. Thus, the star will be located at different positions in image plane. Each imaged star allows for the estimation of the resolving power at this distinct position in image plane. The Fig. 1 shows three samples of imaged Siemens stars taken by the three different cameras. All samples were selected in a way that the Siemens star is almost in the centre of the image. This should give an estimation of best possible resolving power, which is supposed to decrease from centre to image corners. In all cases jpg-images have been analysed, as this is the mostly preferred image data format for most UAV scenarios, even though the possible influence of compression effects is well-known.

As shown from Fig. 1 the Siemens star has a special design. It consists of 32 segments (16 black and 16 white), 4 normalization patterns (2 black and 2 white), 4 rotor shaped external markers, 4 circular shaped external markers and a black square to extract the star's 2D-orientation in image plane. The four normalization patterns (each 2 black and white circles for redundancy if significant noise is present) are used to set maximum and minimum values during determination of spatial resolution. This is not only necessary but very helpful if the imaged Siemens star is too small. The 4 rotor shaped and / or circular shaped external markers are used to find the correct star-centre position which is crucial for correct results. Furthermore, these markers are used to obtain the homographic transformation matrix to correct test target inclination (especially in image corners) if required.

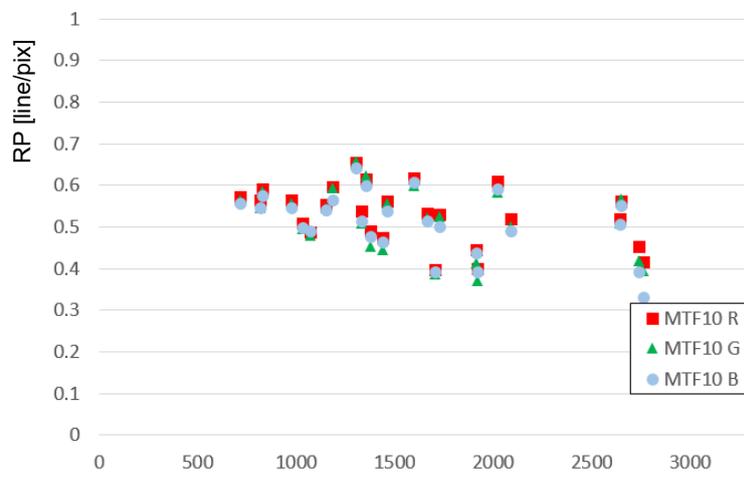
As the GSD is different for each flight, the size of the star in the image is different. The specific image from Phantom 4 shown here, is taken from 28 m flying height above the Siemens star pattern, corresponding to a local GSD of 1.20 cm. Different to this, the selected Phantom 4 RTK image has been captured from 25 m above Siemens star. This is the minimal flying height above ground, which can be realized in Phantom 4 RTK pre-planned automatic mission flights. Due to the smaller sensor pixel size the corresponding GSD is 6.9 mm, which is the minimal GSD due to the minimal flying height limits above. When flown in manual mode no such limits are available. Interesting to note, that the Phantom 4 does not have such limitation. The shown Phase One iXM-100 image was taken from 47 m above Siemens star, corresponding to a GSD of 5.0 mm. The results from MTF analysis are also shown in this figure. The MTF10 value given in line/pix refers to the image space. In order to transform this measure to the object space value the reciprocal of MTF10 is multiplied with the nominal GSD. The corresponding GRD values are given in the figure. If one compares the GRD to GSD, the factors are 128%, 178% and 130% for Phantom 4, Phantom 4 RTK and iXM100-RS respectively. Obviously Phantom 4 provides the same relative resolution as iXM100-RS for the centre image region, but the different absolute numbers and the fact, that Phantom 4 images has be sharpened in internal DJI image processing, has to be noticed.



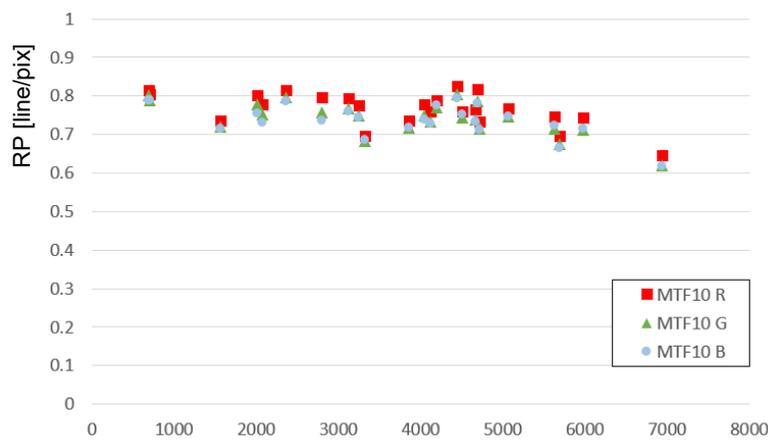
Fig. 1: Siemens star image and analysis from Phantom 4 (left), Phantom 4 RTK (middle) and Phase One iXM100-RS (right). Each image sample is sized 300 x 300 pix.



