Experiences on Automatic Image Matching for DSM Generation with ADS40 Pushbroom Sensor Data

M. Pateraki, E. Baltsavias, U. Recke

1 Institute of Geodesy and Photogrammetry, ETH-Hoenggerberg, CH-8093, Zurich, Switzerland (maria, manos)@geod.baug.ethz.ch
2 Leica Geosystems GIS & Mapping GmbH, CH-9035 Heerbrugg, Switzerland utz.recke@gis.leica-geosystems.com

Commission II, IC WG II/IV

Key Words: pushbroom, sensor, systems, matching, surface, quality, analysis, performance

Abstract:

This paper presents an analysis on automatic image matching for DSM generation with ADS40 airborne push-broom sensor data. ADS40, produced by Leica Geosystems GIS & Mapping (LGGM), offers on a single camera system the possibility to acquire both panchromatic and multispectral images in up to 10 channels (100% overlap). The matching performance is done through a qualitative and quantitative analysis in test areas in Switzerland and Japan, of different types of buildings and roofs. Results are analyzed and compared with manually measured reference data.

1. Introduction

1.1 State-of-the-art and Proposed Evaluation

Single- and multi-line CCDs are employed as research tools in photogrammetric and remote sensing applications. Regarding airborne sensors, several systems have been developed and among them fewer commercial ones, e.g. ADS40 (LGGM), DMC (Z/I Imaging), Starimager (Starlabo), Ultracam-D (Vexcel). New methods, compared to the existing ones for processing of scanned aerial films, are necessary for digital sensors, especially line-CCD-based ones as they have significant differences to the existing film-based cameras, e.g. several (up to 9) CCD-lines with 100% overlap, a non-perspective geometry in flight direction, different radiometric characteristics, simultaneous multispectral imaging capabilities, more complicated imaging geometry and integration of GPS/INS systems for determination of the position and orientation of each line. Investigations regarding airborne linear-CCDs have been already performed regarding camera architecture, direct georeferencing, sensor modeling, ground processing and aerial triangulation for ADS40 (Fricker, 2001; Hinsken et al., 2002; Sandau et al., 2000, Tempelmann et al., 2000; Tempelmann et al., 2003) and for other systems (Fritsch, 1997; Haala et al., 2000; Hoffmann et al., 2000; Leberl et al., 2003; Tianen et al., 2003; Wewel et al., 1998). Notably fewer studies exist on matching methods and DSM generation using airborne linear CCDs (Gwinner et al., 1999; Neukum, 1999; Renouard and Lehmann, 1999; Scholten, 2000). Recently, in the dedicated workshop on digital aerial cameras during the Optical 3D Measurement Techniques conference in 2003, only the papers of Nonin (2003), Pateraki and Baltsavias (2003b) and Zhang and Gruen (2003) have been focused on automatic DSM extraction and only the last two on algorithmic aspects. Regarding digital photogrammetric stations, advancements in matching for tie point extraction and DSM generation have generally not been integrated in the current systems, although images of different sensor orientation models, employing line CCDs, can be imported. To date the algorithms used traditionally for the processing of aerial frame imagery are employed also for airborne linear CCDs.

The aim of this paper is to evaluate the matching performance on ADS40 images, using the commercial SOCET SET digital photogrammetric software package (from this point forward called SS) and the matching software that has been developed at ETH Zurich, adapted to ADS40 characteristics (from this point forward called AIM). However, as each system employs a different matching strategy (see Section 2), the analysis is focused on the quality of the final product, the DSM respectively. The quality of the extracted terrain is evaluated in different areas of terrain relief and land-cover, using manually measured reference data.

