TOWARDS STANDARDIZED EVALUATION OF IMAGE QUALITY FOR AIRBORNE CAMERA SYSTEMS

H. Meißner[a,*], M. Cramer[b], R. Reulke[a]

[a] Institute of Optical Sensor Systems, German Aerospace Center, 12489 Berlin, Germany - (henry.meissner, ralf.reulke)@dlr.de
[b] Institute for Photogrammetry (ifp), University of Stuttgart, 70174 Stuttgart, Germany - michael.cramer@ifp.uni-stuttgart.de

Commission TC I

KEY WORDS: Resolving Power, Image Quality, Siemens Star, Bar Targets, De-Mosaicing, Standardized Process

ABSTRACT:

Digital airborne camera systems and their high geometric resolution demand for new algorithms and procedures of image data analysis and interpretation. Parameters describing image quality are necessary for various fields of application (e.g. sensor and mission design, sensor comparison, algorithm development, in-orbit-behaviour of instruments). The effective sensor resolution is one important parameter which comprehensively estimates the optical quality of a given imaging sensor-lens combination. Although determination of resolving power is a well-studied field of research, there are still some scientific questions to be answered when it comes to a standardized (eventually absolute) determination. This is also research object of a committee of the "German Institute for Standardization" and the given contribution outlines the current state of investigation concerning effective resolving power for airborne camera systems. Therefore an approach using signal processing techniques to calculate the effective image resolution will be described. The open scientific issues will be introduced, explained and answered to some extend.

1. INTRODUCTION

The effective geometric resolution is an essential parameter of imaging systems as it defines the detail of information in any image taken by this camera configuration. This is why resolution estimation is of such importance to quantify the potential of this camera.

Resolution is defined in different ways. A common and well-known criterion was introduced by Rayleigh (Born and Wolf, 1999). There, the response of an imaging system when illuminated with a point light source is defined respectively approximated by a sine cardinal function. Further, Rayleigh defines a resolution limit as the minimal distance between two sources where they are still discriminable. Using the definition that point light sources are approximated as sine cardinal functions the resolution limit is reached if the first maximum position of one function is identical to the first minimum of the other function. In airborne scenarios, on the other hand, image resolution is calculated as pixel-size multiplied by flight altitude divided by focal distance. In almost all cases these calculated values differ from de facto resolution due to numerous reasons (e.g. used lens and aperture, motion during exposure). The resulting, measurable resolution therefore can be defined as effective resolution or resolving power.

Although determination of resolving power is a well-studied field of research, there are still some scientific questions to be answered when it comes to a standardized (eventually absolute) determination. This is also research object of a committee of the "German Institute for Standardization" and the given contribution outlines the current state of investigation concerning effective resolving power for airborne camera systems. Therefore basic radiometric characteristics are discussed in section 2 followed by describing methods for resolving power determination (bar targets and signal processing technique) in section 3.

The open scientific issues will be introduced, explained and answered to some extend in section 4 (e.g. Siemens star center accuracy, used De-mosaicing method).

In the following this paper attends to visualize the influence of resolving power on the later (classical) photogrammetric workflow. The camera is the primary sensor for photogrammetric processing as it delivers the observations of which the later products (tie points for bundle adjustment, sparse and dense 3D point clouds) are being derived. In general one may expect more reliable tie point features, better geometric performance of the bundle block and subsequent dense image matching for sharper input images.

To support this, a real UAV-based data set from DLRs MACS Micro camera system (Kraft et al., 2016) was acquired over a photogrammetric test site. A bundle block adjustment was calculated for all 632 images in cross-flight configuration (overlap: 80% in flight direction, 60% across flight direction). A high precise geodetic network with 45 signalized points has been applied out of which 5 points served as control points and the remaining 40 points were used as check points. In a second run the original 632 images have been blurred through Gaussian blur filter and the bundle block was recalculated. Related experiment description and obtained results are given in section 5.

