

# Digital Photogrammetric Camera Evaluation – Generation of Digital Elevation Models

NORBERT HAALA, Stuttgart, HEIDI HASTEDT, Birmensdorf, Switzerland, KIRSTEN WOLF, Zurich, Switzerland, CAMILLO RESSL, Vienna, Austria & SVEN BALTRUSCH, Schwerin

**Keywords:** Three-dimensional, point cloud, LiDAR, image matching

**Summary:** During the implementation of the DGPF-project on Digital Photogrammetric Camera Evaluation a team “Digital Elevation Models” was established. The main goal was to use the test’s framework for documentation and evaluation of the current state-of-the-art on photogrammetric 3D data capture from automatic image matching. During these investigations the accuracy and reliability of DSM rasters and 3D point clouds as derived from imagery of digital photogrammetric camera systems were evaluated. For this purpose they were compared to reference measurements from ground truth and airborne LiDAR. In addition to the evaluation of standard products, the usability of elevation data from image matching was investigated while aiming at specific applications in the context of urban modeling and forestry.

**Zusammenfassung:** *Digitale photogrammetrische Kamera Evaluierung – Generierung von Digitalen Höhenmodellen.* Während des DGPF-Projektes zur Evaluierung digitaler photogrammetrischer Luftbildkamerasysteme wurde auch eine Auswertegruppe für die Bewertung der Genauigkeit der Höhenmodellgenerierung etabliert. Dabei sollte der DGPF-Test genutzt werden, um den derzeitigen Stand der Technik der photogrammetrischen 3D Erfassung mittels automatischer Bildzuordnung zu dokumentieren. Hierfür wurden DSM Raster und 3D Punktwolken aus Bildern der photogrammetrischen Kamerasysteme abgeleitet und die Qualität dieser Ergebnisse in Bezug auf Genauigkeit und Zuverlässigkeit bewertet. Dabei wurde ein Vergleich zu terrestrischen Referenzmessungen und flugzeuggestützten LiDAR Daten durchgeführt. Neben der qualitativen Bewertung von Standardprodukten wurde auch die Nutzbarkeit der Höhen- daten für spezielle Anwendungen beispielsweise im Kontext der 3D Stadmodellierung und Forstwirtschaft untersucht.

## 1 Introduction

High image dynamic and good signal-to-noise ratio are well known advantages of digital photogrammetric cameras. Compared to the use of scanned analogue images, these improvements of digital imagery are especially advantageous with respect to the accuracy, reliability and density of automatic point transfer. Thus, follow-up products like Digital Elevation Models, which are based on the use of automatic image matching, will potentially benefit, if digital photogrammetric camera systems are used. This progress can be demonstrated and documented very well using comprehensive test data sets as available with-

in the DGPF project on Digital Photogrammetric Camera Evaluation. In order to investigate the current state-of-the-art on image based generation of elevation data, which of course also influences the usability of such products, a special working group with members mainly from academia and administration was established. While the general goal of the DGPF project was to comprehensively analyze photogrammetric digital airborne camera systems, within this group the impact of the captured image data to the available quality of digital elevation models was investigated.

In order to evaluate the quality of such a photogrammetric product, the analysis can of

course not be restricted to image collection but has to pay attention to the respective software for the following data processing. Commercial software systems aiming at the generation of Digital Terrain Models from image matching were already introduced more than two decades ago (KRZYSZEK 1991, COGAN et al. 1991). Nevertheless, the improvements in the available quality of aerial imagery triggered a renaissance in software development to optimally benefit from these advancements. As an example, digital airborne camera systems can capture largely overlapping images at a relatively little additional effort. The availability of such high redundant multi-image information is especially beneficial in situations, where standard stereo matching is hindered due to occlusions. Algorithms which fully exploit this potential of digital aerial cameras by extending the traditional stereo matching to a multiple image matching have been implemented just recently. Such commercial software systems, which will also be used for our investigations are Next Generation Automatic Terrain Extraction (NGATE) from BAE Systems (DEVENEZIA et al. 2007), MATCH-T DSM from INPHO GmbH (LEMAIRE 2008) and SATellite image Precision Processing (SAT-PP) of the ETH Zürich (ZHANG & GRUEN 2004).

One general problem during the evaluation of height data from image matching is to separate the influence of the respective factors on the resulting quality. Major impact results from the quality of the available image data and the sophistication of the used matching algorithms. Additionally, the geometric complexity of the respective object surfaces is of considerable influence. An important factor for image quality is the accuracy and stability of its reconstructed geometry. This is again affected by the geometric configuration of the image block, the geometric stability of the camera and the accuracy and reliability of the camera model. Additionally, image quality depends on the signal-to-noise-ratio of the digitized image signal, which is again influenced by the quality of the sensor system but also by the respective illumination and texture of the depicted surface patches. Finally, the generation of elevation data is influenced by the respective algorithms applied for automatic

point transfer or surface interpolation and filtering. Thus, the wide range of interacting factors, which mutually influence the quality of the generated data complicates a comprehensive analysis of automatic image based elevation measurements.

In our opinion, in addition to a comparative analysis of the respective accuracies, application driven investigations are of even greater interest for potential users. For this reason, accuracy analyses using suitable reference data are complemented by investigations on the usability of elevation data from image matching. Within the paper, special interest is paid to tasks like city model generation or applications in forestry. Firstly, the available test and reference data will briefly be introduced in the following Section. In Section 3 signalized points and selected planar areas are used as reference for a comprehensive analysis of elevation data generated from different imagery and software systems. The quality and usability of DSM from image matching for different applications like data collection in urban areas, investigations in forestry and DTM generation will be discussed in Section 4, while Section 5 will conclude the paper.

## 2 Test Scenario and Reference Data

Within the investigations presented in this paper data sets from the cameras DMC, ADS 40 2<sup>nd</sup>, UltraCamX, Quattro DigiCAM and RMK-Top15 captured at two different flying heights and block configurations were used. In addition to the terrestrial reference points, LiDAR data was made available for comparison to the DSMs from image matching.

### 2.1 Block Configuration and Image Processing

For investigations on the elevation data generation from image matching, the availability of different ground sampling distances [GSD] and image overlaps is of special interest. This was the reason to plan the collection of image blocks with 20 cm GSD and 60% along- and across-track-overlap as well as flights with

GSD 8 cm and 80% along- 60% and across-track-overlap. Due to variations of the different cameras footprint and restrictions from the available ground control, slight deviations from this configuration had to be accepted. The parameters of the investigated camera systems as well as the test design including the respective block configurations are documented in more detail by (CRAMER 2010). During our investigations DSM grids with 0.2 m/0.25 m and 0.5 m raster width were computed for the 8 cm and 20 cm GSD flights in the central of 5.0 × 2.7 km<sup>2</sup> area of the test field.

