ADAPTIVE MULTI-IMAGE MATCHING ALGORITHM FOR THE AIRBORNE DIGITAL SENSOR ADS40

Maria Pateraki, Emmanuel Baltsavias

Institute of Geodesy and Photogrammetry, ETH-Hoenggerberg, CH-8093, Zurich, Switzerland
Tel. +41-1-633 3966/3042, Fax +41-1-6331101, (maria, manos)@geod.baug.ethz.ch

ABSTRACT

In the recent years, the development of digital aerial sensors signifies a revolutionary change in aerial image acquisition, and the possibility of fully digital processing from image acquisition to the generation of value-added products for various applications. LH Systems developed the ADS40, a pushbroom sensor system, which offers on a single camera system the possibility to acquire 3 panchromatic and 4 multispectral images, broadening thus the possible application areas of the sensor. The radiometric characteristics of ADS40 imagery at raw and rectified level are investigated for various data sets. New algorithms of pre-processing of images that sharpen edge definition, enhance contrast and optimally reduce image data to 8 bit are applied prior to matching. The development of an adaptive multi-image matching algorithm for ADS40 data, is the main subject of the work presented in this paper. In the current investigations, only the 3 panchromatic channels have been used and matching is based on cross-correlation or similar measures to provide pixel-accurate positions, which are to be subsequently refined, using sub-pixel matching techniques. The geometry of the sensor is used to apply constraints for matching and the concept of epipolar lines is modified for the ADS40, since each line has its own position and attitude, and the geometry is perspective only on the CCD line direction. Edge pixels are selected for matching, since they can better model discontinuities and have sufficient texture. A modified image pyramid approach, using doublets, is used for derivation of approximations. Quality criteria that are computed during runtime are combined to eliminate errors. Test results are presented, using for the time being only visual control.

KEY WORDS: ADS40, airborne digital sensor, image preprocessing, DSM, matching, quality control

1. INTRODUCTION

Recently, digital photogrammetric cameras have been developed, including the commercial systems ADS40 (LH Systems), DMC (Z7I Imaging) and the Japanese TLS (Starlabo), as well as HSRC of DLR. A crucial factor for the success of the new digital sensors is the development of associated software with high-quality performance, that fits the characteristics of each sensor. For example for ADS40, matching can make simultaneous use of up to 7 images for each strip, and the known interior and exterior orientation of each line can be exploited to enforce geometric constraints. The datasets that were processed come from two regions, with different landscape and density of houses and trees. The dataset of Waldkirch (CH) covers a rural/suburban area (image scale is 1:30000 and pixel footprint 0.2 m) and the dataset of Chofu (JP) an urban one (image scale is 1:25000 and pixel footprint 0.17 m).

1.1 Description of ADS40

ADS40 incorporates latest GPS and INS technology (Applanix) for sensor orientation and new developments in sensor technology, optics, electronics and data transfer and storage. The ADS40 data used come from a camera with seven parallel sensor lines in the focal plane of a single lens system – three panchromatic (forward, nadir, backward), red, green, and blue lines placed next to each other and one infrared. The focal plane has space for two adidtional lines, but these are not currently used. Each panchromatic channel consists of two lines, each with 12000 pixels, staggered (shifted with respect to each other) by 0.5 pixels. In the test data used here, only one staggered line for each panchromatic channel wasused. The viewing angles and characteristics of the sensor are shown in Table 1. Additional information on the ADS40 is given in Reulke et al. (2000), Sandau et al. (2000) and Tempelmann et al. (2000).

Focal length	62.5 mm
Pixel size	6.5 μm
PAN	2*12000 pixels
RGB – NIR	12000 pixels
Forward to nadir stereo angle	28.3° deg
Backward to nadir stereo angle	-14.1° deg
Near infrared to nadir stereo angle	-2.05° deg
RGB to nadir stereo angle	15.9° deg
No. of bits from the ADC	14

Table 1. ADS40 characteristics.

1.2 Ground Processing

Apart from the imagery, GPS position and IMU data are recorded during the flight and written to the mass memory system (MMS). GPS/IMU data from the Applanix system are postprocessed to provide orientation for each image line. To compensate remaining errors (e.g. errors in the GPS/IMU postprocessing, IMU misalignment, calibration errors), a bundle adjustment (aerial triangulation) is applied with a number of GCPs and tie points. Tie points are matched in images rectified on a plane by using the raw GPS/IMU data, and their coordinates are transformed to the raw images before entering the bundle adjustment. The results of the bundle are used for a more accurate rectification. The rectification is needed for stereo viewing, DSM/DTM measurement, mapping etc. within Socet Set. The raw images are referred to as Level 0 images and the rectified ones as Level 1.

