Comeback of Digital Image Matching

NORBERT HAALA, Stuttgart

ABSTRACT

Despite the fact that tools for automatic stereo image matching are available for more than two decades, the collection of high resolution, high accuracy elevation data was mainly dominated by the application of airborne LiDAR systems. However, digital airborne cameras meanwhile enable the area covering acquisition of high dynamic image data with good signal-to-noise ratio. The wide availability of this data also triggered the revival of elevation data collection based on image matching. This trend is currently supported by the development of improved software tools which for example extend traditional stereo- to multi-image matching. Within the paper, this increased performance of photogrammetric DSM generation is documented based on data from a test bed which was originally initiated by the German society of Photogrammetry, Remote Sensing and Geoinformation (DGPF). Data from different digital airborne cameras at various flight scenarios were collected in order to comprehensively analyze the performance of new photogrammetric digital airborne cameras. For this purpose, external control from signalized reference points and LiDAR measurements was additionally made available. Within the paper, the current state-of-the-art image of image based elevation data collection is presented and the improved accuracy and reliability of this approach due to current developments in sensor and software technology is demonstrated.

1. INTRODUCTION

The use of digital image matching for automatic point transfer within photogrammetric applications is a well known standard procedure. Software which integrates feature or intensity based matching for automatic aerial triangulation is commercially available since more than two decades. Meanwhile, the automatic identification of corresponding points is also implemented within close range applications, where the situation is even more complex compared to airborne applications. In such scenarios affine invariant feature detectors are combined with robust orientation estimation techniques to cope with matching problems caused by perspective distortions within images of arbitrary rotation and scale. In contrast to the great relevance of image matching for orientation purposes, the importance of this technique for 3D surface reconstruction is subordinated. Similar to automatic aerial triangulation, software tools for image based surface reconstruction are available for more than two decades. However, automatic stereo image matching was frequently substituted by LiDAR measurements. Compared to image based surface reconstruction, LiDAR measurements were more competitive especially while aiming at very accurate and dense elevation data. Meanwhile, this gap could be considerably narrowed since now imagery from digital airborne camera systems is available as standard data source. Compared to scanned film, these images feature an increased dynamic and improved signal-to-noise-ratio. This is highly beneficial for automatic image matching especially for surfaces with relatively little surface texture. As a result, the quality and accuracy of image based point transfer as basic observation for 3D surface reconstruction could be improved considerably. In addition to the better radiometric quality, digital airborne camera systems can also capture largely overlapping images at a relatively little additional effort. The availability of such high redundant multi-image information is especially beneficial in situations, were standard stereo matching is hindered due to occlusions. The increasing potential of such data for surface reconstruction based on automatic image matching additionally triggered a renaissance in the research and development of suitable software tools. As an example, techniques like semi-global matching (Hirschmüller, 2008) were developed for pixelwise matching based on mutual information and the approximation of a global smoothness constraint. Additionally, commercial tools like Next Generation Automatic Terrain Extraction (NGATE) from BAE Sytems

(DeVenecia et. al, 2007) or MATCH-T DSM from INPHO GmbH (Lemaire, 2008), which are based on standard feature and intensity based matching approaches, meanwhile enable the use of multiple overlapping imagery for dense and accurate 3D data collection.

Within the paper this comeback of digital image matching for DSM generation is documented based on data, which was originally captured during a test initiated by the German society of Photogrammetry, Remote Sensing and Geoinformation (DGPF). This test aims at an evaluation of airborne large format photogrammetric sensor systems in combination with the analysis of the photogrammetric software performance (Cramer et. al. 2009). Within the test, different working groups were organized to structure the evaluation phase. Topics which are covered by the network of participating institutions are the analysis of geometric accuracy and sensor calibration, the radiometric performance including on-site radiometric calibration and multi-spectral land classifications. Within this paper, investigations on the performance of photogrammetric generation of elevation data will be used to present and discuss the state-of-the-art on digital image matching.

2. PERFORMANCE EVALUATION – THE DGPF TEST

The accuracy, reliability and density of elevation data as generated from automatic image matching is influenced by a number of factors. Important factors are the quality of the available image data and the sophistication of the used matching algorithms, but also the geometric complexity of the respective object surfaces. Image quality again depends on the accuracy of image geometry which is affected by the geometric configuration of the image block, the geometric stability of the camera and the accuracy and reliability of the camera model. Further on, it 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. Another considerable impact to the final result results from the subsequent data processing, where the employed software usually consists of different modules for automatic point transfer as well as surface interpolation and filtering. The wide range of interacting factors which mutually influence the quality of the generated elevation data of course complicates the analysis of the complete process. However, the data as it is available from the DGPF test provides a suitable scenario to discuss a number of aspects.

2.1. Test scenario and available data sets

In order to independently provide actual data of photogrammetric systems during the DGF test, flight campaigns were realized in the Vaihingen/Enz photogrammetric test site. This site which is established and maintained by the Institute for Photogrammetrie, Universität Stuttgart, consists of approximately 200 signalized and coordinated reference ground points, which are distributed in a 7.4 x 4.7 km² area. In summer 2008 data from 12 different photogrammetric systems were collected and distributed to the network of more than 25 participating institutions. The investigations on photogrammetric generation of elevation data currently concentrate on frame based camera systems. Thus, results for data from the DMC, the Ultracam-X and the quattro DigiCAM will be discussed. For comparison also elevation data based on scanned analogue images from a RMK-Top15 were investigated. In addition to the image data, a LiDAR flight was realized by a Leica ALS 50 scanner at a point density of 5 points/m^2. By these means an additional reference is available to evaluate the results from photogrammetric image processing.