2. RADIOMETRIC CHARACTERISTICS

Radiometric quality of a sensor system is causally determined by multiple factors. Light rays which are being reflected by an object and detected by a camera sensor partially traverse the atmosphere and lose some of their energy due to diffusion and absorption. In airborne applications this part could be considered very small and won't be discussed further here.

Next the light passes a (complex) lens system where an aperture is integrated and limits the effective solid angles for every ray. As

---

a consequence the aperture directly affects the amount of light which in turn determines the amount of photons that reach the sensor plane and contribute to the imaging process. The smaller the aperture is chosen the more diffraction of light limits a sharp optical imaging. On the other hand, if the aperture is chosen too large spherical and chromatic aberrations gain influence. The amount of photons passing through the lens system and reaching the sensor at a distinct time frame directly influences the exposure time needed to create an equivalent sensor signal. In aerial photogrammetry the exposure time however affects a sharp optical imaging in terms of motion blur that is a result of the systems change of location / movement whilst the sensor is exposed. This change of location (translative or rotative) can be compensated with stabilized mounts and several aerial camera systems offer some techniques. But many systems (especially for UAV applications) retain as additional parts increase the total weight limiting flight endurance and operation time. Still, situation is changing, when looking on the video recording. Here quite sophisticated stabilized mounts are available to minimize blur in images. The influence of image blur, comparing imaging in static (laboratory) and dynamic (operating) conditions has been shown in Kraft et al. (2016b).

Another interfering aspect is the gain of shading (or inverse the luminous intensity decrease) starting from the principle point to image corners. This effect is often described as vignetting and is caused by the lens system itself and by the integrated aperture. The vignetting can be measured and corrected as an image processing step whilst determine the Photo Response Non-Uniformity (PRNU) (Kraft et al., 2016a). After the light rays passed the lens system they hit the sensor surface. That is the part of the camera system that creates a digital interpretable signal directly depending on the amount of collected photons during the exposure time window. The quality of that signal is affected by several electronic components (e.g. sensor read-out electronic, analog-digital converter). A measure of this quality is the signal noise ratio (SNR). The SNR also is characterized by a) the ambient noise level that unavoidably occurs when a semi-conductor is connected to its supply voltage and b) to the photo-effective area of each sensor element (pixel). The larger the effective area the more photons contribute to the signal assuming identical time frames and therefore increase the signal. The electronic ambient noise can be determined pixel by pixel as part of the Dark Signal Non-Uniformity (DSNU) (Kraft et al., 2016a).

During post-processing images of sensors using micro filter patterns (e.g. Bayer pattern) for color acquisition need to be converted from fractured color array to true color array. The chosen de-mosaicing (or de-bayering) method directly affects the effective resolving power (Meißner et al., 2017).

3. DETERMINATION OF RESOLVING POWER

Sharpness as an image property is characterized by the modulation transfer function (MTF) which is the spatial frequency response of an imaging system to a given illumination. "High spatial frequencies correspond to fine image detail. The more extended the response, the finer the detail - the sharper the image." (Mix, 2005).

The effective image resolution or resolving power of an imaging device can be determined in different ways. A classic approach is the use of defined test charts (e.g. USAF resolution test chart with groups of bars). There, the (subjectively) identified image
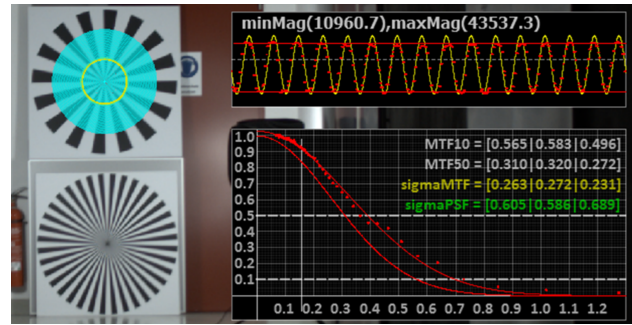

Figure 1. Designated test pattern Siemens star (left), radial modulation analysis (top-right), resulting MTF and PSF (bottom-right).

resolution corresponds to that distance where the smallest group is still discriminable. This is very similar to the Rayleigh criterion (Born and Wolf, 1999) that defines the minimum distance between two point sources in order to be resolved by an imaging system. Besides subjective components included in this process the function values (resolving power) are discrete instead of continuous, depending on resolution steps between groups of bars.