2. RADIOMETRIC ASPECTS AND IMAGE PREPROCESSING

2.1 Radiometric Quality Analysis

Although ADS data are 14-bit, the effective number of bits is less (ca. 12 according to LHS). Histograms of the Level 0 images that were processed show a strong peak towards the darker grey values (about 310-340 in the panchromatic channels). The effective grey level range was 13 bits for the panchromatic channels and 12 for the multispectral, excluding grey values at the histogram ends with frequency less than ca. 0.0012% (5 times less the frequency occuring if all 14-bit values were equally occupied).

A noise analysis has been performed for different datasets in Level 0 and Level 1 images, according to the method presented in (Baltsavias et al, 2001). The noise characteristics were analysed in non-homogeneous areas and the standard deviation, as an indication of noise, was calculated with a 3x3 mask size every 3 x 3 pixels, for various grey value ranges. Table 2 shows results for the Waldkirch data, whereby the Pan results represent the average of the 3 channels.

Grey value range	1 - 511	512 - 1023	1024 - 1535	1536 - 2047	2048 - 2559	2560 – 3071
Pan level 1	1.3	2.8	3.2	3.6	3.7	5.0
Pan level 0	2.3	4.4	5.0	5.6	5.5	7.2
RGB + NIR level 1	1.3	1.9				

Table 2. Noise estimation (mean standard deviation) of Waldkirch data using non-homogeneous areas.

The standard deviations indicate that noise is generally low and even more for the Level 1 images, due to the resampling of pixels during rectification. The multispectral channels exhibit slightly less noise than the panchromatic. Noise is increasing with increasing grey values. Some grey value ranges have no reliable standard estimation, due to the small number of samples. These results are consistent with the noise analysis that has been performed by DLR (Börner et al., 2000), stating that the standard deviation is about 2 grey values.

2.2 Image Preprocessing

In order to optimise the images for subsequent computer processing, various preprocessing methods were developed, implemented and tested. All preprocessing was applied to the 14-bit data. First, two adaptive local filters were developed (Baltsavias et al., 2001). They reduce noise, while sharpening edges and preserving even fine details such as one-pixel wide lines, corners and line end-points. The effect of the two local filters is generally quite similar, although they use a different number and size of masks, and one employs a fuzzy method. They require as input an estimate of the noise, which may be known or estimated by the method mentioned above. Table 3 shows the noise estimation before and after the adaptive filtering. After filtering, noise is reduced by about factors 2 to 3.

	1-511	512 - 1023	1024 - 1535	1536 - 2047
Backward / before NR	1.2	3.0	3.1	3.8
Nadir / before NR	1.3	3.0	3.6	4.0
Forward / before NR	1.3	2.3	2.9	3.2
Backward / after NR	0.4	1.3	1.4	1.7
Nadir / after NR	0.5	1.3	1.5	1.7
Forward / after NR	0.5	1.1	1.4	1.6

Table 3. Noise estimation (mean standard deviation) using non-homogeneous areas before and after noise reduction, using the Waldkirch dataset.

Next, a new version of the Wallis filter (Baltsavias, 1991), which estimates automatically some of the filter parameters, is applied, and finally a reduction to 8-bit imagery by histogram equalisation is performed. With Wallis, the contrast and thus also the noise is enhanced, but since the noise has been reduced in the first preprocessing stage, the noise level still

remains under the noise level of the original images. Histogram equalisation is used, since it preserves more grey values that are more frequently occurring. The histogram equalisation was iterative with the aim being to occupy all 8 bits with similar frequencies. The histogram equalisation may lead to too strong bright and dark regions for visual interpretation, and thus in such cases it can be replaced by a histogram normalisation (Gaussian form). Figure 1 shows the radiometric improvement of 8-bit Level 1 images after pre-processing.
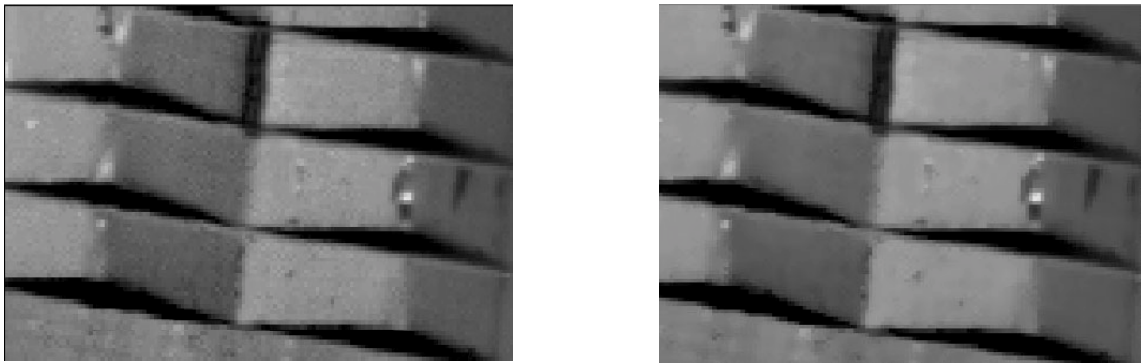

Figure 1. Detail of roof (Level 1 data after Wallis filtering and reduction to 8-bit by histogram normalisation) without (left) and with (right) noise reduction. On the right, noise is reduced, edges are sharpened and small details are preserved.