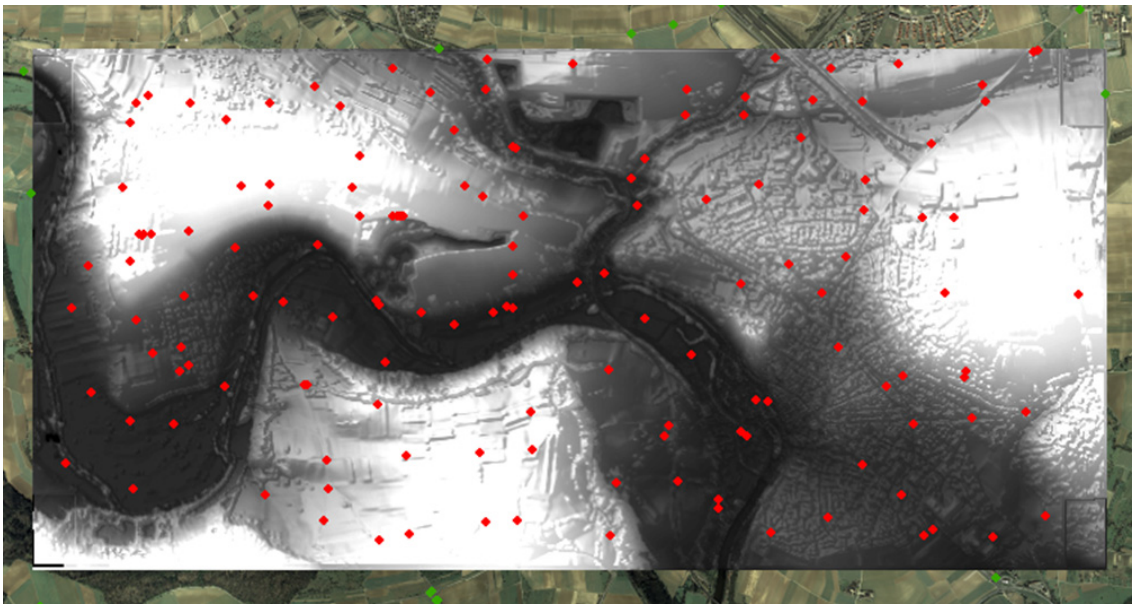


Fig. 1: Test area with collected reference DSM and signalized reference points.

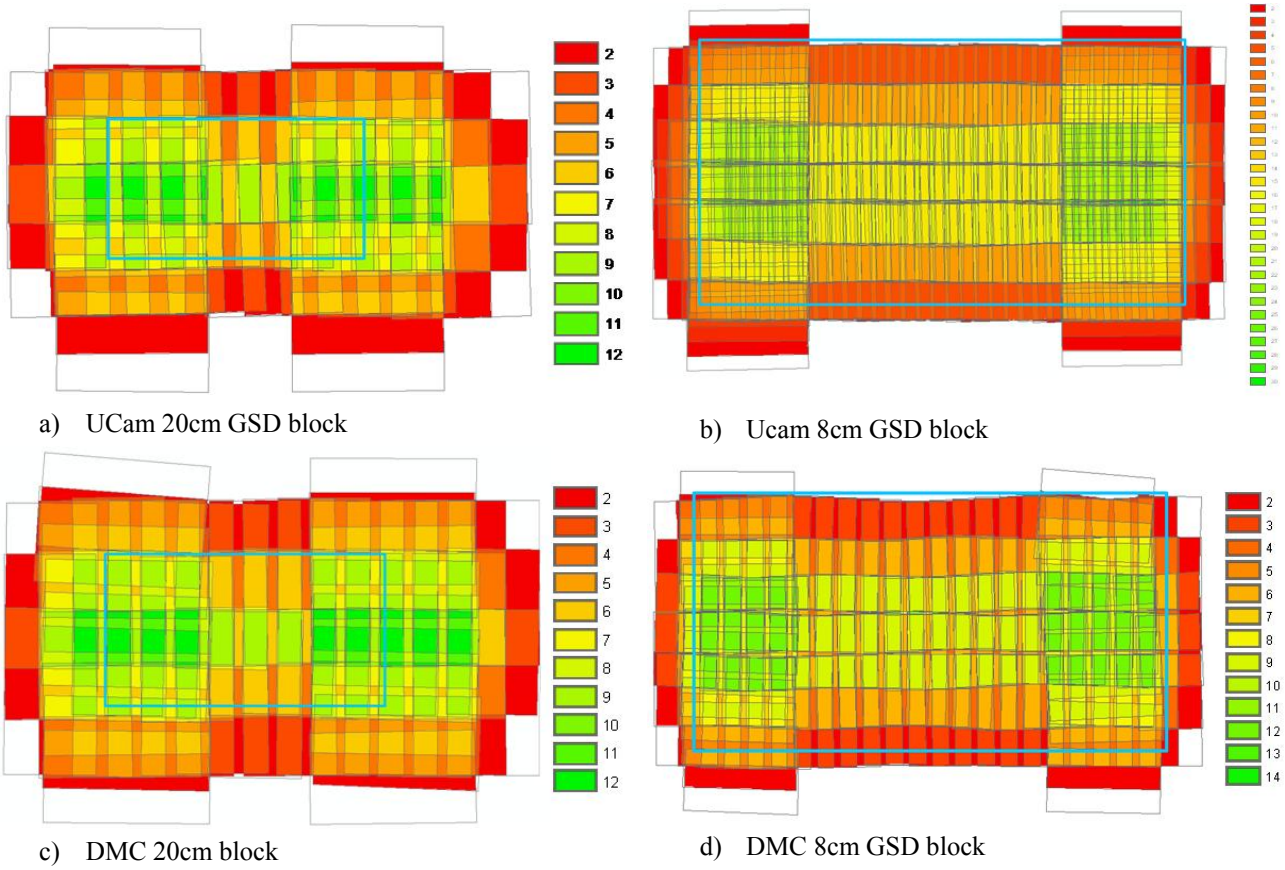


Fig. 2: Color coded image overlap for the blocks from Ultracam-X (top) and DMC (bottom) with test area marked by blue polygon.