To reduce subjective influence with bar charts during the determination process and to convert discrete function values to continuous some approaches use signal processing techniques to calculate the effective image resolution. The method described by Reulke et al. (2004, 2006) is one of the latter approaches. There, the contrast transfer function (CTF), modulation transfer function (MTF) and subsequently point spread function (PSF) is calculated for images with a designated test pattern (e.g. Siemens star). According to the above mentioned approaches the smallest recognizable detail or "the resolution limit is reached if the distance between two points leads to a certain contrast in image intensity between the two maxima." Using a priori knowledge of the original scene (well-known Siemens star target) CTF, MTF and PSF are approximated by a Gaussian shape function (see Figure 1).

Coordinate axis X for CTF and MTF is the spatial frequency f (eq. 1) and is calculated as quotient of target frequency f_s divided by current scan radius r multiplied by π. Target frequency f_s is constant and equivalent to the number of black-white segments of the well-known Siemens star. Related (initially discrete) values for contrast transfer function $C_d(f)$ are derived using intensity maximum I_{max} and minimum I_{min} for every scanned circle (eq. 2). Simultaneously the function value is normalized to contrast level C_0 at spatial frequency equal to 0 (infinite radius). Continuous function values C are derived by fitting a Gaussian function into discrete input data (eq. 3). According to Coltman (1954) the obtained CTF describes the system response to a square wave input while MTF is the system response to a sine wave input. The proposed solution is a normalization with $\frac{\pi}{4}$ followed by series expansion using odd frequency multiples (eq. 4). MTF describes the effective resolving power in frequency domain while PSF is the image domain equivalent. For this reason both functions are linked directly by fourier transform (eq. 5).

$$f = \frac{f_s}{\pi r} \tag{1}$$

$$C_d(f) = \frac{I_{max}(f) - I_{min}(f)}{I_{max}(f) + I_{min}(f)} * \frac{1}{C_0} \tag{2}$$

$$C = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \qquad (3)$$

$$MTF(f) = \frac{\pi}{4} \left[C(f) + \frac{C(3f)}{3} + \frac{C(5f)}{5} + \ldots \right] \qquad (4)$$

$$PSF \quad \circ\!\!-\!\!\bullet \quad MTF \qquad (5)$$

There are several criteria for the resolving power of cameras. The parameter σ (standard deviation) of the PSF (Gaussian-shape function) is one criterion. It directly relates 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 this frequency is the approximation for the number of the smallest line per pixel. The width of PSF at half the height of the maximum is another criterion (full width half maximum - FWHM).

4. STANDARDIZATION ASPECTS

Standardized procedures in terms of repeatable and comparable measurements for resolving power is research object of a committee of the "German Institute for Standardization". When looking down the path to define a standard for the determination process there are still some scientific questions to be answered. This section explains these open questions, categorizes the issue's origin (e.g. algorithmics or measurement arrangement) and provides answers to some extend.

4.1 Comparison with slanted edge approach and bar target method

Given the variety of approaches and techniques to determine the effective resolving power of airborne images it is consistently necessary to compare their respective results and answer the question if used techniques do or do not perform equivalently and what are reasons for particular observations. Contemplated approaches are: bar targets (e.g. USAF resolution test chart with groups of bars), slanted-edge method (Burns, 2000) and the approach described in section 3.