### 2.2 Preprocessing and Accuracy Analysis of Collected LiDAR Data

For investigating the height accuracy of the derived height models, 63 reference ground points were made available to the test participants (see Section 3). This way, however, the height accuracy can only be checked at discrete locations. For a continuous accuracy check the entire area was therefore surveyed by LiDAR. In total 10 strips were captured with a Leica ALS50 system at 45° FOV with a mean flying height above ground of 500 m and

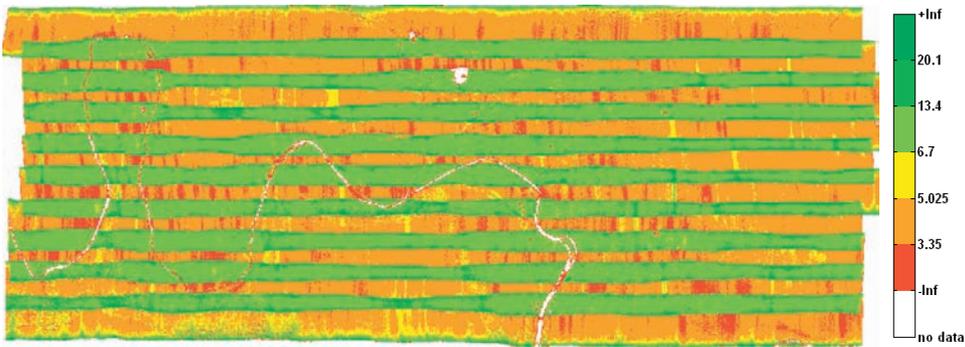


Fig. 1: Color-coding of the point density of all 10 LiDAR strips. The legend is in points/m<sup>2</sup>.



Fig. 2: Sample of a color-coded strip difference for the original georeferencing (top) and for the improved georeferencing after strip adjustment (bottom). Right: Legend of color coding. Black is used for the area outside the overlap of neighboring strips, but also for the parts covered by the roughness mask.

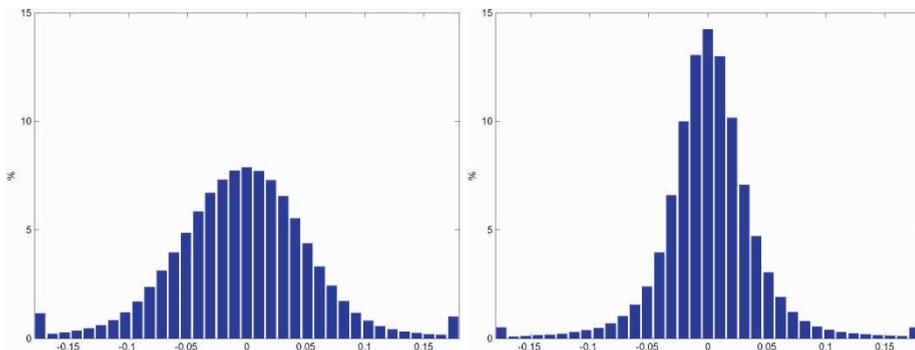
a mean strip overlap of 30 %. The median of the point density is 6.7 points/m<sup>2</sup>, however, the point density varies a lot over the whole block (see Fig. 1). Regions covered by only one strip have a mean density of 4 points/m<sup>2</sup>.

In order to use the LiDAR data as a reference for the height models derived from the aerial images, the georeferencing of the LiDAR data should be checked in advance. A simple and effective tool for checking the quality of the relative orientation of the LiDAR strips are strip differences (RESSL et al. 2008). For this a DSM is interpolated for each strip (with 1 m grid width) and then the difference of pairs of overlapping strip DSMs is computed. Because of the interpolation involved, the differences derived at rough surface areas, e. g., vegetation, are not suitable for judging the accuracy. For considering only smooth surfaces a roughness mask for each strip is used (RESSL et al. 2008).

Fig. 2 (top) shows a sample of a color-coded masked strip difference for the original georeferencing and Fig. 3 (left) shows the histogram of all 9 masked strip differences between the 10 strips. From this histogram a  $\sigma$ MAD value of 4.5 cm is derived for the masked strip differences.  $\sigma$ MAD is the standard deviation derived from the median of absolute differences (the so-called MAD) as  $\sigma$ MAD = 1.4826 · MAD. Although 4.5 cm may appear acceptable, the color-coding reveals large systematic errors visible at the buildings. There the large vertical differences exceeding 18 cm are caused by horizontal displacements between the neigh-

boring strips (i.e. errors of the relative orientation), which themselves result from residual errors in the individual system components: GNSS (Global Navigation Satellite System), INS (Inertial Navigation System), the laser scanner and the mounting calibration (which describes the rotation and translation between these individual components).

In order to minimize the systematic error patterns visible in the color-coded strip differences, a LiDAR strip adjustment was performed following the procedure described by (KAGER 2004). For this the GNSS/INS trajectory of the strips and 1110 corresponding tie planes were used in order to correct internal systematic errors (like a wrong mounting calibration) and to improve the relative orientation of the strips by minimizing the residuals at corresponding planes in the strips. For comparing the LiDAR data with the DSMs derived from the images, both should refer to the same datum. Therefore the absolute orientation of the LiDAR data should be adapted in case their GNSS/INS data refers to a wrong datum. However, no suitable ground control planes for the LiDAR data were available from terrestrial measurements. Therefore, 49 ground control planes were derived from the aerial triangulation (with available ground control points) of the DMC photos with 8 cm GSD. These control planes were used simultaneously in the strip adjustment together with the tie planes. This entire LiDAR strip adjustment therefore can very much be compared with block adjustment by integrated sensor orienta-



**Fig. 3:** Histogram of the strip differences (considering the roughness mask) based on all overlapping strips (ca. 6.5 million values). Left: original georeferencing ( $\sigma$ MAD = 4.5 cm). Right: improved georeferencing after strip adjustment ( $\sigma$ MAD = 2.9 cm).

tion in case of aerial images. The effect of the strip adjustment on the LiDAR data in flight direction is 1 cm (mean), 13 cm (RMS) and 44 cm (max), across flight direction -6 cm (mean), 10 cm (RMS) and -23 cm (max), in height 0 cm (mean), 3 cm (RMS) and -15 cm (max). After the strip adjustment new strip differences were computed; see Fig. 2 (bottom). By comparing top and bottom of Fig. 2 one can clearly see that the systematic errors are largely removed. Fig. 3 (right) shows that the  $\sigma$ MAD of the strip differences improves from 4.5 cm to 2.9 cm.

Although the images are now used to adapt the reference LiDAR data, the effect on the later is only in the absolute orientation. The positive effect is that deviations between the DSM from the LiDAR data (with the improved georeferencing) and the DSM derived from the images cannot be attributed to residual orientation errors, but can primarily be attributed to the different quality of the local surface description of the LiDAR data and the DSM derived from the images. Although in this comparison it should be considered, that another LiDAR flight (with different flying height and/or point density) would give a different result.