Fig. 1 shows the DSM for the test area as derived from the LiDAR data. This was realized at a grid width of 20cm. The available reference points from static GPS measurement are additionally overlaid. As it is visible in Fig. 2 during the test campaign, image blocks of 20 cm GSD and 60% along and cross track as well as blocks with 8cm GSD and 80% in flight and 60% cross flight

overlap were collected. By these means, the influence of different ground sampling distances [GSD] on the quality of computed elevation data can be evaluated. Fig. 2 exemplarily shows the resulting image overlap for the Ultracam-X and the DMC flights. The block configurations for the 20cm GSD flights are depicted on the left, the overlap for the 8cm GSD flights are shown on the right. As it is visible, considerable overlap is available in the given scenario. The test area for the investigations on image matching is marked by the light blue rectangle. This area corresponds to the reference DSM depicted in Fig. 1.

The geometric accuracy of a DSM from image matching is influenced considerably by the quality of the reconstructed image geometry, which is usually provided from bundle block adjustment. Within the DGPF test this step is investigated as an independent topic, thus it is only briefly discussed in the following. Usually, the geometric accuracy from bundle block adjustment is evaluated based on RMS values from coordinate differences at the signalized points. These empirical RMS values were rather similar for all investigated flights. For the horizontal component an absolute accuracy RMS in the range of ¼ pixel and better was achieved both for the 8cm and 20cm GSD blocks. The vertical component resulted in an accuracy of ½ pix and better. The coordinates of the reference points within the test field were determined with static GPS base line observations. Since this provides an accuracy of 1cm for horizontal and 2cm for vertical coordinates, the accuracy of object points from bundle adjustment is already in the same range of accuracy.

2.2. DSM comparison to reference points

During processing of the respective image data considerable improvement of object point quality from bundle block adjustment was feasible by self-calibration. By these means the remaining systematic effects within the images could be modeled by a suitable set of additional parameters. In order to apply these corrections for the preceding DSM generation process so-called absolute orientations were computed in a two step process (Cramer & Haala, 2009). These absolute orientations are based on adjusted object coordinates from the self-calibrating block adjustment which are used as fixed observations in a second step. Thus, the effects from self-calibration are transferred into the orientation elements. During further processing, the application of these orientations for DSM generation helped to avoid interface problems with respect to different software packages and thus simplified the comparison of results from different groups.

For our investigations DSM grids of 0.2m raster width were generated from the 8cm GSD imagery using the software MATCH-T DSM. Additionally, 0.5m raster DSM grids were computed from the 20cm GSD blocks. As a reference a 0.2m raster DSM was derived additionally from the ALS 50 LiDAR point cloud. In order to allow a first evaluation of the resulting DSM qualities, differences between the respective raster surfaces and the available reference points were computed. The results for the respective camera systems DMC, Ultracam-X, quattro DigiCAM and RMK-Top15 and the LiDAR reference measurement are summarized in Table 1. For further analysis of the respective point differences, for each DSM the maximum and minimum values Δ Max/Min are additionally made available in Table 1.

	Sensor	RMS[cm]	Δ Max/Min [cm]		# points
LiDAR-reference	ALS 50	7.2	6.4	-61.0	145
GSD 8cm Raster 0.2m	DMC	7.6	74.1	-24.3	145
	Ultracam-X	4.2	11.7	-10.8	143
	DigiCAM	5.6	15.5	-23.4	145
	RMK	8.8	77.4	-19.9	145
GSD 0.2m Raster 0.5m	DMC	16.4	57.5	-30.5	146
	DigiCAM	10.6	27.1	-41.0	146
	Ultracam-X	8.0	21.3	-30.0	146
	RMK	10.7	37.2	-35.1	146

Table 1: Differences between DSM and all reference points.

Typically, the largest differences for each data set occurred at areas, which were potentially compromised to occlusions. Such a situation of a reference point partially occluded by a car is depicted exemplarily in Fig. 3. In order to eliminate these potential gross errors, a simple threshold was used to filter out points with differences outside a range of $\pm 3 \cdot$RMS. This was sufficient for our investigations, however, more advanced methods are for example described in (Höhle & Höhle, 2009)

	Sensor	RMS [cm]	Mean [cm]	Δ Max/Min [cm]		Elim. Points
LiDAR- reference	ALS 50	3.4	-1.1	6.4	-11.0	3
GSD 8cm Raster 0.2m	DMC	3.9	-0.8	21.1	-0.9	2
	Ultracam-X	4.2	-1.4	11.7	-10.8	0
	DigiCAM	5.3	-1.1	15.5	-15.7	1
	RMK	5.2	2.4	15.6	-19.9	2
GSD 20 cm Raster 0.5m	DMC	15.7	-9.3	36.9	-30.5	1
	DigiCAM	10.1	-0.1	27.1	-30.5	1
	Ultracam-X	7.6	0.7	21.3	-17.9	1
	RMK	9.9	1.4	31.8	-25.9	2

Table 2: Differences between DSM and reference points after elimination of gross errors.