4.2 Influence of test target inclination

There are two basic assumptions regarding the used approach for resolving power determination described in section 3. First premise is best as possible center position estimation, explained in sub-section 4.3. The second assumption is an orthogonal image acquisition configuration (see Figure 2, left). According to equation (1) the test pattern imaging is scanned in concentric circles. If given configuration violates this premise the measurement unavoidably will be falsified (see Figure 2, right). In airborne scenarios sensor inclination can be reduced (e.g. stabilized mounts) but not entirely excluded.

There are three possible solutions at this point. In terms of defining a standardized process it could be appropriate to set a maximum inclination angle (e.g. ± 5 degree). While this would preserve quality of the measurement it is assumingly difficult to operate in airborne scenarios. Another solution is to determine the

inclination angle (external measurement or included markers) and to rectify the image accordingly. However, this approach demands for image re-sampling and affects the measurement depending on chosen interpolation method. As an alternative it could be appropriate to transform the sampling pattern to image space with derived homography parameters (Hartley and Zisserman, 2004) and conduct all scans using the new (local) coordinates system.
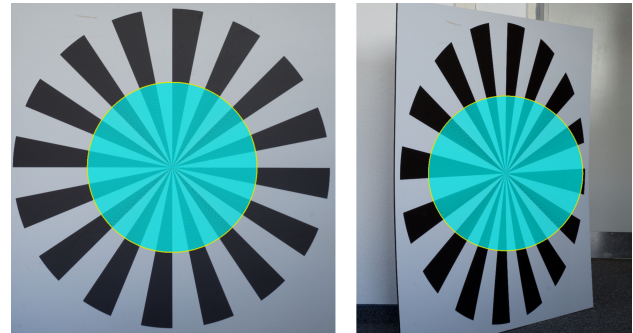

Figure 2. Orthogonal image acquisition configuration (left), inclined image acquisition configuration (right)

4.3 Siemens star center position

Within the process of MTF / PSF estimation the centre of the Siemens star resolution target has to be identified. Most often this is done by manual measurement. The correctness of this manual centering will influence the later MTF and PSF measurement. In order to clarify this, the corresponding MTF and PSF has been calculated from slightly different star centre measurements, varying in a given pixel neighbourhood window (e.g. $\pm 6px$) and subpixel step size. The results shown in Figure 3 indicate that there are negligible variations in PSF only in very close vicinity (e.g. $\pm 1px$) of the centre but there is significant decrease in PSF estimation with increasing distance from the true position of Siemens star centre.
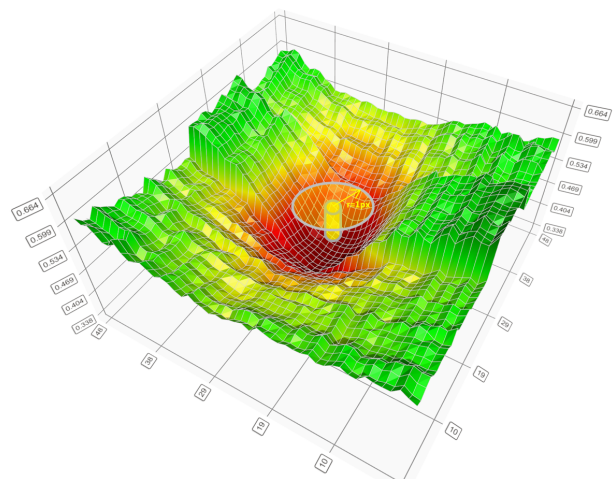

Figure 3. Plot of σ PSF in relation to Siemens star center position

Conclusion at this point, besides correct centre position determination (using external markers or geodetic survey) a derived confidence level in close surrounding of the initial centre position (see Figure 3) can be helpful to validate the quality.

4.4 Normalization of contrast magnitude

The normalization factor C_0 according to equation (2) is the intensity ratio at spatial frequency f equal to zero (infinite radius r). Based on the fact that an image does not yield information for an infinite radius it is common to use intensity ratio provided by the largest available circle (see Figure 1). However, if the designated test pattern is too small (or the imaging distance too large) it can not be guaranteed that the outermost circle's intensity ratio is equivalent to maximum intensity ratio of the lens-sensor combination to be measured.