Phantom 4 – MTF10



Phantom 4 RTK – MTF10



iXM100-RS – MTF10

Fig. 2: Resolving power (RP) of tested cameras give for the three colour channels. MTF10 values [line/pix] are plotted against the radial distance from image centre to image corners given [pix]. Notice the different sensor formats.

A comparison of each camera regarding spatial resolution in relation to distance of image centre is given in Fig. 2. The iXM100-RS with its huge sensor format and 100 megapixels shows very low variation in resolving power over the complete image space (Fig. 2, bottom). Imaging performance of both DJI systems (Fig. 2, top and centre) is fairly homogenous but the loss of spatial resolution towards image corners is significantly higher in comparison to the iXM100-RS sensor-optic combination. Unfortunately, only six images are available for the Phantom 4 resolution determination.

Phantom 4 RTK images have been used without DJI internal pre-processing, i.e. without sharpening and pre-distortion corrections as implemented in the DJI raw image processing. This causes quite strong distortions in the image borders and especially those Siemens star targets imaged there show strong ellipsoidal deformation. The homographic transformation has been applied to correct the target's distortion. This transformation however is not bijective (HARTLEY & ZISSERMAN 2004) and thus, the transformed measurement coordinate system is not absolutely accurate and obtained values contain noise. While the overall measurement accuracy of spatial resolution is in the range between 5 % – 7 % (iXM100-RS), it rises to approximately 10 % - 12 % (Phantom 4 RTK) when homographic transformation is applied.

Two observations have to be mentioned at this point. First, exposure-time usually is set to obtain enough information (in form of signal strength) even in darker, shadowed areas. The Siemens star target with its (ideal) black and white segments then often is over-exposed which influences the measurement of spatial resolution. The Siemens stars in Phantom 4 RTK images show this effect, as well as some of the Phantom 4 images. In addition, images of Phantom 4 have been sharpened in the internal DJI image pre-processing. This is the reason, why spatial resolution in Phantom 4 RTK in general is lower compared to the Phantom 4 imagery.

Second, if the test target is too small in the corresponding images the selected target design is crucial to normalize the data during measurement process. Alternatively, the used target and corresponding diameter needs to satisfy the requirement that each segment (black or white) of the outermost ring (low spatial frequency) is covered of at least three pixel. This then ensures unmingled values for black and white. Both issues can be solved when planning and executing a dedicated test flight for spatial resolution determination. Especially the optimal setting of exposure parameter to get optimal Siemens star images is not feasible in many practical scenarios. Regarding the size of the Siemens star in the images, the resolution target diameter itself can be increased, but this limits the mobility and flexible use of the target.

3.2 3D object point determination

The performance of 3D object point determination is derived from extended aerial triangulation, which includes camera self-calibration and (optional) directly measured GNSS/inertial exterior orientation elements. Data is captured in a test site, which provides significant number of signalized and highly precise coordinated 3D object points. These points are considered as control and/or check points. The test field for the following experiments was the ship lock at Hessigheim, Neckar river, about 30 km north of Stuttgart, which is sized about $200 \times 70 \text{ m}^2$. The site was established as part of a research and development project initiated by the German Federal Institute of Hydrology (BfG) in Koblenz in partnership with the Office of Development of Neckar River Heidelberg

(ANH). The 3D monitoring of the ship lock infrastructure and its surrounding area is the underlying application. Different to traditional monitoring, where classical terrestrial survey methods are applied, the 3D structure should be precisely mapped from UAV data, in order to move from point to surface capturing. This is why the test site was established with extremely accurate 3D signalized object points. Checkerboard discs with approximately 30 cm diameter are used for signalization. Their 3D point accuracy is within 1 mm (std.dev.). The Fig. 3 depicts the test area and the distribution of signalized points.



Fig. 3: Test area of ship lock Hessigheim/Neckar with distribution of signalized points. The image on the right shows one of the signalized points. All points are fixed to the ground except three, which are mounted on top of approximately 140 cm high concrete pillars (red squares).