Based on the remaining reference points the RMS values of the respective DSMs were then recomputed. These results summarized in Table 2, where the number of eliminated point differences is given in the last column. After filtering, the resulting RMS value for the LiDAR DSM measured by the ALS 50 sensor is 3.4cm, which is almost in the order of the vertical accuracy of the used check points. Compared to this accuracy, the RMS values of the DSMs for the DMC, Ultracam-X and DigiCAM are only slightly larger.

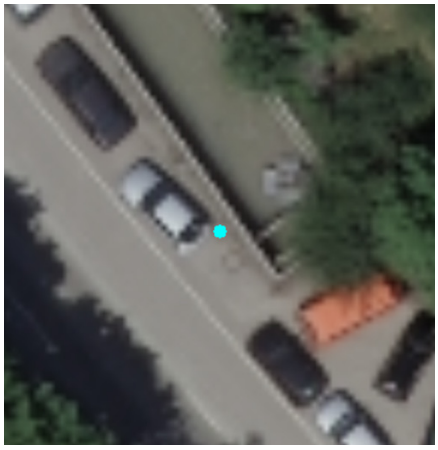


Fig. 3: Reference point.

As it is visible in Table 2, for the 8cm GSD blocks the values are rather similar for the different camera systems. The results of these highly overlapping blocks also correspond very well to the vertical component of the preceding block adjustment which gave an accuracy of ½ pix. The same holds true for the results from the 20cm GSD blocks, where again the quality of bundle block adjustment seems to be the most dominating factor. For interpretation of these results, it still has to be mentioned, that similar to the exemple in Fig. 3, the ground control points were typically installed at paved areas like small roads or parking lots. If not occluded, such flat neighborhoods are of course beneficial for the filtering and interpolation process during generation of the DSM raster in that area. For this reason differences between respective DSMs can be computed to allow for a more area covering analysis in addition to the singular elevation values at the available reference points.

Fig. 4: Differences between DSM from DMC 8cm GSD image matching and LiDAR (left), corresponding ortho (right).

The difference between the DSMs from DMC 8cm GSD image matching and the LiDAR measurement for a part of the test area is given in the left image of Fig. 4, while the right part of Fig. 4 shows the corresponding section of the ortho image. As it is visible, the differences between both surfaces are rather small and mainly correspond to vegetated areas. Due to the time gap of four week between the DMC and the LiDAR flight this indicates that these differences 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. This is especially true for the relatively larger differences which occur at trees.

2.3. Evaluation of 3D point clouds

In addition to automatic point transfer by image matching, suitable algorithms for filtering and interpolation are required during DSM raster generation. For this reason, 3D point clouds were generated as additional output to allow for investigations of the respective matching accuracy while avoiding the influence of such interpolation processes. In order to evaluate the accuracy of the generated 3D point clouds a sports field is determined as planar test area. After estimation of a best fitting plane, the geometric accuracy of these points can be determined based on the remaining vertical distances. An additional indication for the matching quality is available from the corresponding point density.