This leads to the conclusion that either minimum size of the Siemens star has to be defined (e.g. in relation to an approximate of effective ground resolution) or sufficiently large normalization patterns (black-white) need to be included.

4.5 Mathematical PSF and MTF approximation and transition

The two-dimensional PSF is approximated with a Gaussian-bell function under the assumption of a rotation symmetrical lens-sensor PSF. Hence, the one-dimensional PSF is derived as cross-section of the Gaussian-bell. The one dimensional PSF as Gaussian-shape function, however, requires the shape of MTF to be a Gaussian-shape function, too (see eq. 5).

This leads to two open issues.

First, is the approach approximating CTF, MTF and subsequently PSF according to equation (3),(4) and (5) with Gaussian-shape functions applicable? Or would it be more precise calculating discrete inverse Fourier transformation using discrete non-approximated input data? A further option is to use continuous polynomial approximated input data. But both in turn would eliminate σ of PSF as quality parameter and raise the question for an alternative criterion.

Second, is the approach approximating PSF with a 2D-Gaussian bell function accurate enough? Or would it be better to measure a 2D-MTF (then closely related to slanted-edge methods) and expand equation (5) to a 2D-Fourier transformation?

4.6 De-Mosaicing method

The predominant majority of colour cameras use micro filter arrays (one filter evaporated onto every single pixel, so-called Bayer-pattern) to capture colour information. The array usually is arranged with alternating colour filters (e.g. green – red) for a single sensor line and then a complementary arrangement in the following sensor line (e.g. blue – green), etc. During the colour reconstruction process the missing information is determined by interpolating between the neighbourhood values. Hence, demosaicing or debayering is another factor influencing image resolution and the corresponding measurement.

The simplest way to restore the missing information is to interpolate each channel separately using neighbouring values. bi-linear interpolation is the most commonly used mode, but it would be possible to use nearest neighbour or bi-cubic interpolation instead. This method is efficient and straight forward to implement, but images will exhibit colour fringing at edges.

VNG (Various Number of Gradients) reduces colour fringing by using edge detection (Chang et al., 1999). A set of 8 gradients is
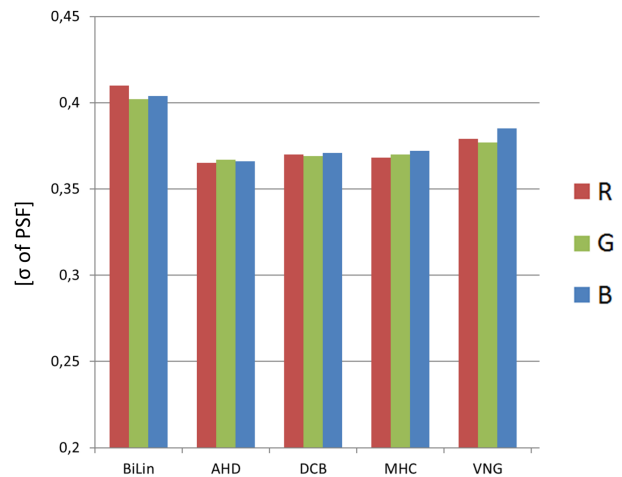

Figure 4. Comparison of different de-mosaicing methods for (R)ed, (G)reen and (B)lue channel

calculated for each pixel by comparing values in the 5x5 neighbourhood. The gradient is calculated by summing up the absolute difference of like-coloured pixels. A threshold is used to determine if the pixel lies on a smooth area and averaging can be used to approximate the missing values, or if the pixel lies on a steep gradient, where it is better to use one of two neighbouring values.