1.2 ADS40 and Sensor Plate Design

To date the ADS40 camera system and architecture has been described in several publications, as mentioned already in Section 1.1. In brief, ADS40 consists of seven parallel CCD lines in the focal plane of a single lens system – three panchromatic (forward- nadir, backward looking), red, green and blue placed next to each other and near infrared. In addition, ADS40 incorporates GPS and INS technology (Applanix) for direct georeferencing.
In this study, datasets have been acquired with two different sensor plate (SP) designs. The standard and most known configuration applies to SP1 design with staggered panchromatic lines in backward (PANB), forward (PANF) and nadir (PANN) position, the multispectral red, green and blue lines (RGB) between forward and nadir lines and the near infrared line (NIR) close to nadir. In the SP2 design, the RGB triplet is set at the nadir position (preferred for true orthophoto generation), two panchromatic lines in forward viewing (PANF), one panchromatic line for backward viewing (PANB) and one near infrared line between nadir and forward. In SP2, the staggered mode for the PAN CCDs is not used. In Figure 1, the SP1 and SP2 designs are illustrated and in Table 1, the respective viewing angles of the CCDs are listed.

![Figure 1. Configuration of the CCD lines on the two different sensor plate designs, SP1 (a) and SP2 (b). The panchromatic lines forward, nadir and backward are indicated as PANF, PANN and PANB respectively, near infrared as NIR and the color triplet as RGB.](image)

<table>
<thead>
<tr>
<th>CCD lines</th>
<th>SP1</th>
<th>SP2</th>
</tr>
</thead>
<tbody>
<tr>
<td>PAN</td>
<td>+28°, 0°, -14°</td>
<td>+28°, +16°, -14°</td>
</tr>
<tr>
<td>RGB</td>
<td>+16°</td>
<td>0°</td>
</tr>
<tr>
<td>NIR</td>
<td>-2°</td>
<td>+14°</td>
</tr>
</tbody>
</table>

Table 1. Viewing angles of the CCDs for the two different focal plate designs.

The acquired ADS40 channels are further rectified onto a height plane (Lev1 images) in order to be used for stereo viewing and in automatic matching processes for tie point and DTM/DSM extraction. In Lev1 images, differences between channels due to scale, but also rotation and shear that exist in raw images (Lev0 images) are removed to a large extent. However, the option of utilizing Lev0 images in aerial triangulation (AT) is being investigated (not handled in this paper), in order to accelerate a part of the ground processing chain (rectification and automatic tie point extraction).

2. DATASETS AND SYSTEMS

2.1 Datasets

Two ADS40 datasets were used in these investigations. The first was acquired over the rural area of Waldkirch area in Switzerland and the second over the dense city center of Yokohama in Japan. The Waldkirch block was flown with SP1 camera by LGGM in May 2002, and consisted of 4 parallel and 2 cross strips. The coordinate system used was the WGS84. The block of Yokohama consisted of three parallel strips and was flown by Pasco Corp. The coordinate system used was the Japanese grid. Both datasets included panchromatic and multispectral imagery in Lev0 (raw) and Lev1 product with 0.20 m ground sampling distance.

In terms of radiometric quality, the Yokohama compared to the Waldkirch dataset exhibited higher noise. Interpretability of objects was more difficult, as denser and higher buildings existed in combination with strong shadows (in many cases saturated) and poorer radiometric quality (Fig. 2). All images used for DSM extraction have been pre-processed in order to reduce noise, improve feature definition and minimize radiometric differences among channels (Pateraki and Baltsavias, 2003a). This part was essential, notably for the Yokohama dataset, in order to help matching in shadowed areas (Fig. 3).

Regarding geometric quality, each individual camera with SP1 and SP2 has been calibrated over a test field with precisely measured control and check points, and interior orientation and IMU misalignment parameters have been estimated. These have been later used in AT which was carried out for both datasets, using ORIMA software, in order to adjust and refine the orientation parameters acquired from the GPS/INS systems on board. Tie points were automatically measured using Automatic Point Matching (APM) module of SS software. Blunders could be visually controlled and iteratively eliminated. GCPs in the Waldkirch dataset were distributed at the block corners, whereas for Yokohama at the block center. The derived geometric accuracy from bundle adjustment in terms of sigma a–posteriori was 2.5 μm for the Waldkirch dataset. For the Yokohama dataset the geometric accuracy was lower, due to the poorer radiometric quality (more blunders in automatic tie point measurement), small errors in the recordings of the GPS and the poorer block geometry (no cross strips). Table 2 summarizes acquisition and bundle adjustment parameters of the two datasets.