### 3 Accuracy Investigations for DSM and Point Clouds

The Vaihingen/Enz photogrammetric test site where the flight campaigns of the DGPF test were realized consists of approximately 200 signalized and coordinated reference ground points, distributed in a  $7.5 \times 5.0$  km<sup>2</sup> area. The central area of the test field, where the investigations on elevation data generation from image matching were concentrated has a size of  $5.0 \times 2.7$  km<sup>2</sup> with approximately 63 reference ground points available for the test participants. The coordinates of all reference points were determined with static GPS base line observations, which provide an accuracy of 1 cm for horizontal and 2 cm for vertical coordinates. Using the vertical differences between the elevation data from image matching and the available reference points a quality estimate can be realized. For our investigations, DSM grids of 0.2 m raster width were generated from the 8 cm GSD imagery of the UltraCamX, Quattro DigiCAM and the scanned RMK images using the software MATCH-T DSM. Similarly, the 20 cm GSD blocks were used for computation of 0.5 m raster DSM grids. From these DSM, differences to the signalized points were computed and further analyzed. The results of these analyses are summarized in Tab. 1.

**Tab. 1:** Differences between DSM and reference points after gross errors elimination – MATCH-T DSM.

	Sensor	RMS [cm] no gross errors	Mean [cm]	$\Delta$ Max/Min [cm]		# points
LiDAR	ALS 50	3.3	0.4	9.4	- 6.7	59
GSD 8 cm Raster 0.2 m	DMC	3.3	0.9	9.5	- 6.9	60
	UltraCamX	4.8	0.6	11.7	-10.0	60
	DigiCAM	6.0	-1.7	15.5	-15.7	61
	RMK	4.6	2.4	8.2	-11.5	61
GSD 20 cm Raster 0.5 m	DMC	16.2	-7.5	36.9	-30.5	61
	UltraCamX	7.5	-0.7	14.9	-16.8	60
	DigiCAM	9.6	0.5	18.9	-23.1	61
	RMK	9.5	0.7	23.9	-25.9	61

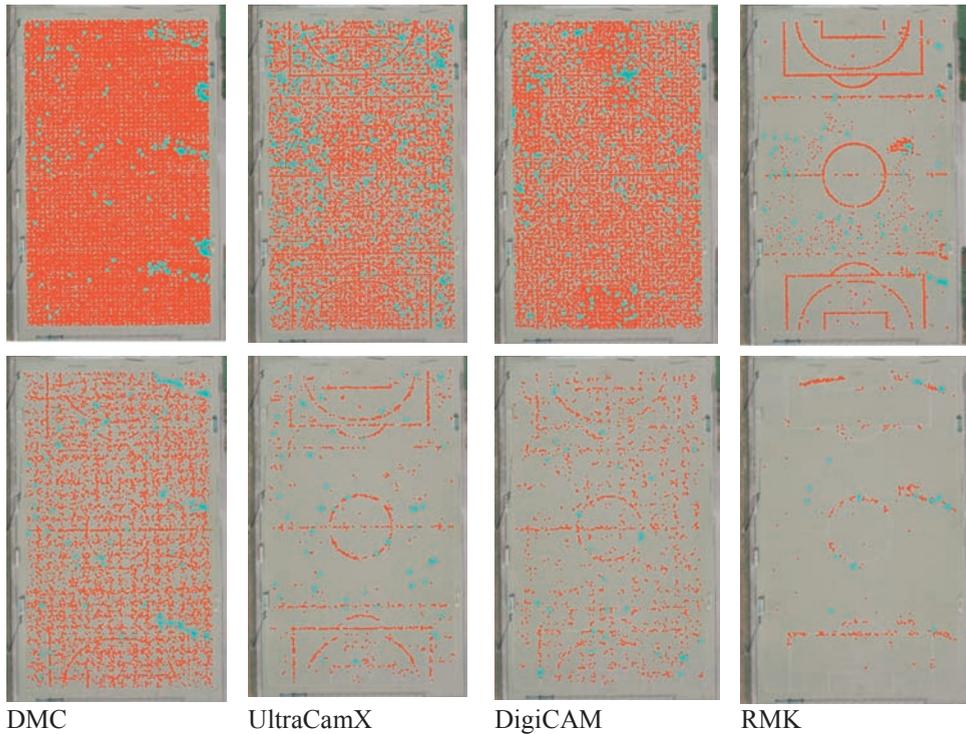
As it is visible in the first column of Tab. 1, DSM grids with 0.2 m and 0.5 m raster width were computed for the 8 cm and 20 cm GSD flights, respectively. The second column gives the investigated camera systems DMC, UltraCamX, Quattro DigiCAM and RMK-Top15. As it is also visible, a DSM as generated from the available LiDAR flight was evaluated for comparison. The third column of Tab. 1 gives the RMS values calculated from the filtered differences between reference point and respective DSM surfaces. An analysis of all differences between the respective DSM and the available reference points shows, that in all cases the largest differences occur at areas potentially compromised to occlusions. In order to eliminate these potential gross errors, a simple threshold was used. First a RMS value was calculated from the height differences to all signalized points, which were available for the respective DSM area. In a second step all points with differences outside a range of  $\pm 3 \cdot \text{RMS}$  were eliminated as gross errors and the remaining differences were used to calculate the filtered RMS. Typically, one or two points were filtered out from the complete of signalized points. This was sufficient for our investigation. However more advanced methods for accuracy assessment of digital elevation models by means of robust statistical method are for example described in (HÖHLE & HÖHLE 2009). Tab. 1 also gives the mean as well as the maximum and minimum values  $\Delta \text{Max/Min}$  from all differences for each DSM. Again the point set with gross errors eliminated was used to calculate these values. The final column gives the number of points after filtering.

While the results in Tab. 1 are based on the use of the software MATCH-T, Tab. 2 shows the results for the DSM grids alternatively generated with BAE Systems NGATE. There the camera systems ADS 40, DMC and UltraCamX were investigated, while DSMs were computed with 0.25 m and 0.5 m raster width for the 8 cm and 20 cm flights, respectively.

As given in Tab. 1, the RMS value for the LiDAR DSM measured by the ALS 50 sensor is 3.3 cm. This is almost in the order of the vertical accuracy of the used reference points. Compared to this accuracy, the RMS values of the DSMs for the DMC, UltraCamX, Quattro DigiCAM and ADS 40 as given in Tab. 1 and 2 are only slightly larger. They correspond very well to the vertical component of the preceding block adjustment, which gave an accuracy of  $\frac{1}{2}$  GSD (JACOBSEN et al. 2007). Typically, the ground control points used to evaluate the DSM quality in Tab. 1 and 2 were installed at paved areas like small roads or parking lots. Such flat neighborhoods are of course beneficial for the filtering and interpolation process during DSM raster generation. For this reason, the results presented in Tab. 1 and 2 might give too optimistic accuracies for regions of higher geometric complexity. As an alternative 3D point clouds can be used to evaluate the matching quality during accuracy analyses without the influence of interpolation processes. Such point clouds can be optionally generated from modern photogrammetric software systems, and can for example be used as an alternative to the traditional 2.5D raster representations of Digital Surface Models during tasks like 3D object reconstruction.