The Adaptive Homogeneity-Directed (AHD) demosaicing approach (Hirakawa and Parks, 2005) identifies three different classes of artifacts: misguidance colour artifacts, interpolation colour artifacts and aliasing. They set out to minimize aliasing by using filter-bank techniques. Misguidance colour artifacts, which arise when the direction of interpolation is erroneously selected (interpolation along an object boundary is preferable to interpolation across the boundary), are addressed through a nonlinear iterative process. The image is interpolated twice - once vertically f_h and once horizontally f_v. The final output f is calculated by combing f_h and f_v based on a homogeneity matrix H_f which aims to minimize colour artifacts.

MHC is a simpler algorithm than VNG or AHD, it has higher performance than such nonlinear algorithms and doesn't suffer from artefacts due to (sometimes wrong) assumptions about gradients in the image (Malvar et al., 2004). It works linearly in a 5x5 neighbourhood by first filling in values using bilinear interpolation. It tries to analyse local luminance changes by comparing the actual value at the current pixel position to the value arrived at by interpolating same-colour neighbours. It then factors a corresponding gain term when calculating the other two colour values at the same position.

Finally, the iterative DCB algorithm is included. It is largely undocumented, but an open source implementation which performs well is available. (Góźdź, 2009).

Figure 4 provides a comparison of all described de-mosaicing methods using the resolution determination approach described in section 3. All approaches except bi-linear interpolation differ only by $\pm 3\%$. Further investigations will have to clarify if these differences between de-mosaicing methods remain stable for other optic-lens combinations.

5. SIMULATED IMPACT OF IMAGE SHARPNESS ON BUNDLE ADJUSTMENT

As stated in section 1 the camera is primary sensor for further photogrammetric processing as it delivers the observations of which the later products (tie points for bundle adjustment, sparse and dense 3D point clouds) are being derived. To support this, a real UAV-based data set from DLRs MACS Micro camera system (Kraft et al., 2016) was acquired over a photogrammetric test site.

5.1 Simulation Setup

A bundle block adjustment was calculated for all 632 images in cross-flight configuration (overlap: 80% in flight direction, 60% across flight direction, see Figure 5).
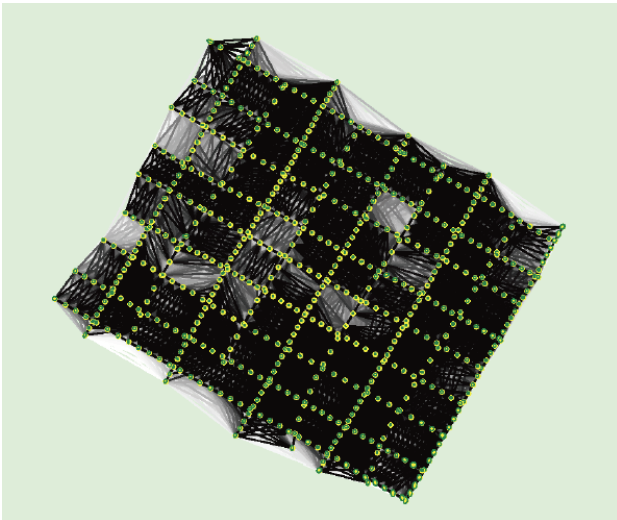

Figure 5. Block configuration

A high precise geodetic network with 45 signalized points has been applied out of which 5 points served as control points and the remaining 40 points were used as check points. In a second run the original 632 images have been blurred through Gaussian blur filter (see Figure 6) and the bundle block was recalculated.

Figure 6. Visual impression of original image (left) and the same image blurred by Gaussian filter (right)

5.2 Results

It is shown that both the number of detected features and check point accuracy decreases. Figure 7 compares both results.

While total amount of automatic extracted SIFT-features decreases by 8.2% the amount of valid features used for bundle block adjustment diminishes by 13.7% from original to blurred

images. Resulting 3D-Error differs by 18.6% with Z-Error component (19.9%) as main factor of influence. Based on the assumption that real lens systems not only are degraded through Gaussian blur but also aberration effects, the real differences could assumingly be even higher than within this simulation.