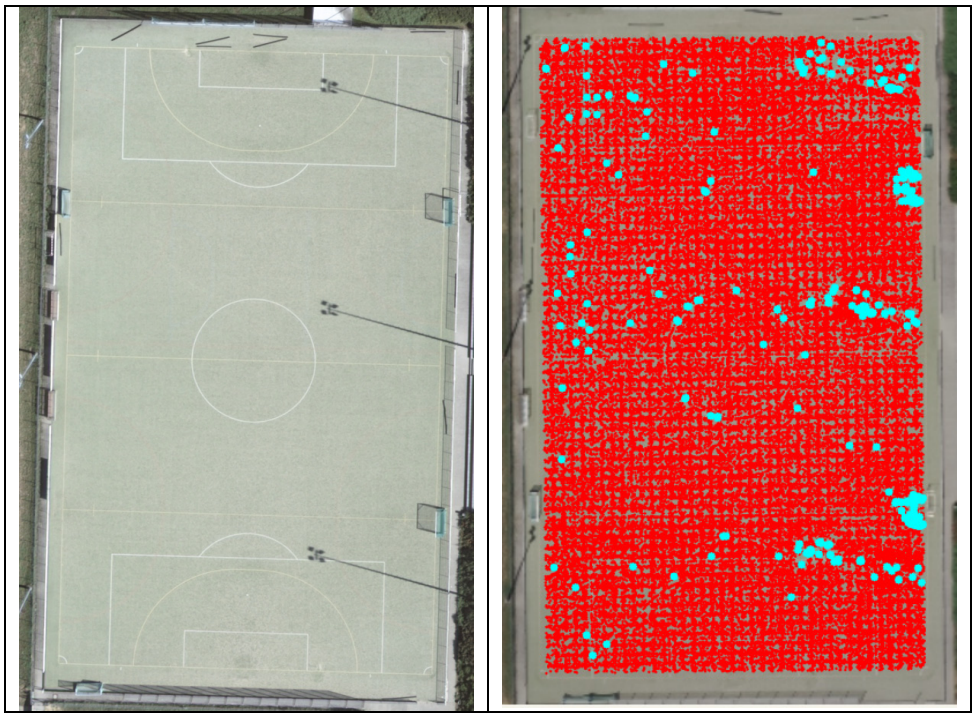
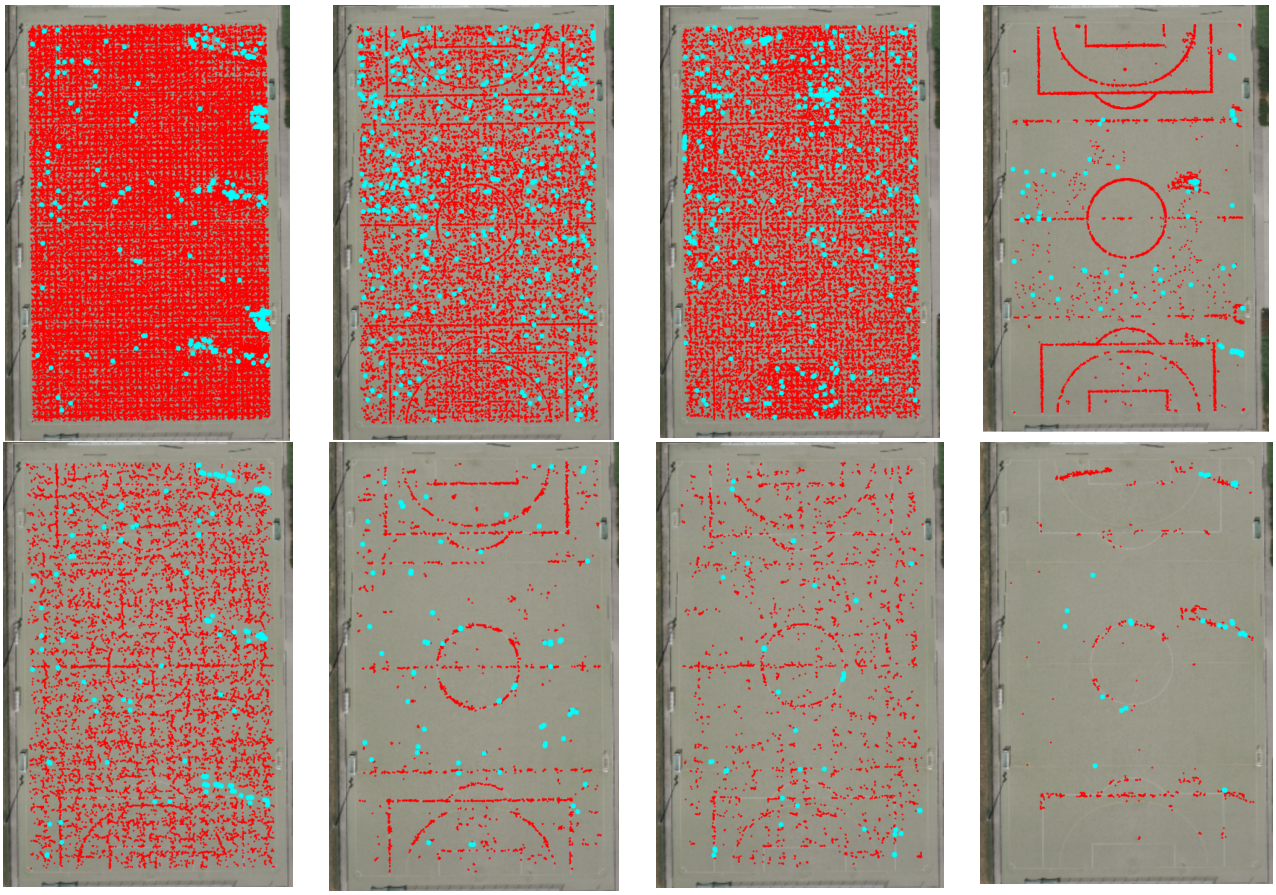


Fig. 5: DMC image (left) and ortho image with cloud overlaid (right).

An example of this process for GSD 8cm imagery from the DMC camera is given in Fig. 5. On the left an image of this block is shown, while the generated 3D point cloud is overlaid to the corresponding ortho image on the right. For this test area a point density of 22.41 Pts/m^2 was generated by image matching. The vertical distances to a best fitting plane result in an RMS value of 9.7cm for the respective point measurements. Similar to the investigation presented in section 2.2. , a simple threshold can then be used to eliminate points outside a range of ±3·RMS. Within Fig. 5 (right) these potential gross errors are marked in light blue. As it is visible, they 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.



a) DMC b) Ultracam-X c) DigiCAM d) RMK

Fig. 6: Generated point clouds for investigated camera systems. The top row shows the results for the 8cm GSD block, the bottom row for the 20cm GSD block.

	Sensor	RMS [cm] after filter	RMS [cm] no filter	Elim. Pts [%]	Density [Pts/m^2]
	LiDAR	1.47	1.50	0.33	11.58
GSD 8cm	DMC	3.19	5.94	0.46	22.31
	UltracamX	7.14	7.46	0.74	15.96
	DigiCAM	6.24	6.94	0.52	22.67
	RMK	6.85	19.86	0.63	5.71
GSD 20cm	DMC	10.46	15.10	1.09	2.65
	Ultracam-X	11.58	13.21	1.68	1.21
	DigiCAM	20.70	22.14	1.00	1.44
	RMK	35.27	43.97	1.48	0.34

Table 3: Results of plane approximation from 3D point clouds at sports field test area.