**Tab. 2:** Differences between DSM and reference points after gross errors elimination – NGATE.

	Sensor	RMS [cm]	Mean [cm]	$\Delta \text{Max/Min}$ [cm]		# points
		No gross errors				
GSD 8 cm Raster 0.25 m	ADS 40	6.7	-1.1	13.9	-18.1	57
	DMC	4.4	-1.2	9.0	- 8.8	53
	UltraCamX	7.2	1.6	16.1	-11.8	59
GSD 20 cm Raster 0.5 m	ADS 40	4.8	1.9	12.9	- 8.8	60
	DMC	19.0	-2.7	51.3	-31.4	61
	UltraCamX	11.5	1.6	27.2	-21.6	61



**Fig. 4:** Point clouds for investigated camera systems generated by MATCH-T DSM. The top row shows the results for the 8 cm GSD block, the bottom row for the 20 cm GSD block.

Within our investigations 3D point clouds were computed and evaluated for the data from the frame based camera systems DMC, UltraCamX, Quattro DigiCAM and RMK-Top15. In order to evaluate the accuracy of the generated 3D point clouds by a relatively simple process, a test area at a planar sports field was defined. The respective point clouds as generated from MATCH-T DSM are depicted

in Fig. 4. Results for the 8 cm GSD blocks are shown in the top row, while the matching results from the 20 cm GSD blocks are presented in the bottom row.

Since the matched 3D points are restricted to a planar area, their geometric accuracy can be determined based on the estimation of an approximating plane. After a best fitting plane is determined the perpendicular point distanc-

**Tab. 3:** Accuracy of 3D point clouds MATCH-T DSM and NGATE – GSD 8 cm.

Sensor	STD after filter [cm]		STD no filter [cm]		Elim.Pts [%]		Density Pts/m <sup>2</sup>	
	M-T	NGATE	M-T	NGATE	M-T	NGATE	M-T	NGATE
DMC	5.2	2.1	9.7	2.3	1.3	0.9	19.7	8.2
UltraCamX	6.8	13.1	8.0	15.6	0.4	1.5	19.0	8.2
DigiCAM	10.2		11.2		0.7		20.8	
ADS 40		2.3		2.6		0.7		8.2
RMK	17.2		27.3		3.2		0.8	
ALS 50	1.8		1.9		0.5		8.25	

**Tab. 4:** Accuracy of 3D point clouds MATCH-T DSM and NGATE – GSD 20 cm.

Sensor	Stdv. after filter [cm]		Stdv. no filter [cm]		Elim.Pts [%]		Density Pts/m <sup>2</sup>	
	M-T	NGATE	M-T	NGATE	M-T	NGATE	M-T	NGATE
DMC	17.2	7.5	25.4	9.1	1.1	1.7	2.7	4.0
UltraCamX	22.6	25.0	34.2	38.1	0.4	1.0	1.6	2.6
DigiCAM	34.1		48.2		2.5		2.6	
ADS 40		7.4		8.3		1.4		4.0
RMK	60.6		66.2		0.7		0.3	

es are used to determine the respective standard deviations, which represent the accuracy of 3D point measurement. Again, a threshold is used to eliminate points outside a range of  $\pm 3$ -RMS. These erroneous points usually correspond to shadow areas from the goals and the floodlight poles. Such time dependent shadow movement can result in considerable errors of automatic point transfer especially if high resolution images from different strips are matched. Within Fig. 4 points eliminated by the filter process are marked in light blue, while the remaining points are shown in red.

Tab. 3 and 4 summarize the results of point cloud analysis. There, the standard deviations from the matched points are given in addition to the percentage of points eliminated in this filtering process. The final column gives the point density as provided from image matching, which is an important indicator for the quality of this process. Using the software MATCH-T DSM, on average, a point density of about 20 pts/m<sup>2</sup> was reached using the GSD 8 cm images from the digital camera systems. In contrast, the matching of scanned RMK images gives less than 1 pt/m<sup>2</sup>. Obviously, the higher radiometric quality of digital images allows for much denser point matching while RMK-Top15 imagery is not suitable for the automatic derivation of high accurate surface models. This supremacy is verified for all digital camera systems. This result is especially relevant for the DMC and RMK images, which were recorded almost simultaneously at identical atmospheric and illumination conditions by using a double-hole aircraft.

Additionally, the results presented in Fig. 4 and Tabs. 3 and 4 show a considerable advan-

tage of point matching for the GSD 8 cm blocks compared to the GSD 20 cm blocks for all digital camera systems. For MATCH-T DSM, the point density using the GSD 8 cm images from the digital camera systems is even higher than the approximately 10 pts/m<sup>2</sup>, which were generated by the ALS 50 laser scanner at the sports fields. However, the standard deviation for the LiDAR data is better than 2 cm, almost without any gross errors, while an average of 5.5 cm for the filtered points is achieved from image matching. Thus, for the 8 cm block an accuracy of below 1 pixel GSD was achieved for the single point measurements. For the GSD 20 cm this value is slightly worse with an average standard deviation of 14.1 cm for the digital cameras. Compared to the 8 cm GSD block, the average point density of 1.8 pts/m<sup>2</sup> is much lower. For this reason, especially height data as it can be provided from largely overlapping high resolution imagery like the GSD 8cm blocks seems to be at least comparable to 3D data from LiDAR measurement.

While aiming at a joint evaluation of the different digital camera systems DMC, UltraCamX, Quattro DigiCAM and ADS 40 it has to be considered, that due to the test period of more than 2 months, there were significant changes in vegetation as well as atmospheric conditions and illumination. Some of the flights were done quite early in the morning, others were flown around noon. These differences in illumination of course influence the matching quality also for areas of little texture like the investigated sports field. Together with the variations of the block geometry these differences considerably influence the results

as available from the digital camera systems. Furthermore, the variations of the respective results with respect to the two applied software systems MATCH-T DSM and NGATE clearly indicate the influence of the respective matching and filtering algorithms on the generated elevation data. However, a comprehensive analysis of such influences is beyond the scope of this paper.

#### 4 Usability of Elevation Data from Image Matching

Dense and accurate elevation data are required for a large number of applications, like 3D-landscape visualization or the generation of products like true orthophotos, 3D-building models or DTMs. Especially if aerial images are already collected for other purposes, image matching is economically advantageous compared to the additional use of alternative sensors like RADAR or LiDAR. As an example, most national mapping agencies collect digital aerial images countrywide and resume the acquisition within short time periods mainly for generation of actual ortho imagery. This leads to the possibility of generating DSMs from image matching within the same time period.

Fig.5 exemplarily depicts a DSM from DMC 8 cm GSD image matching and the ALS 50 LiDAR measurement for a part of the test area. As it is visible, the differences between

both surfaces are rather small and mainly correspond to vegetated areas. During the DGPF test, the DMC 8 cm GSD imagery was captured at July 24th 2008 while, airborne LiDAR (ALS) was collected at August 21th 2008. Due to the time gap of four week between the DMC and the LiDAR flight the differences between both DSM most probably result from plant growth. Additionally, as a result of the different measurement principles, the surface which is actually captured might be different in these areas. As an example, ALS measurement will partially penetrate a tree canopy, while matching will most probably relate to the visible surface. As discussed in the following sections such effects are especially important if the resulting elevation data is further analyzed for applications in forestry or for DTM generation.