Several tests have been flown in this site, only a small sub-set will be presented in the following, the main flight parameters are given in Tab. 2. Only three control points have been introduced, all remaining points are for check only, in order to have efficient operational scenario. Regular blocks are compared to cross-pattern flight scenarios. In all cases the additional self-calibration parameters are estimated from zero initial values as part of the bundle adjustment. The corresponding accuracy as obtained from 15-17 check points (not all of the points were available in all the different flight campaigns) are given in Tab. 3. The accuracy is given in absolute numbers and is also referred to individual GSD of flight test.

In flight test #1, the result is only GCP-based, as the on-board GNSS of Phantom 4 only provides navigation grade initial camera exteriors. The horizontal RMS is consistent and all within one pixel, while the vertical RMS reached 2.2 pixel. The distribution of Z error of check points shows a ‘bowling effect’ – the error in the centre of the survey area was positive while it became negative at the border. This is the typical situation for minimal GCP constellations with no perspective centre coordinate measurements and additional self-calibration. Using a denser GCP distribution, the solution stabilizes. In an extra test, based on a dense control point distribution using eleven GCPs, the RMS Z dropped to 0.7 pixel and the RMS 3D is 1.17 pixel.

Tab. 2: Tested flight constellations, all versions using 3 GCPs only.

#	System	Block design	GSD (cm)	Number of images	Overlaps & Remarks
1	Phantom 4	Regular 4 NS lines	1.29	64	long-track 70%, cross-track 76% GCP-based only
2	Phantom 4 RTK	Regular 8 EW lines	1.29	82	long-track 80%, cross-track 60% with RTK-GNSS
3	Phantom 4 RTK	Regular 6 NS lines	0.68	298	long-track 80%, cross-track 60% with RTK-GNSS
4	Phantom 4 RTK – CalibFlight	Calibration flight scenario 6 NS lines / 8 EW lines	0.68 NS / 1.29 EW	388	Overlaps as in test #2 & #3 with RTK-GNSS
5	iXM100-RS – Cal- ibFlight	Calibration flight scenario 4 NS lines / 11 EW lines	0.4 NS / 0.5 EW	148	NS: long 70%, across 50% EW: long 77%, across 57% with post-processed GNSS/inertial positioning from APX-15el UAV board

Tab. 3: 3D object point performance from check point differences, all versions using 3 GCPs only.

#	System	RMS X [cm] / [pix]	RMS Y [cm] [cm] / [pix]	RMS Z [cm] [cm] / [pix]	RMS 3D [cm] / [pix]
1	Phantom 4	0.93 / 0.7	0.92 / 0.7	2.90 / 2.2	3.19 / 2.4
2	Phantom 4 RTK	0.62 / 0.4	0.93 / 0.7	4.46 / 3.4	4.59 / 3.5
3	Phantom 4 RTK	0.31 / 0.4	0.50 / 0.7	0.80 / 1.1	1.00 / 1.4
4	Phantom 4 RTK – CalibFlight	1.11 / 1.6	1.70 / 2.5	1.81 / 2.6	2.73 / 4.0
5	iXM100-RS – Cal- ibFlight	0.14 / 0.3	0.23 / 0.5	0.19 / 0.4	0.33 / 0.8

The processing of Phantom 4 RTK data in test #2 and #3 takes the initial exteriors from the RTK-GNSS. The two flights are with different strip direction and GSD. The flight #2 is ‘cross-strip’, i.e. with strips in east-west direction, perpendicular to the long side of the rectangular survey area. It has eight strips and nine images per strip; the GSD is equivalent to the Phantom 4 flight test #1. In test #3 the flight direction is in north-south axis, similar to test #1, but it has six strips and 47 images per strip, because of the smaller GSD. The ‘cross-strip’ pattern of #2 weakens the connection of image block, as more images were connected side-by-side and the side-lap is lower than

the front-lap. This mainly influences the vertical RMS, which is significantly larger in test #2 compared to test #3. However, flight test #3 shows consistent results in all axes. In terms of absolute numbers, the accuracy obtained here is best, as this test provides the smallest GSD. Relating the accuracy with respect to the pixel size, horizontal accuracy is comparable to the previous flight tests #2 and #3, still the vertical accuracy RMS is significantly better around 1.1 pixel. This proves the positive influence of high-quality GNSS perspective centre coordinates in combination with a sufficient image block geometry.

It is worth noting that the RMS 3D of flight test #4 is worse than #3 and #2. The two flying heights in cross-pattern should have strengthened the image block geometry, and additional RTK-GNSS observations should have increased the precision as well. However, both horizontal and vertical RMS do not increase. There may exist systematic error in the GNSS observation that disagree with the image block geometry. Additionally, when introducing ten more GCPs, the overall RMS decreased to 3.4 pixel, as more GCPs fixed the image block in a stronger way. Still, this result does not satisfy expectations and needs further investigations.