Similar to the results for the DMC camera already presented in Fig. 5, point clouds were generated for the remaining frame based camera systems quattro DigiCAM, Ultracam-X and RMK-Top15. Again the software MATCH-T DSM was used. The respective point clouds are depicted in Fig. 6. Results for the 8cm GSD blocks are shown in the top row, while the matching results from the

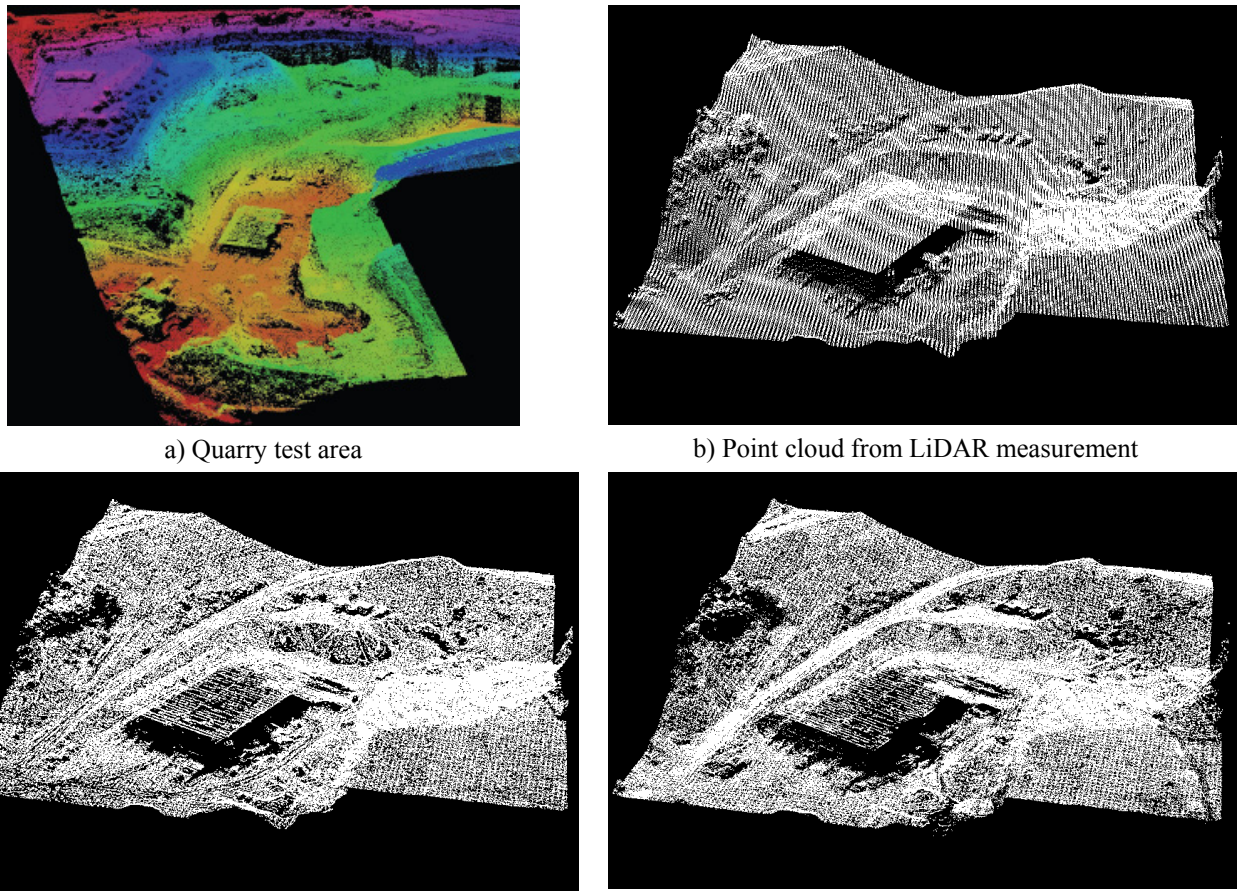
20cm GSD blocks are presented in the bottom row. Again, points eliminated by the filter process are marked in light blue, while the remaining points are shown in red. Supplementary to Fig. 6, the results of point cloud analysis for all investigated flights are summarized in Table 3. On average, a point density of about 20 pts/m² was reached using the GSD 8cm images from the digital camera systems, while matching of scanned RMK images gives less than 1 pt/m². 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. 6 and Table 3 show a considerable advantage of point matching for the GSD 8cm blocks compared to the GSD 20cm blocks for all digital camera systems. The point density using the GSD 8cm images from the digital camera systems is even higher than the approximately 10 pts/m², which were generated by the ALS 50 laser scanner at the sports fields. However, the standard deviation for the LiDAR data is better than 2cm, almost without any gross errors, while an average of 5.5cm for the filtered points is achieved from image matching. Thus, for the 8cm block an accuracy of below 1 pixel GSD was achieved for the single point measurements. For the GSD 20cm this value is slightly worse with an average standard deviation of 14.12cm for the digital cameras. Compared to the 8cm GSD block, the average point density of 1.8 pts/m² 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, Ultracam-X and quattro DigiCAM 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 noontime. Together with the variations of the block geometry these differences considerably influence the results as available from the digital camera systems.