##### 4.1 First Investigations in Urban Areas

For a first investigation of the potential of the camera systems for DSM generation in urban areas, interpolated, regular DSMs (25 cm grid) using the software package SAT-PP (Satellite Image Precision Processing, ETH Zurich) based on data from the frame based camera systems DMC and UltraCamX (8 cm GSD) were generated. Especially in urban areas time depending changes and the differences of the level of detail between the image matching

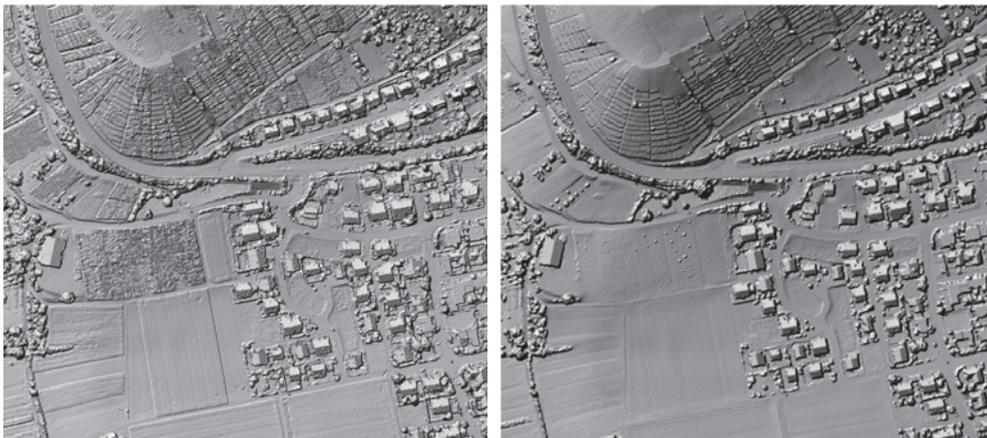
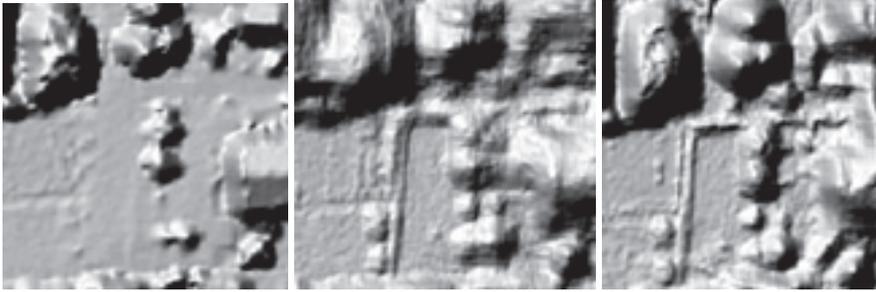


Fig. 5: DSM from image matching (left) and airborne LiDAR (right).

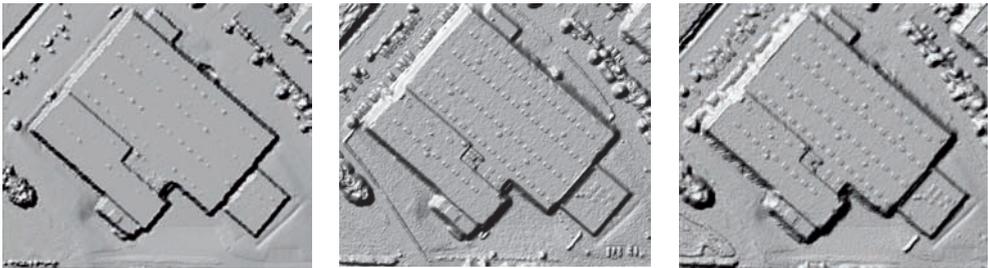
DSMs and the interpolated reference DSM (25 cm grid) from the LiDAR point cloud are problematically for the evaluation process. These differences make an area based comparison between the generated DSMs and the reference LiDAR DSM less representative and

less significant. Figs. 6 and 7 illustrate the different level of detail of the data sets. A small wall or hedge is visible in both image matching DSM, but not in the LiDAR DSM.

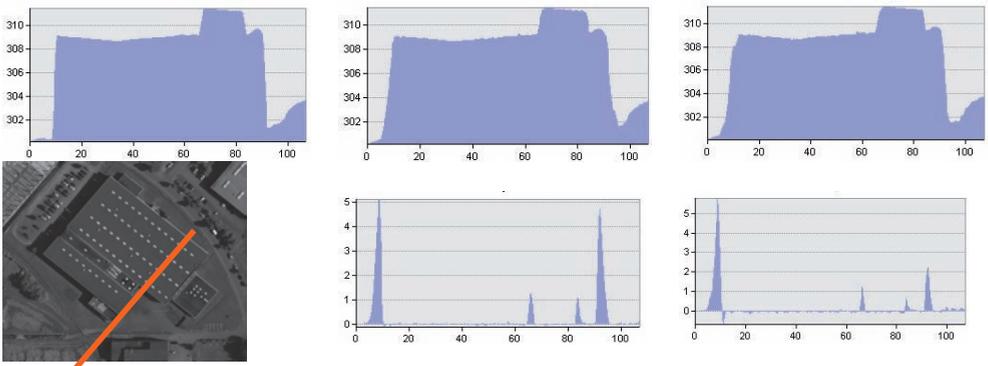
For simpler comparison DSM profiles are used instead of an area based evaluation meth-



**Fig. 6:** DSM from airborne LiDAR (left) and image matching for the data of the DMC, 8 cm GSD, 6 images overlapping (middle) and the UltraCamX, 8 cm GSD, 13 images overlapping (right).



**Fig. 7:** DSM from airborne LiDAR (left) and image matching for the data of the DMC, 8 cm GSD, 6 images overlapping (middle) and the UltraCamX, 8 cm GSD, 10 images overlapping (right).



**Fig. 8:** Upper line: profile through the interpolated DSM from airborne LiDAR (left) and image matching for the data of the DMC, 8 cm GSD, 6 images overlapping (middle) and the UltraCamX, 8 cm GSD, 10 images overlapping (right). Lower line: original DMC, 8 cm GSD image (left), the position of the profile (red line) and the differences between the LiDAR DSM and the image matching DSM along the profile. All values are given in meters.

od; in general, the latter is more preferable. In the following results for an industrial building with a length of 113 m are discussed, exemplarily. Fig. 7 shows the different DSMs of the building and again the different level of detail.

Fig. 8 gives the original DMC 8 cm GSD image, the different profiles and the comparison between the LiDAR DSM and the image matching DSMs. The level of detail of the image matching DSMs is high, edges are reconstructed well, on top of the building the difference to the LiDAR DSM is very small, and blunders are detectable only in the area of buildings borders. The investigations discussed in this section were done before the georeferencing of the LiDAR data was improved by strip adjustment. Therefore, here the original georeferencing was used. Thus, the deviations at the building borders also indicate the need for improving the original georeferencing of the LiDAR data as discussed in Section 2.2.