![Figure 2. Original image quality of Waldkirch (top) and Yokohama (bottom) dataset.](image)
In SS, the adaptive method or AATE (Adaptive Automatic Terrain Extraction) was used. Adaptive matching can use more than two images, can generate regular grids or triangulated irregular networks (TINs), changes some of the strategy parameters based on the grid matching approach utilised in SS (no interpolating of surface discontinuities. The TIN method is inherently based on the grid matching approach utilised in SS (no interpolation is performed at the last stage, for grid points that have not been successfully matched).

The AIM method is based on a combination of area and feature based matching techniques. Different types of primitives (area patches, single edgels that belong to contours, edges) are combined based on the type of the terrain (rugged, steep, flat). However, since AIM is still an experimental system under fine-tuning, several parameters are set by the user, according to the area and terrain type. The description of the algorithmic approach exists already in the literature (Pateraki and Baltsavias, 2003b), and below only a brief overview of AIM is given. Two types of matching strategies can be utilized, namely single and multi-template strategy. The first is applied in case of relative flat terrain, whereas the second in more complex areas. Multi-resolution levels are employed in a doublet approach (Pateraki and Baltsavias, 2003a) in order to acquire approximate values. More than two images are matched simultaneously, geometrical constraints are enforced by means of quasi-epipolar curves, and 3D position is computed only from the good rays, following correlation and blunder detection. In the upper levels, a surface approximation is derived by matching of grid points (favorable for faster processing) and which is subsequently refined in the lower levels by inclusion of linked contour points. Initial positions at each level are derived by a multi-patch approach, utilizing cross-correlation and three masks of different size. Least squares matching (LSM) with geometrical constraints is further used for verification and refining the matching solution and is applied for straight edges and single points (edgels and grid points). The main reason for extending LSM to straight edges is to improve modeling of discontinuities and minimize surface smoothing (Pateraki and Baltsavias, 2004).

As automatic matching in each system is based on different strategies, the assessment is focused on the quality of the final product, the DSM respectively. Alternatively, an analysis on a different level, namely forcing the systems parameters to be relatively similar, would not be realistic for SS as it has certain limitations for full control of the matching strategy and blunder detection. For AIM, modifications would be feasible in terms of implementation, to a certain extent (to adapt some of its parameters to the ones of SS, e.g. using area-based grid matching).

However, this would be less favorable as AIM was applied in case of relative flat terrain, whereas the second in more complex areas. Multi-resolution levels are employed in a doublet approach (Pateraki and Baltsavias, 2003a) and uses different primitives for an optimal matching strategy, in contrast to SS. In both cases, the pyramid levels and the initial mask sizes have been set to equal values and the same number of images has been used. The three stereo panchromatic channels and the Green channel taken out of one strip have been used as input in all systems. In SS, the TIN version without additional filtering (elimination of trees/buildings/other objects) has been used. Similarly for AIM, additional smoothing has been excluded and raw matched data have been used in the analysis (irregularly distributed points).

Table 3 lists the basic strategy parameters used in each system.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>SS</th>
<th>AIM</th>
</tr>
</thead>
<tbody>
<tr>
<td>No pyramid levels/matching passes</td>
<td>6/8</td>
<td>6/6</td>
</tr>
<tr>
<td>No of images</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Type of matching</td>
<td>Image matching using two “best” images</td>
<td>Image matching all images simultaneously</td>
</tr>
</tbody>
</table>

Table 3. Matching parameters.