3. USABILITY OF DSM DATA FROM IMAGE MATCHING

Despite the fact that LiDAR still provides data of highest quality, the results presented in the preceding section clearly indicate, that a considerable number of applications will be feasible based on height data from image matching, if digital airborne cameras are used. As an example, investigations on the DSM generation in agricultural areas like vineyards and orchards using the software NGATE are presented by Hasted & Ginzler (2009). Wolff (2009) shows results on point cloud and DSM raster generation in urban areas using the software SAT-PP. Another example for a test area of relatively high geometric complexity is given in Fig. 7. Similar to the data sets presented in section 2.3. point clouds were generated. This test area at a quarry was defined together with other regions at residential, industrial and agricultural areas during preparation of the DGPF test in order to allow for investigations at areas of different image texture or surface complexity.



a) Quarry test area

b) Point cloud from LiDAR measurement

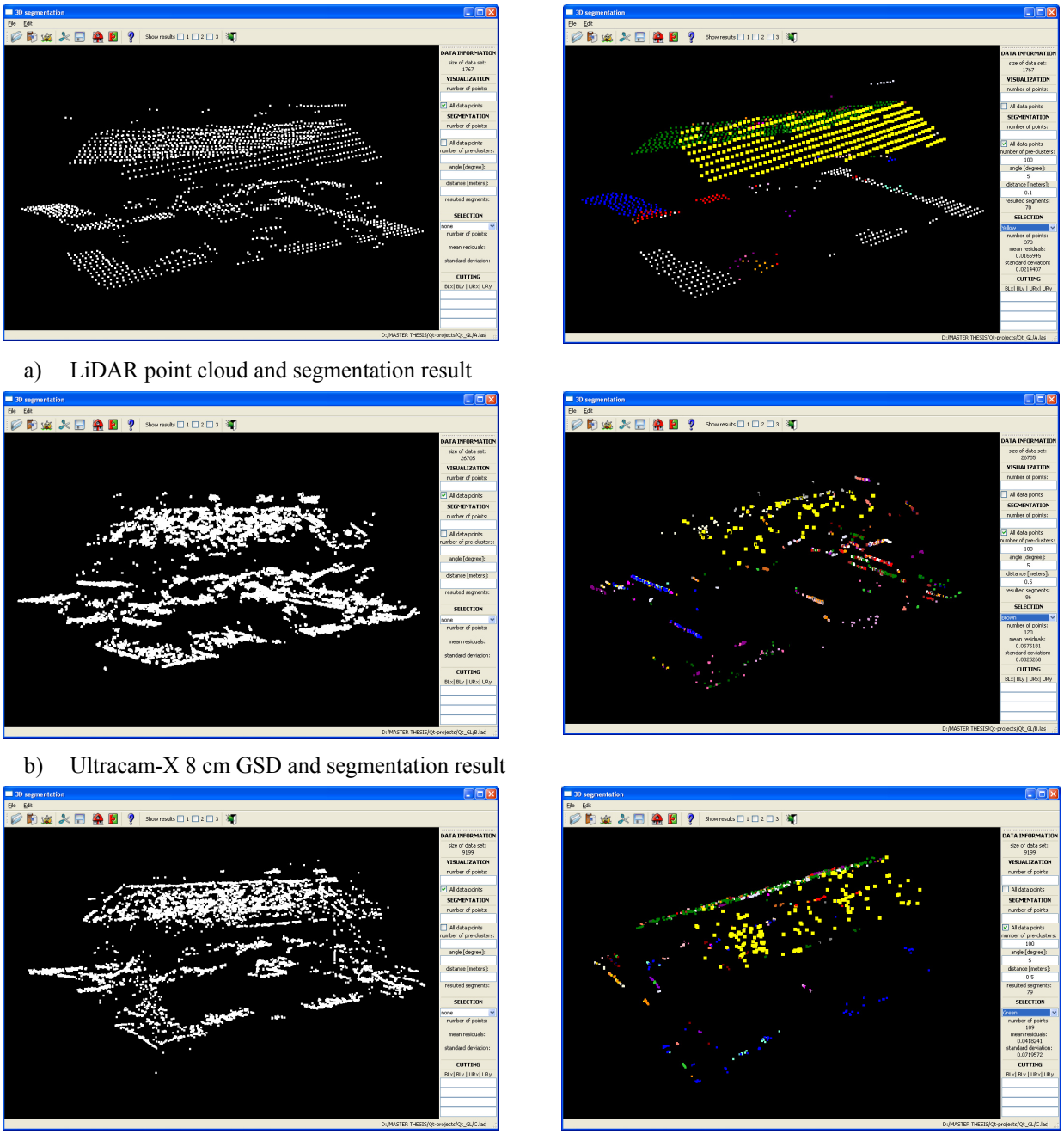
c) Point cloud from Ultracam-X, 8 cm GSD

d) Point cloud from DMC, 8 cm GSD

Fig. 7: Quarry: Original area and cut out data sets.

Within Fig. 7 the complete area and a smaller section from LiDAR measurement is depicted in the top row. The bottom row shows point clouds from matching of the 8cm GSD data, which were again computed by the MATCH-T DSM software. The bottom left depicts points derived from the Ultracam-X imagery, while bottom right shows results from the DMC. As it is visible, even relatively complex surface geometries can be represented at an accuracy and amount of detail, which is sufficient for a number of applications.

Frequently, the further analysis and interpretation of such 3D point cloud aims at tasks like 3D object reconstruction. Within this context, a segmentation of the respective point cloud is usually realized as a first processing step. Thus, in order to evaluate the feasibility of elevation data from image matching for further interpretation like point cloud segmentation a standard region growing approach was implemented by Jarzabek (2009). In this algorithm, the n-nearest neighbors of each point are used to estimate a local best fitting plane. In addition to normal vector computation, points with corresponding local planes of small residuals are then selected as seed regions. Within the following region growing step, points adjacent to this seed region are iteratively added to the plane if they are compatible to the actual plane. In order to fulfill this compatibility criterion the vertical distance to the plane and the normal vector difference has to be beyond a certain threshold.



a) LiDAR point cloud and segmentation result

b) Ultracam-X 8 cm GSD and segmentation result

c) DMC with 8 cm GSD and segmentation result

Fig. 8: Point clouds derived from different sources (left images) and their colour coded segmentation results (right images).