## 4.2 Forestry Applications

Information on height and 3D structure is a strongly needed input in many forestry applications. A dense, accurate and up-to-date digital surface model (DSM), assuming a digital terrain model (DTM) being available, is therefore required in order to get an appropriate canopy height model. Several investigations aim on the combination of LiDAR data and multispectral images to develop and evaluate methods for the determination of tree and forest attributes. Frequently, aerial image information is applied for classification purposes while LiDAR data is used for the estimation of DSM and DTM. Estimated as the difference from the DSM to its corresponding DTM, the canopy height model (CHM) is the base model within the aspired forestry applications.

$$\text{CHM} = \text{DSM} - \text{DTM} \quad (1)$$

DTMs are constant over a long time and supported by the national mapping agencies. DSMs not, in particular not with respect to environmental purposes, thinking of flooding, wind damages or shrub encroachment. In or-

der to be able to estimate a dense and accurate up-to-date CHM investigations on the generation and reliability of DSMs are required. In forestry applications CHMs are used to derive different forest attributes on single tree and stand level (e.g. height, crown closure, volume, structure). Using these attributes one can derive ecological data like above ground biomass, carbon pools, economical data like timber yield for forest management and input data for forest inventories (WASER et al. 2009). These parameters are also used to extract potential tree areas for semi-automatic estimations of main tree species and fractional tree covers. Other applications apply the CHM for the areal acquisition of forest gaps or use CHM in conjunction with Color and Intensity information of aerial images to derive forest areas (BÖSCH et al. 2007). The investigations on semi-automatic extractions of main tree species or forest areas are at present done within small test areas; at the Swiss Federal Institute for Forest, Snow and Landscape Research (WSL) they are currently in use for the evaluation of the canton Appenzell. Due to missing up to date LiDAR data but usage of most recent aerial images, semiautomatic methods for forestry applications are difficult to refine especially converting them to large areas like a whole country. High accurate and dense up to date DSMs from image matching are therefore the future for the derivation of the base model CHM for forestry applications. ST.-ONGE et al. (2001) evaluated the potential of using DSMs from image matching with digitised analogue photographs. HEURICH et al. (2004) determined differences of forest surfaces comparing LiDAR and image matching methods using DMC data and clearly showed the potential of DSMs from image matching. In this context, investigations using data of the DGPF test are of considerable importance.

Besides the overall accuracy with respect to control points, which are mainly on ground and good textured surface, the differences of image matching with respect to LiDAR and 3D stereo measurements of single surface points in vegetation covered areas are interesting for environmental applications. First, the image matching results from matching with one stereo image pair are compared to LiDAR original and corrected data. In order to ex-

clude a benefit in areas of LiDAR overlapping data lines only one set of LiDAR data was used for each analysis. Secondly they are compared to stereo measurements in order to determine the differences from stereo matching to stereo interpretation, whereas the stereo interpretations could be considered as a most probable result of image matching at the specific stereo measurement positions. The differences of DSM to the different comparison values estimate the true values of canopy heights using image matching methods.

Therefore different areas in the testsite Vaihingen/Enz are examined to estimate the potential and deviations with respect to LiDAR data and manual 3D measurements. The analyses give a good overview on the accessible data and its density and the limits in the usage of matching methods for environmental applications using BAE Systems NGATE. ADS 40 (CIR Images), DMC (RGB images) and UltraCamX (UCX; RGB Images) data with 20 cm resolution was used for analyses; all DSMs are calculated with a resulting Ground Sample Distance of 50 cm. The DSMs are generated using one stereo image pair on BAE Systems NGATE in conjunction with a specific parameter set defined at WSL.

**Example 1:  
Area with Compact Crown Closure**

The first example area was chosen due to a compact crown closure with small height differences in the surface. Differences are calculated as an actual-target-comparison; actual = DSM and target = LiDAR – equation (2), its Root Mean Square Error is calculated with equation (3).

$$dZ = \text{actual} - \text{target} \tag{2}$$

$$RMS = \sqrt{\frac{\sum dZ^2}{n}} \tag{3}$$

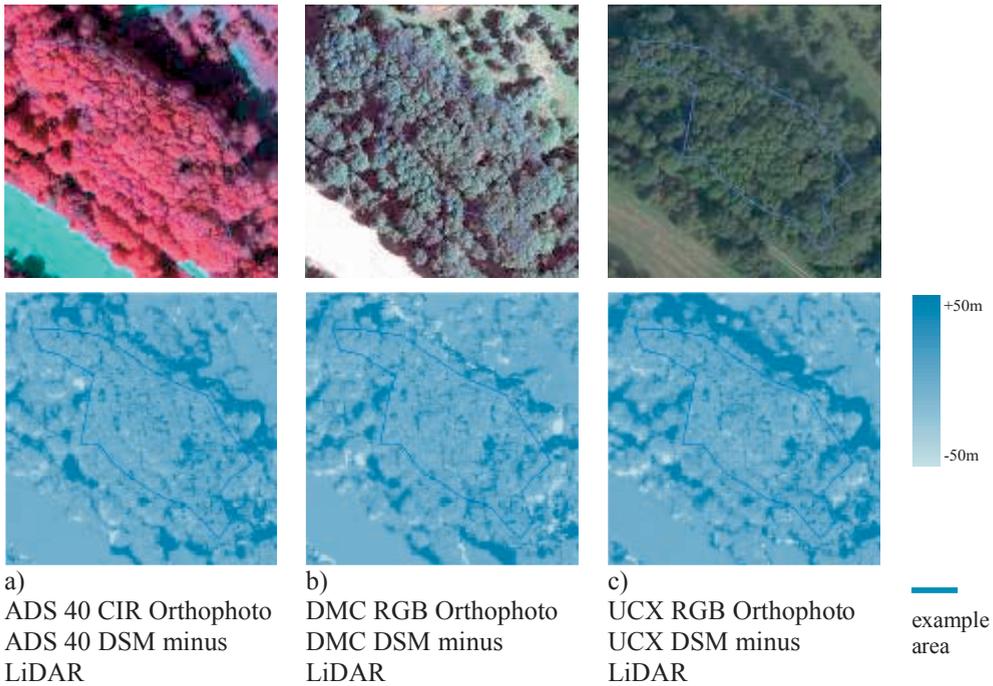
For a first comparison with dense 3D data the generated DSMs and the LiDAR data (original and corrected) are taken for an actual-target comparison. In order to be able to evaluate the most probable deviation to the surface that can be calculated with image matching methods,

~110 manual 3D stereo measurements are done for each data set. This estimates the difference of LiDAR due to leafiness and variation in pulse responds with respect to stereo plotting. All 3D data sets are cross-calculated; the results are listed in Tab. 5. All results rely on a gross error filtering, only differences in the range of  $\pm 3 \cdot RMS$  are used for evaluations. Fig. 9 shows in the upper row the orthoimage section for each data set including the example area that leads to the results in Tab. 5. Beneath each image the corresponding difference model from raster calculation of DSM – LiDAR (original) is shown as a color-coded raster dataset (dark blue – high positive difference, bright blue – high negative difference). The LiDAR original point clouds are imported and merged to a raster data set using ESRI ArcGIS. All figures include the defined example area as blue polygon.