Flight test #5 with the Phase One iXM100-RS camera and post-processed GNSS/inertial positioning is the best in absolute numbers as well as in pixel. This in a way was expected, as the much larger sensor format allows less images to get the same project area captured with similar overlap, even though the GSD is again smaller than the smallest Phantom 4 RTK GSD. Interesting to see that, comparing test #3 with #5, the Phantom 4 RTK can achieve similar horizontal accuracy. Still, vertical accuracy exceeds the result in #3 by a factor of two.

4 Summary and Conclusions

This paper exemplarily highlights two methods to evaluate the quality of UAV-based or other airborne sensor image data. The methods start with the resolution testing of images using defined resolution charts, the overall geometric performance is then derived from test site with sufficient number of highly accurate check point information. The first step is using a software, which comes as part of an official DIN standard soon, and gives objective, comparable and transferable results as the testing scenario is clearly defined. Results have shown that the image pre-processing has significant influence on the estimated spatial resolution, in addition to the camera-lens combination itself. It therefore has to be mentioned, in which processing level the resolution testing has been done. In principle the resolution of derived products like orthoimage can also be analysed. The typical sharpening of raw imagery, implemented in the manufacturers' internal raw data processing has to be mentioned here. In our example the spatial resolution of Phantom 4 sharpened images was better than the Phantom 4 RTK unsharpened images, which has to be expected. In general, sharper images contain sharper details and more information, this is why many images are internally sharpened before the image is provided to the user. However, it has to be investigated how this sharpening influences the subsequent tie point transfer, especially when it comes to later dense image matching.

Additional environmental conditions, like platform vibrations, motion blur or atmospheric conditions also play a role. These are mission dependent, nevertheless, when resolution testing is done within the mission itself, this gives a quality measure from the same data used for the application.

The advantage of this method is the flexibility as the defined resolution target can be provided in each project area with minimal effort.

The testing of 3D object point quality from test sites is well established. Still, as discussed earlier, it is quite complex to define and install test sites, which provide very similar test environments. Special care has to be laid on the high quality of reference object data. In addition, the area should provide enough texture, which is helpful for the later tie point transfer. This is the essential part in bundle adjustment. From this point, the selected ship lock area is not optimal, even though it provides numerous reference points, as the surrounding water areas prevent reliable tie point transfer, especially when camera only has smaller image formats. Additional depth information in the test field support the in-situ camera calibration. This is easier to reach for UAV-applications as the factor between flying height above mean terrain level and height variations in object space is much smaller compared to classical large area airborne mapping. Assuming a mean flying height of 40 m, a height variation of only 4 m already corresponds to 10 % of the flying height. This is easy to reach in every test area. Flying in 2000 m above ground, the same 10 % height variation already corresponds to 200 m height difference, which already is a significant height jump.

It was intention of the paper to compare the performance of standard, more low-cost UAV-camera systems to the high-end camera based UAVs. As shown before, some differences are visible from the empirical check point accuracy. Although DJI systems are well-built and easy-to-use, some drawbacks still exist compared to professional unmanned aerial photogrammetric systems. First of all, the systems are close-sourced, which means that users cannot get access to the raw observations and logs to perform further processing to increase accuracy. For example, users are not allowed to get access to the raw image data but only JPG of Phantom 4 RTK. The JPG image may have already been ‘corrected’ by the firmware so that it does not fit the camera distortion model (PEPPA et al. 2019). Other limitations by the manufacturer may include the minimum flying height and the adjustable camera parameters. Users usually have to set the exposure settings empirically and adjust them during the flight to achieve an optimal image quality. However, the aperture of Phantom 4 RTK camera is not adjustable during the image capture, which may introduce additional problems. On the other hand, both cameras have large FOV and significant distortion. This may cause the actual image overlap being smaller than the planned one. Moreover, these cameras require careful in-situ calibration (CRAMER et al. 2017), despite that DJI provides the lab camera calibration result of each Phantom 4 RTK. The interaction of camera calibration parameters and directly measured GNSS perspective centre coordinates with the tie point observations and ground reference data is essential. Since DJI systems are compact, the possibility of hardware upgrading is zero — users cannot change any of the components of Phantom 4 or Phantom 4 RTK.