3. MATCHING

3.1 General principles

Methods for automatic DTM and DSM generation exist already in various commercial digital photogrammetric systems, including SOCET SET of LH Systems. All existing matching algorithms are geared towards frame aerial imagery, with a usual overlap of 60% per stereopair, while in matching only two images are used. The ADS40 however, offers new capabilities, which make necessary the development of new matching algorithms based on another philosophy. ADS40 provides full 100% overlap of 3 to 7 images for each strip. Thus, a multi-image matching approach is followed, leading to substantial reduction of problems caused by occlusions, multiple solutions, image noise, and surface discontinuities and higher measurement accuracy through the intersection of more than two image rays. The known interior and exterior orientation are used to enforce geometric constraints, restricting the search space along quasi-epipolar lines. The sensor model developed by LH Systems is used in all transformations between image and ground coordinates. The main aim of the developed methods is to derive a DSM/DTM from Level 1 images, and a by far secondary one is to measure tie points or derive a coarse DSM (using Level 0 or Level 1 images), results which can be used in the bundle adjustment. In the first matching stages, cross-correlation or other similarity or difference measures are used delivering pixel accurate results, which can be improved by subsequent sub-pixel matching methods, e.g. least squares matching. Level 1 images generally exhibit small scale and rotation differences, especially in the upper pyramid levels, and thus the use of only shifts in matching, without additional shaping parameters, is justified, leading to faster processing than other time consuming matching methods like least squares matching. Apart from normalised zero-mean cross correlation, the sum of squared differences or of the absolute differences can be used and thus further decrease processing time. Although investigations have reported that these difference measures are inferior to normalized zero-mean cross correlation, e.g. sensitive to radiometric differences, after the preprocessing and radiometric equalization that we apply, first tests show that these measures perform equally well. Matching is performed at pixels of extracted features (Sec. 3.2) and approximations are derived by a modified image pyramid concept (Sec. 3.3).

In matching, different type of image information can be used: original grey value images, binary images by thresholding the edge magnitude, images where edge magnitude is kept and non-edges are set equal to 0, variations of the two previous image types, by encoding through appropriate values information on the edge pixel orientation or sign. The rationale behind using image types other than the original grey values, and especially binary ones, was a possible matching speed-up, especially in the upper pyramid levels, where not the outmost accuracy is sought for. Initial tests showed that coarse matching with binary images was slightly faster and the accuracy was similar compared to the points matched using gray level images. Therefore, binary images may be used in the upper pyramid levels without significant loss in accuracy but not in the finer pyramid levels.

3.2 Extraction of features for matching

Extraction of features was applied after preprocessing of the images to detect match points. Extracted match features should be as dense as possible, thus making the blunder detection and correction easier, reducing propagation of mismatches to the lower levels. Our consideration was to develop or to improve existing algorithms in such a way that we would gain in speed and robustness. Since each image could be kilometers long, which corresponds to more than 1 Gbyte, processing