The trend for the deviations is similar in all data sets. For all datasets the DSMs are 25 cm (average) above the 3D stereo measurements surface. The deviation of LiDAR original with respect to the Stereo measurements only results for the ADS40 in a high value of about 70 cm, for DSM and UCX the value is about 35 cm. For all deviations the Stereo measurements are above the LiDAR data set. The mean

**Tab. 5:** Deviations of DSM, LiDAR and Stereo measurements cross-calculated by equation (2).

Sensor	actual	Target	Mean	RMS
			[m]	[m]
ADS 40	DSM	Stereo	0.3	0.5
	LiDAR	Stereo	-0.7	2.7
	DSM	LiDAR	2.0	4.2
	DSM	LiDARcorr	1.9	4.1
DMC	DSM	Stereo	0.2	0.8
	LiDAR	Stereo	-0.3	1.9
	DSM	LiDAR	1.8	4.2
	DSM	LiDARcorr	1.7	4.0
UltraCamX	DSM	Stereo	0.2	0.5
	LiDAR	Stereo	-0.4	1.9
	DSM	LiDAR	1.8	4.1
	DSM	LiDARcorr	1.7	4.0



**Fig. 9:** Image sections and resulting differences from DSM to LiDAR data.

difference between DSM and LiDAR original and corrected is +1.9 m and +1.8 m respectively. Due to the higher density of LiDAR and DSM compared to the sample size of approximately 110 points for stereo measurements we assume the results for deviations of DSM to LiDAR to be more reliable and valid for overall and further analyses.

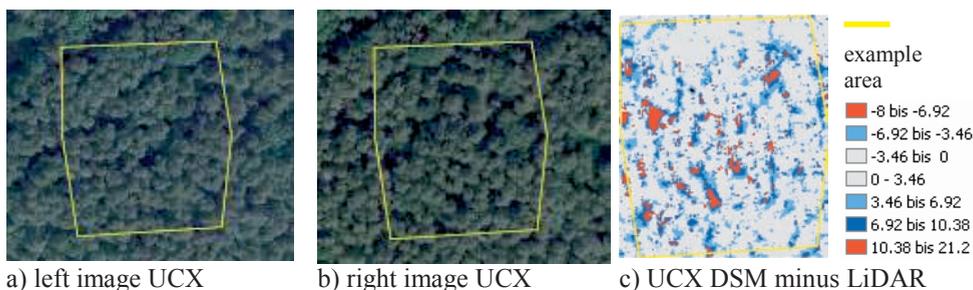
**Example 2: Area with Normal to Light Crown Closure**

The second example covers an area with normal to light crown closure. Gaps and local height differences characterize the example area. For a first comparison with dense 3D data the generated DSMs and the LiDAR data (original and corrected) are taken for an actual-target comparison. Due to the morphological structure and the strong shadows the deviations from DSM to LiDAR are higher compared to Example 1. Fig. 10 shows the corresponding stereo image pair sections of the UltraCamX data; the yellow polygon illus-

trates the example area. Here the strong differences of object space due to the different perspectives are obvious, which lead to higher deviations for image matching. Additionally the resulting color-coded difference model for DSM – LiDAR original is shown on the right side. The colors are chosen with respect to 1–3 sigma of the RMS values. Grey values are within 1 sigma, light blue within 2, dark blue within 3 and red areas are outside 3 sigma, they highlight the strong shadowed parts and areas with high perspective differences.

Tab. 6 concludes the results. All results rely on a gross error filtering, only differences  $< 3 \cdot \text{RMS}$  are used for evaluations. For all data sets the mean value of all differences from DSM to LiDAR original and corrected data is within 3–3.9 m and 2.9–3.8 m respectively; the DSMs therefore result in average 3.5 m above LiDAR.

The analyses for forestry applications show similar trends for the deviations in all data sets, even though the UltraCamX data relies on significant disadvantages in image quality due to weather conditions, which are mani-



**Fig. 10:** Image sections of stereo image pair UltraCamX and resulting differences from DSM to LiDAR data.

**Tab. 6:** Deviations of DSM and LiDAR for example area 2.

Sensor	actual	target	Mean	RMS
			[m]	[m]
ADS40	DSM	LiDAR	3.0	5.9
	DSM	LiDARcorr	2.9	5.8
DMC	DSM	LiDAR	3.8	6.4
	DSM	LiDARcorr	3.7	6.3
UltraCamX	DSM	LiDAR	3.9	6.5
	DSM	LiDARcorr	3.8	6.3

fested in the results of plane analyses. Despite the correction of LiDAR data as described in Section 2.2, the benefit for forestry applications is not as significant as expected. The results show the high potential of new aerial images and the usage of new matching methods for vegetation areas, but further investigations are needed in order to evaluate overall accuracies and more reliable results for different areas and vegetation types.

### 4.3 Generation of ATKIS-DTM

As a component of the Authoritative Topographic-Cartographic Information System (ATKIS), the surveying and mapping agencies of the federal states in Germany provide area covering and actual Digital Terrain Models and digital ortho images. Originally, for high quality DTM generation airborne LiDAR data was used. Since most national mapping agencies take digital aerial images countrywide

and resume the acquisition within short time periods, the use of airborne imagery for automatic DTM update would be highly advantageous to improve the cost efficiency. In order to evaluate the potential of image based ATKIS®-DTM generation, 3D point clouds as provided from the software MATCH-T DSM were further analyzed. For this purpose, filter algorithms available within the software SCOP++ LIDAR from INPHO GmbH were used. This tool has so far been applied by the national mapping agency of Mecklenburg-Vorpommern in order to classify terrain points from airborne LiDAR measurements during the generation of DTM grids.

Fig. 11 shows six test areas used as reference for the following investigations. For these areas, filtered point clouds from different camera configurations Quattro DigiCAM, UltraCamX and DMC were compared to the LiDAR reference measurement. As expected, the existing setup for classification and filtering of LiDAR-points could not directly be used for the evaluation of point clouds from image matching. The available point density from image-matching especially from the 8 cm GSD flights is always higher than from LiDAR measurement. However, image matching provides point distributions, which are suitable for DSM generation, while in forest regions almost no points are available at the terrain surface. This is of course a prerequisite for DTM generation and already motivated the introduction of airborne LiDAR in the nineteen eighties.

In general, the generation of DTM in complex regions like urban areas depends very much on the quality of automatic filtering,



**Fig. 11:** Test areas for DTM generation.

sometimes still the use of additional map data or manual editing is required. At present, LiDAR data filtering still seems to be more advanced than post processing of photogrammetrically derived elevation data, however, it will be interesting to follow the future developments. Since filtering of the 3D points is the main problem during DTM generation, changes in overlap and GSD as well as the use of different camera systems (DigiCAM, UCX, DMC) did not result in significant differences of the result. Despite these problems, image matching can be useful at least for change detection as a prerequisite of DTM update. Usually, high vegetation, gives hint to constant terrain surface, where no update is required. However, low vegetation up to 50 cm like shrub can hide DTM changes. This is especially a problem for longer time periods like 3 years, which is the current flight interval for national mapping agencies.