It was already expected that Phase One iXM100-RS performs better in several aspects. What should be highlighted is the consistent resolution of the tested camera-lens combination, with almost no decrease when it comes to the image borders. The true GSD is close to the nominal GSD throughout the whole image plane. Larger image format helps to cover given areas with less images per strip which gives a stronger geometry, especially for low number of GCPs. Even though the tested camera was lab-calibrated right after the empirical test, using these lab calibration parameters did not deliver highest object point accuracy. Therefore, similar to the DJI systems, the in-situ calibration is mandatory. In principle, an additional lab calibration is not of need at all, but

might be used to provide good initial values for the in-situ calibration. One of the biggest advantages is the accessibility of raw camera data. With the given software, the user has full access to the pre-processing of images. In general, the overall flexibility is much higher compared to the DJI integrated systems. The iXM offers different lens options and provides defined interfaces for other platforms and combinations with other sensors, like aiding navigation sensors. Still, in order to realize such multiple system integrations, the integration and fine-tuning of all the sensor-platform interactions is under the full control but also responsibility of the user, which requests for some hardware experience.

In the end, it is always up to the user to select between low-cost versus high-end, compact versus mid-format, close-sourced versus flexible UAV solutions according to the specific application. The proper selection of the UAV imaging system in high-resolution mapping highly relies on the quality and accuracy that users want to achieve. The in-field image resolution testing can help users to better understand the imaging quality, which is as necessary as the in-situ camera calibration. The current version of software to estimate the spatial resolution of imagery in a defined way according to DIN 18740-8 together with the appropriate template of the Siemens star can be made available to interested readers. Please contact the main author of this paper.

5 Acknowledgments

Special thanks has to be expressed to the German Federal Institute of Hydrology (BfG) in Koblenz financing the research project on area wide 3D UAV-based monitoring, which makes the test site Hessigheim possible. The Office of Development of Neckar River Heidelberg (ANH) kindly supports the terrestrial measurements and all on-site work. Applanix kindly provided the APX-15el UAV GNSS/inertial board for the Phase One camera flights and delivered the processed trajectory information for Phase One iXM100-RS flights. All this support is kindly acknowledged!

6 Bibliography

- CRAMER, M., 2010: The DGPF-Test on Digital Airborne Camera Evaluation – Overview and Test Design. *Photogrammetrie – Fernerkundung – Geoinformation (PFG)*, **2010(2)**, 75-84.
- CRAMER, M., PRZYBILLA, H.-J. & ZURHORST, A., 2017: UAV Cameras: Overview and Geometric Calibration Benchmark. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* **42(2/W6)**, 85-92. doi: 10.5194/isprs-archives-XLII-2-W6-85-2017.
- DJI, 2019a: DJI GS Pro, Mission-Critical Flight Simplified. DJI Innovation, <https://www.dji.com/uk/ground-station-pro>.
- DJI, 2019b: Phantom 4 RTK Specs. DJI Innovation. <https://www.dji.com/uk/phantom-4-rtk/info#specs>.
- DJI, 2019c. Phantom 4 Specs. DJI Innovation, <https://www.dji.com/uk/phantom-4/info#specs>.
- HARTLEY, R. & ZISSERMAN, A., 2004: Multiple View Geometry in Computer Vision. Second Edition. 2nd edition, Cambridge University Press, ISBN: 0521540518.
- MEIBNER, H., CRAMER, M., & REULKE, R., 2018: Towards standardized evaluation of image quality of airborne camera systems. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* **42(1)**, 295-300, doi: 10.5194/isprs-archives-XLII-1-295-2018.

- MEIBNER, H., CRAMER, M. & REULKE, R., 2019: Evaluation of structures and methods for resolution determination of remote sensing sensors. *Advances in Image and Video Technology, Proceedings of 9th Pacific-Rim Symposium on Image and Video Technology*, November 18-22, Sydney, Australia, Springer, Lecture Notes in Computer Science (LNCS) **11994**, 59-69.
- PEPPA, M. V., HALL, J., GOODYEAR, J. & MILLS, J. P., 2019: Photogrammetric Assessment and Comparison of DJI Phantom 4 Pro and Phantom 4 RTK Small Unmanned Aircraft Systems. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* **42**(2/W13), 503-509, doi: 10.5194/is-prs-archives-XLII-2-W13-503-2019.
- Phase One (2020): Phase One iXM-100 specifications. https://industrial.phaseone.com/iXM_Camera_Series.aspx.
- USGS, 2020: US Geological Survey Online Test Site Catalog. https://calval.cr.usgs.gov/apps/test_sites_catalog.
- WIESER, C., 2018: Personal correspondence, 26.6.2018.