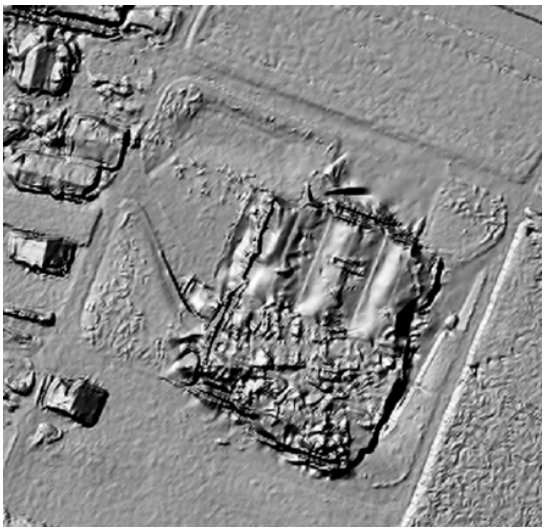
Results of this segmentation process are shown in Fig. 8. While for the top row measurements from the ALS 50 laser scanner were used, the second and the third row provide results for point clouds, which were derived from the 8cm GSD images of the Ultracam-X and the DMC camera, respectively. Within Fig. 8 the original point clouds are depicted on the left, while the results from the segmentation process are presented on the right. There, points which are grouped to the same plane are represented by corresponding color values. During segmentation the most plausible results are obtained from processing airborne laser scanning data. Meaningful entities are generated since the roof planes are detected and clearly distinguished. Moreover, the residuals of the estimated best fitting planes from LiDAR measurement are superior to the result obtained by Ultracam-X and

DMC. In case of point clouds derived from multiple image matching, the segmentation gives a relatively poor performance for the investigated scenario. However, the implemented region growing approach is tailored for point sets that feature some level of order. Such even distribution or regular spacing is usually available from LiDAR measurement. Thus this type of data gives the most plausible results. In contrast, point clouds from image matching are denser at the edges than at the homogenous surfaces since MATCH-T DSM implements a feature based matching approach. The segmentation of such point clouds by standard region growing therefore frequently results in relatively small parts of the respective objects.

a) Ortho image

b) DSM from LiDAR measurement

c) DSM from DMC 8 cm GSD



d) DSM from RMK 8 cm GSD

Fig. 9: Comaprison of shaded DSM from different data sets.

Despite the fact, that for our segmentation LiDAR measurement is still superior, the advances of digital airborne camera systems compared to scanned analog images for matching is obvious. As it is demonstrated in Fig. 9, this is especially true for regions with limited surface texture. The bottom left picture of Fig. 9 shows a shaded DSM from image matching based on the DMC 8cm GSD block. The corresponding result for the scanned RMK data is depicted on the bottom right. As already mentioned, this comparison is especially interesting, since both image blocks were captured simultaneously at identical atmospheric conditions. For comparison, the top left image of Fig. 9

additionally shows the corresponding ortho image, while the shaded DSM from LiDAR measurement is depicted in the top right.

4. CONCLUSIONS

Within the paper the benefits of digital image recording for elevation data generation by image matching were demonstrated. Recent developments in sensor and software technology 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. Compared to this technique, results from image matching still are more error-prone, i.e. due to problems like moving shadows and still provide results of partly varying geometric quality. Despite these limitations, the current performance of digital image matching enables a number of standard applications like DSM and DTM generation at sufficient quality. Remaining challenges to ameliorate the further use of elevation data from image matching are a further improvement of filter approaches but also by an optimal adaption of algorithms for surface interpretation and object reconstruction. Additionally, the full use of jointly collected high resolution radiometric and geometric information for the collection of detailed geo-data 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.

5. REFERENCES

Cramer, M. & Haala, N. (2009): DGPF Project: Evaluation of Digital Photogrammetric Aerial Based Imaging Systems – Overview and Results from the Pilot Centre. ISPRS Hannover Workshop 2009.

Cramer, M., Krauβ, H., Jacobsen, K., von Schönermark, M., Haala, N. & Spreckels, V. (2009): Das DGPF-Projekt zur Evaluierung digitaler photogrammetrischer Kamerasysteme. DGPF Jahrestagung Band 18.

DeVenecia, K., Walker, S. & Zhang, B. (2007): New Approaches to Generating and Processing High Resolution Elevation Data with Imagery. Photogrammetric Week ’07, pp. 297-308.

Hastedt, H. & Ginzler, C. (2009): Generierung digitaler Oberflächenmodelle im DGPF-Projekt unter Verwendung von NGATE − Erste Ergebnisse, CD-Publikation, DGPF Jahrestagung Band 18.

Hirschmüller, H. (2008): Stereo Processing by Semi-Global Matching and Mutual Information. IEEE Transactions on Pattern Analysis and Machine Intelligence, 30 (2), pp. 328-341.

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, pp. 398-406.

Jarząbek, M. (2009): Segmentation and presentation of multiple 3D point clouds, Master Thesis, Institute for Photogrammetry, Universität Stuttgart.

Lemaire, C. (2008): Aspects of the DSM Production with High Resolution Images IAPRS, Volume XXXVII, Part B4, S. 1143-1146.

Wolff, K. (2009): DGPF Project: Evaluation of Digital Photogrammetric Aerial Based Imaging Systems – Generation of Digital Surface Models, ISPRS Hannover Workshop 2009.