Similar to LiDAR measurement, an image based generation of DTM requires flights outside the vegetation period – however for ATKIS ortho image generation usually data collection in summer or spring is preferred. Another point to be solved is the amount on computational power and time, which is still re-

quired especially if large areas like a complete federal state have to be covered by high resolution at large overlap. Estimates with actual hard- and software configuration range from 40 to 1470 days for a complete federal state like Mecklenburg-Vorpommern with an area of approximately 23.000 km<sup>2</sup>.

## 5 Outlook and Conclusions

The tests and investigations within the DGPF-project on Digital Photogrammetric Camera Evaluation clearly demonstrated the benefits of digital image recording for elevation data generation by image matching. Data from up-to-date digital airborne cameras facilitate the generation of 3D point clouds and 2.5D raster representations at a quality, which in the past was only feasible by LiDAR measurements. Elevation data from image matching can be used in deriving 3D-building models, roof shapes, canopy models, producing true orthophotos, 3D-landscape visualization and – at least partially – for generating and updating DTMs.

However, compared to LiDAR measurement results from image matching still are

compromised to errors. Potential problems, for example resulting from changing illumination or moving shadows still provide results of partly varying geometric quality. Despite the very promising results, current matching software does not yet fully exploit the complete potential of the new generation of aerial images. Further developments, investigations and tests are still required in the field of multi image matching to broaden potential applications. In order to allow for standard workflows while for example introducing this method into the working practice of National and State Mapping Agencies, also the question of the actually required magnitude of forward and sideward overlap has to be solved. Since any increase of sideward overlap results in longer flying time and therefore raises costs, the especially for large flight mission is very crucial from an economical point of view. Remaining challenges to ameliorate the further use of elevation data from image matching are a further improvement of filter approaches, the reduction of computational cost and an optimal adaption of algorithms for interpretation of surfaces or point clouds from image matching. Additionally, the full use of jointly collected high resolution radiometric and geometric information for the collection of detailed geodata is just at the beginning. It is the aim of efforts like the DGPF test to encourage such developments and further support the current comeback of digital image matching.

## References

- BÖSCH, R., GINZLER, C. & WANG, Z., 2007: Von Decken und Lücken. – Landesforstinventar (LFI) Info Nr. 8, Eidgenössische Forschungsanstalt für Wald, Schnee und Landschaft (WSL), Birmensdorf, Switzerland, 1–4, [www.lfi.ch](http://www.lfi.ch).
- COGAN, L., LUHMANN, T. & WALKER, A.S., 1991: Digital Photogrammetry at Leica, Aarau. – ISPRS Conference on Digital Photogrammetric Systems.
- CRAMER, M., 2010: The DGPF-Test on Digital Airborne Camera Evaluation – Overview and Test Design. – this issue.
- DEVENECIA, K., WALKER, S. & ZHANG, B., 2007: New Approaches to Generating and Processing High Resolution Elevation Data with Imagery. – Photogrammetric Week 17 (5): 1442–1448.
- HEINZEL, J., WEINACKER, H. & KOCH, B., 2008: Full automatic detection of tree species based on delineated single tree crowns – a data fusion approach for airborne laser scanning data and aerial photographs. – *SilviLaser 2008*, 8<sup>th</sup> international conference on LiDAR applications in forest assessment and inventory, 76–85.
- HEURICH, M., SCHADECK, S., WEINACKER, H. & KRZYSTEK, P., 2004: Forest Parameter Derivation From DTM/DSM Generated From LiDAR And Digital Modular Camera (DMC). – *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 37 (B): 84–89.
- HÖHLE, J. & HÖHLE, M., 2009: Accuracy assessment of digital elevation models by means of robust statistical methods. – *ISPRS Journal of Photogrammetry and Remote Sensing* 64: 398–406.
- JACOBSEN, K., CRAMER, M., LADSTÄDTER, R., RESSL, C. & SPRECKELS, V., 2010: DGPF-Project: Evaluation of Digital Photogrammetric Camera Systems – Geometric Performance. – this issue.
- KAGER, H., 2004: Discrepancies between overlapping laser scanning strips – simultaneous fitting of aerial laser scanner strips. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 35 (B1): 555–560.
- KRZYSTEK, P., 1991: Fully Automatic Measurement of Digital Elevation Models with MATCH-T. – *Photogrammetric Week 1991*.
- LEMAIRE, C., 2008: Aspects of the DSM Production with High Resolution Images. – *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 37 (B4): 1143–1146.
- PERSSON, A., HOLMGREN, J., SÖDERMAN, U. & OLSSON, H., 2002: Detecting and Measuring Individual Trees using an Airborne Laser Scanner. – *Photogrammetric Engineering & Remote Sensing* 68 (9): 925–932.
- RESSL, C., KAGER, H. & MANDLBURGER, G., 2008: Quality checking of ALS projects using statistics of strip differences. – *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 37: 253–260.
- ST-ONGE, B. & ACHAICHA, N., 2001: Measuring forest canopy height using a combination of LiDAR and aerial photography data. – *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 34 (3/W4): 131–137.
- WASER, L.T., GINZLER, C., KÜCHLER, M., LANZ, A. & BALTSAVIAS, E., 2009: Semiautomatic prediction of main tree species using multi-temporal ADS 40 data. – *IUFRO Division 4 Symposium on Extending Forest Inventory and Monitoring over space and time*.

ZHANG, L. & GRUEN, A., 2004: Automatic DSM Generation from Linear Array Imagery Data. – International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences **35** (B3): 128–133.

Addresses of the Authors:

Apl. Prof. Dr. NORBERT HAALA, Institut für Photogrammetrie, Universität Stuttgart, Geschwister-Scholl-Str. 42D, D-70174 Stuttgart, Tel.: +49-711-685-83383, Norbert.Haala@ifp.uni-stuttgart.de.

Dipl.-Ing. HEIDI HASTEDT, Eidgenössische Forschungsanstalt für Wald, Schnee und Landschaft WSL, now: Fachhochschule Wilhelmshaven/Oldenburg/Elsfleth, Institut für Angewandte Photogrammetrie und Geoinformatik, Ofener Str. 16, D-26121 Oldenburg, Tel. +49-441-7708-3164, heidi.hastedt@hs-woe.de.

Dr. KIRSTEN WOLFF, Institut für Geodäsie und Photogrammetrie, ETH Zürich, now: swisstopo, Seftigenstrasse 264, CH-3084 Wabern, kirsten.woff@swisstopo.ch

Dr. CAMILLO RESSL, Institut für Photogrammetrie und Fernerkundung, Technische Universität Wien, 1040 Wien, Austria, Gußhausstraße 27–29, Tel: +43-1-58801-12234, car@ipf.tuwien.ac.at.

Dipl.-Ing. SVEN BALTRUSCH, Landesamt für innere Verwaltung Mecklenburg-Vorpommern, Amt für Geoinformation, Vermessungs- und Katasterwesen, Lübecker Str. 289, 19059 Schwerin, Tel.: +49-385-48013202, sven.baltrusch@laiv-mv.de.

Manuskript eingereicht: Dezember 2009  
Angenommen: Januar 2010