2.3 Reference Data

In order to check the matching accuracy, reference datasets were derived from ADS40 images. Mass points and breaklines have been manually collected in stereo mode in SS with an estimated accuracy of 20-40 cm and 40-60 cm for Waldkirch and Yokohama dataset respectively. The same orientation used later for
the automatic extraction was used also for the manual measurements. Mass points have been separated in two classes: ground points close to the perimeter of buildings and ground points defining the bare earth (BE) surface excluding points close to buildings. This classification is justified as errors arising from matching should be treated separately from modeling errors of the surface in 2.5D. Breaklines have been extracted along discontinuities on the ground and on buildings or man made objects (MMO), even small features on roofs. Points along the breaklines have been interpolated every 0.20 m, corresponding to approximately 1 pixel in image space. Forest areas and trees have been excluded from the measurements as it was difficult to extract reliable 3D coordinates, due to their imprecise and unclear shape. Moreover, due to the radiometric quality of the Yokohama dataset, many roofs, especially the smaller ones lying partly in the shadowed areas, could not be accurately extracted and for several manually measured points the estimated accuracy degraded (~ 60 cm). Roadmarks were not always easily visible due to the shadowing. Less points could be extracted in open surfaces and close to building outlines and therefore the class used to analyze separately modeling errors was not used.

3. RESULTS OF IMAGE MACHING AND QUALITY EVALUATION

DSM extraction was tested in selected regions, excluding large shadowed areas (especially in the Yokohama dataset). As the Waldkirch region had more variations in land cover, three regions have been selected. The relief in area W1052A (Fig. 4 (a)) was undulating and sparse houses, trees and small salt lakes existed. The area W1052B (Fig. 4 (a)) was relatively flat with good texture in agricultural fields. The area W1052C (Fig. 4 (c)) included complex objects, well defined roads, single and groups of trees, in many cases close to the houses. One representative region has been selected in the Yokohama dataset (Fig. 5) and it included high buildings (some over 20 m high) and thus large discontinuities and occlusions. In combination with the low radiometric quality, the degree of difficulty in extracting a reliable DSM of this area increased. The parameters of each system have been modified according to the characteristics of each area. As mentioned already in Section 2.2, in SS the AATE method was used in all tests in combination with the TIN option. For the Waldkirch test areas W1052A and W1052C, 1 meter grid interval in object space was used and for W1052B, as the terrain was less undulating, the grid interval was set to 1.5 m. For the Yokohama test area a denser matching was employed with 0.5 m grid spacing due to the density of buildings and large parallax differences. In AIM, the strategy of each area varied. Single and multiple template strategies were used for areas W1052A, W1052B and Y0624C, W1052C respectively. An approximate surface in the upper pyramid levels was generated by matching of grid points (5 m interval), and further refined by matching of edges in the lower levels. However, for area W1052B edge matching was used only in the two lower levels as the area was less undulating.

The results of the quantitative analysis, performed for the above areas, are shown in Table 4. The raw DSM points without any postprocessing (e.g. filtering, modeling of breaklines) and the interpolated elevations of the reference-measured points were utilized to compute statistics of elevation differences (RMS, mean with sign, absolute maximum and standard deviation).

Figure 4. Test areas in the Waldkirch dataset. The size of W1052A (a), W1052B (b) areas was approximately 2000 x 2000 pixels and of W1052C (c) area 2500 x 1800 pixels.

Figure 5. Part of test area Y0624C in the Yokohama dataset. The total size of the area is 1000 x 3000 pixels.