		Difference %		
Nr. of features per Image (mean)	39569	8.2	Nr. of features per Image (mean)	36315
Matched features per Image (mean)	21706	13.7	Matched features per Image (mean)	18739

RMSE [cm]	Check points		RMSE [cm]	Check points
X-Error	0.349		X-Error	0.337
Y-Error	0.236		Y-Error	0.247
Z-Error	1.372	19.9	Z-Error	1.712
3D-Error	1.435	18.6	3D-Error	1.762

projection error (mean)[pix]	0.408		projection error (mean)[pix]	0.415

Figure 7. Simulated effects of image sharpness on bundle block adjustment for original images (left), blurred images (right) and corresponding differences (center)

6. CONCLUSION AND OUTLOOK

The significance of effective geometric resolution as an essential quality parameter of imaging systems has been accentuated and corresponding measurement methods have been described. Ensuing open issues defining a standardized process using signal processing techniques have been categorizes with respect to origin (e.g. algorithmics or measurement arrangement) and answered to some extend.

Future work is going to deepen the investigation and thoroughly clarify all open issues as support to research of the "German Institute for Standardization".

References

Born, M. and Wolf, E., 1999. *Principles of Optics*. Cambridge University Press.

Burns, P. D., 2000. Slanted-edge mtf for digital camera and scanner analysis. In: *Is and Ts Pics Conference*, SOCIETY FOR IMAGING SCIENCE & TECHNOLOGY, pp. 135–138.

Chang, E., Cheung, S. and Pan, D., 1999. Color filter array recovery using a threshold-based variable number of gradients. *Sensors, Cameras, and Applications for Digital Photography*.

Coltman, J. W., 1954. The specification of imaging properties by response to a sine wave input. *J. Opt. Soc. Am. 44* pp. 468–471.

Góźdź, J., 2009. DCB demosaicing. http://www.linuxphoto.org/html/dcb.html (April 2016).

Hartley, R. and Zisserman, A., 2004. *Multiple View Geometry in Computer Vision*. Second edn, Cambridge University Press.

Hirakawa, K. and Parks, T. W., 2005. Adaptive homogeneity-directed demosaicing algorithm. *IEEE Transactions on Image Processing* 14(3), pp. 360–369.

Kraft, T., Geßner, M., Meißner, H., Cramer, M., Przybilla, H.-J. and Gerke, M., 2016a. Evaluation of a metric camera system tailored for high precision uav applications. In: *International Society for Photogrammetry and Remote Sensing*, Volume XLI-B1, 2016.

Kraft, T., Geßner, M., Meißner, H., Przybilla, H.-J. and Gerke, M., 2016b. Introduction of a photogrammetric camera system for rpas with highly accurate gnss/imu information for standardized workflows. In: J. Skaloud and I. Colomina (eds), *EuroCOW 2016, the European Calibration and Orientation Workshop (Volume XL-3/W4)*, pp. 71–75.

Malvar, H. S., He, L.-w. and Cutler, R., 2004. High-quality linear interpolation for demosaicing of bayer-patterned color images. In: *Acoustics, Speech, and Signal Processing, 2004. Proceedings.(ICASSP'04). IEEE International Conference on*, Vol. 3, IEEE, pp. iii–485.

Meißner, H., Cramer, M. and Piltz, B., 2017. Benchmarking the optical resolving power of uav based camera systems. In: *UAVg 2017, The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 42.*, p. 243.

Mix, P. E., 2005. *Introduction to Nondestructive Testing: A Training Guide*. JOHN WILEY & SONS INC.

Reulke, R., Becker, S., Haala, N. and Tempelmann, U., 2006. Determination and improvement of spatial resolution of the CCD-line-scanner system ADS40. *ISPRS Journal of Photogrammetry and Remote Sensing* 60(2), pp. 81–90.

Reulke, R., Tempelmann, U., Stallmann, D., Cramer, M. and Haala, N., 2004. Improvement of spatial resolution with staggered arrays as used in the airborne optical sensor ADS40. In: *Proceedings of the XXth ISPRS Congress*.