Errors with respect to points on bare earth (BE) and on man made objects (MMO) are analyzed separately (Pts, Brklin). The number of single points on BE was less than the interpolated breakline points and the first were measured in locally flat areas. Modeling errors are separately computed as error statistics from points that lie close to the buildings and on the ground (hPts). In general, AIM delivered more accurate results than SS in all test areas. In relative flat areas (W1052B), less differences in the accuracy could be observed. Blunders in DSMs derived with SS were significantly higher, in locally flat areas and in the case of MMO but also at breaklines. AIM shows a better performance in case of discontinuities, as accuracy increased by approximately two times and blunders were less than SS for all areas in the case of MMO. For both systems accuracy degraded for points close to building outlines, as less points could be automatically derived in these areas. Here, elevation was interpolated from the nearest points, which apparently were either on the ground or on buildings, leading thus often to large errors. In addition, planimetric errors, even if small, can also cause at surfaces discontinuities large height errors. SS has generated more points than AIM in all areas except W1052B. The reason for this is the inclusion of edges with AIM, leading to extraction of many points in highly-textured areas, e.g. the agricultural fields. In the case of Yokohama, as one could expect, accuracy degraded more, as the geometric and radiometric quality of the data was poorer and the acquisition of the reference data less accurate, compared to Waldkirch. Still, AIM delivered better results, especially along discontinuities on MMO, with an RMS...
<table>
<thead>
<tr>
<th>Test data/ number of matched points [SS/AIM]</th>
<th>Type of points/number of comparison points</th>
<th>System</th>
<th>RMS (m)</th>
<th>Mean with sign (m)</th>
<th>Maximum absolute (m)</th>
<th>Std. dev. (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>W1052A/ [292157/166009]</td>
<td>BE {Pts}/40</td>
<td>SS</td>
<td>1.29</td>
<td>-0.82</td>
<td>3.26</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AIM</td>
<td>0.56</td>
<td>-0.22</td>
<td>1.00</td>
<td>0.52</td>
</tr>
<tr>
<td></td>
<td>BE {Pts, Brklin}/2500</td>
<td>SS</td>
<td>0.95</td>
<td>-0.75</td>
<td>2.80</td>
<td>0.60</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AIM</td>
<td>0.70</td>
<td>-0.49</td>
<td>1.69</td>
<td>0.44</td>
</tr>
<tr>
<td></td>
<td>BE {hPts}/43</td>
<td>SS</td>
<td>2.27</td>
<td>-1.90</td>
<td>4.78</td>
<td>1.26</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AIM</td>
<td>2.22</td>
<td>-1.56</td>
<td>3.47</td>
<td>0.90</td>
</tr>
<tr>
<td></td>
<td>MMO {Brklin}/1000</td>
<td>SS</td>
<td>1.20</td>
<td>0.09</td>
<td>4.11</td>
<td>1.18</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AIM</td>
<td>0.66</td>
<td>-0.02</td>
<td>2.32</td>
<td>0.66</td>
</tr>
<tr>
<td>W1052B/ [141512/342692]</td>
<td>BE {Pts}/30</td>
<td>SS</td>
<td>0.86</td>
<td>0.64</td>
<td>1.84</td>
<td>0.68</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AIM</td>
<td>0.58</td>
<td>0.34</td>
<td>1.17</td>
<td>0.50</td>
</tr>
<tr>
<td></td>
<td>BE {Pts, Brklin}/1100</td>
<td>SS</td>
<td>0.37</td>
<td>-0.10</td>
<td>1.84</td>
<td>0.33</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AIM</td>
<td>0.45</td>
<td>-0.36</td>
<td>1.27</td>
<td>0.30</td>
</tr>
<tr>
<td>W1052C/ [736749/283956]</td>
<td>BE {Pts}/37</td>
<td>SS</td>
<td>0.72</td>
<td>0.43</td>
<td>2.25</td>
<td>0.59</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AIM</td>
<td>0.55</td>
<td>0.33</td>
<td>1.09</td>
<td>0.45</td>
</tr>
<tr>
<td></td>
<td>BE {Pts, Brklin}/24840</td>
<td>SS</td>
<td>1.38</td>
<td>0.19</td>
<td>15.05</td>
<td>1.36</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AIM</td>
<td>0.89</td>
<td>0.24</td>
<td>7.42</td>
<td>0.86</td>
</tr>
<tr>
<td></td>
<td>BE {hPts}/334</td>
<td>SS</td>
<td>1.79</td>
<td>-0.26</td>
<td>19.24</td>
<td>1.77</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AIM</td>
<td>1.23</td>
<td>-0.46</td>
<td>4.66</td>
<td>1.14</td>
</tr>
<tr>
<td></td>
<td>MMO {Brklin}/10000</td>
<td>SS</td>
<td>0.96</td>
<td>0.33</td>
<td>7.89</td>
<td>0.89</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AIM</td>
<td>0.60</td>
<td>0.17</td>
<td>2.83</td>
<td>0.58</td>
</tr>
<tr>
<td>Y0624C/ [128950/40750]</td>
<td>BE {Pts}/33</td>
<td>SS</td>
<td>4.55</td>
<td>-1.96</td>
<td>14.31</td>
<td>-1.96</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AIM</td>
<td>2.45</td>
<td>-1.71</td>
<td>4.99</td>
<td>1.78</td>
</tr>
<tr>
<td></td>
<td>BE {Pts, Brklin}/6553</td>
<td>SS</td>
<td>4.10</td>
<td>-2.65</td>
<td>19.09</td>
<td>3.10</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AIM</td>
<td>3.28</td>
<td>-2.50</td>
<td>9.80</td>
<td>2.13</td>
</tr>
<tr>
<td></td>
<td>MMO {Brklin}/41738</td>
<td>SS</td>
<td>7.56</td>
<td>1.43</td>
<td>103.96</td>
<td>7.42</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AIM</td>
<td>1.30</td>
<td>0.30</td>
<td>7.87</td>
<td>1.26</td>
</tr>
</tbody>
</table>

Table 4. Error statistics of reference data minus automatically generated DSMs. Single and interpolated points from breaklines are indicated as Pts and Brklin respectively. Points on bare earth (BE) close to building outlines are indicated as hPts.

error of 1.30 m compared to 7.56 m of SS. In addition, SS exhibited gross errors of over 100 m. The good performance of AIM at MMO is also shown by the smaller mean values compared to SS. The mean values for both methods show that matching generally measures higher than the manual measurements, while for MMO matching results are lower. The fact that the accuracy was less for BE than MMO for Y0624C and W1052C is due to the density of the buildings and the relative short distances of the measured points to neighboring MMO. BE breaklines in these areas have been extracted along discontinuities on the ground, which were close to MMO. Therefore, it is possible that modeling errors were introduced.

The example in Fig. 6 shows the elevation differences computed for 900 reference points along a specific breakline defining a roof edge. The differences are smaller for points derived with AIM, compared to SS, because the use of contour points and edges in conjunction with the multi-image matching approach used in AIM, improves the modeling of discontinuities.

4. CONCLUSIONS

In this study, two systems have been evaluated regarding DSM generation using ADS40 imagery acquired over two different test areas. The AIM system showed better performance, compared to SS system, especially along building discontinuities. Accuracy increased, often by factor two or more, and blunders decreased, even in difficult areas, as Yokohama. Various components of the AIM method, especially those that take advantage of the characteristics of ADS40, lead to this improved performance. With the exception of breaklines in W1052C, the performance of AIM in the rural areas, even at MMO, was close to the theoretically expected and the accuracy of the manual measurements. However, dense urban areas still pose a problem and need algorithmic improvements but also denser matching and use of more images from neighbouring strips that should have a high overlap. A better configuration of the lines on the focal plane, having e.g. at least two lines with a sufficient base to height ratio in both forward and backward to minimise occlusions, would also be beneficial for urban mapping. Currently, further research in AIM is focused on improving surface modeling, fine-tuning adaptivity of parameters and extending the matching algorithm on Lev0 images with non-adjusted orientation data for tie point extraction in AT.
ACKNOWLEDGEMENTS

This work was performed within the project AIM, a cooperation of ETH Zurich and LGGM. The authors are grateful to Kikuo Tachibana, Tadashi Sasagawa and Hiroyuki Okada, PASCO Corp., for providing the ADS40 dataset of Yokohama. Natasha Vassilieva, IGP-ETH, is especially thanked for measuring the reference data in both Yokohama and Waldkirch datasets. The general support of the ADS40 development team of LGGM and the assistance of Muzzafar Adiguezel and Fernando Shapira, LGGM support team, in data processing are gratefully acknowledged.

REFERENCES