time should be short and memory should be efficiently managed. We performed several tests with existing operators, like Foerstner and Harris (used for extracting points), and Canny and Susan (used for extracting edges), and developed an edge extractor, which performs a radiometric analysis of the image content, calculates statistical measures and thresholds depending on a user defined value. A comparison of the two point detectors, showed similar results, with Foerstner permitting a feature localisation, while Harris was by 50% faster. Results showed that Canny and the in-house developed algorithm performed better than the other algorithms in terms of processing time and amount of match points. Compared to Canny, the in-house method needs one third of the processing time and extracted edges have width more than one pixel, allowing a better modelling of discontinuities. During edge extraction, additional information on edges, like orientation and sign, are derived and can be used in further matching stages. The edge extraction can be combined with the derivation of the image types, mentioned at the end of Sec. 3.1. All algorithms were modified to work also on 14-bit data, so that the full image information is used during processing. Figure 2 shows the extracted points with the Harris operator and the extracted edges with the Canny operator. The feature extraction is either performed in all channels used in matching or only in the template image (if the grey value images are used and not the other image types).
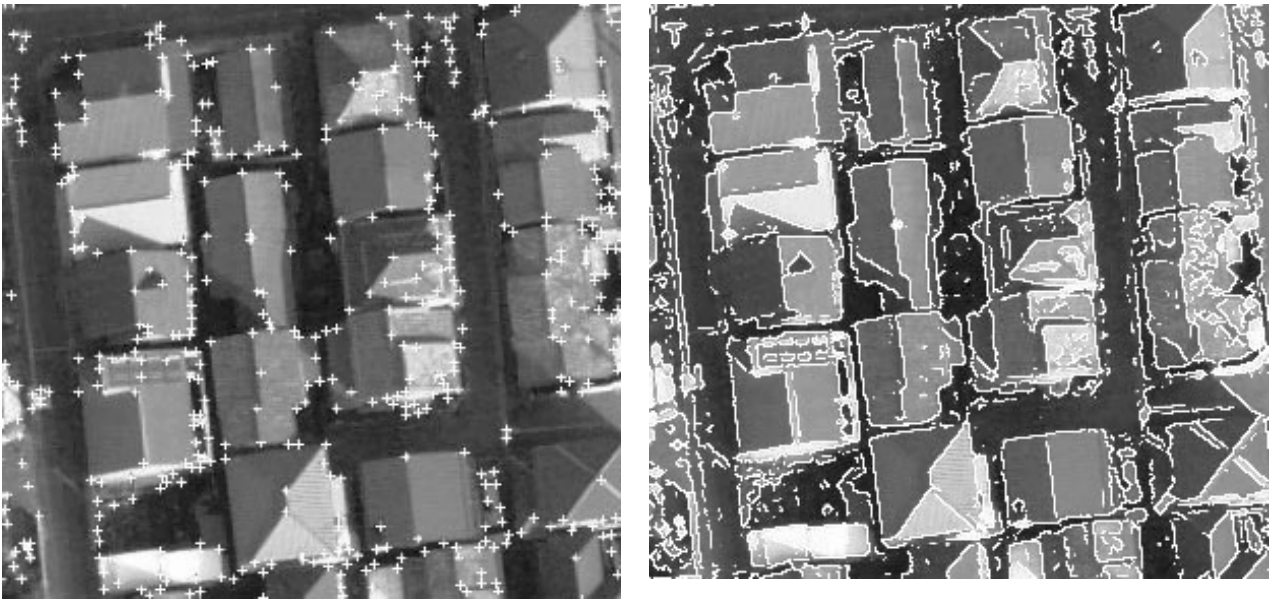

Figure 2. Left: extracted points with Harris operator. Right: Extracted edges with Canny operator.

3.3 Approximate values and use of doublets in image pyramids

The concept of image pyramids is utilised in the algorithm, so that parallaxes become small and matching is applied in small search ranges. Level 1 images have small y-parallaxes, e.g. less than 100 pixels maximum with bias. For Level 0 images, parallaxes may be larger although the known sensor orientation is taken into account and offsets are calculated depending on the viewing angles of the channels that are used. Image pyramids have the main disadvantage that texture may disappear in the upper levels. Therefore, we avoid using the commonly used of average filter for pyramid generation and we apply optimal filters that permaintain features as much as possible in the upper levels (Baltsavias, 1991). In addition, images are enhanced in each level and are radiometrically balanced using the Wallis filter.

Most matching methods employ a time consuming interpolation to pass coarse matching results to the lower levels. Therefore, the doublet strategy is being used. This strategy aims at reducing processing time and interpolation of match results from one pyramid level to the other, thus also restricting propagation of match errors . Doublets are consisting of 2 consecutive pyramid levels. Extraction of features is performed in the lower level of doublets and these features are transferred one level up, and kept only if on this level an extracted feature (e.g. edge) also exists. Then, matching in all 2 levels from top to bottom is performed. Thus, interpolation and propagation of errors to neighboring points is avoided. For example, when 6 pyramid levels are used, instead of interpolating 5 times between the different levels, interpolation is applied between the three defined doublets, performing in total 2 interpolations.

3.4 Multi-patch matching, use of geometric constraints and combination of multiple channels

In matching, a multi-patch approach is adopted where at each match point 3 passes of matching are performed with different parameters (such as search range and patch dimensions). The use of multi-patch approach can be justified since the larger patches aim at reliability of a coarse solution and the smaller ones at accuracy. The large patch is less sensitive to noise, occlusions, multiple solutions etc. while the small one is more accurate and better preserves height

discontinuities. Furthermore, the matching results for each of the 3 passes can be compared and used in the quality control for error detection.

Constraints may also be used in matching. First, the behavior of epipolar lines in rectified images has been investigated. It is assumed that a feature is defined in the so-called template image and the corresponding features are searched for in the remaing images, called patch images. Epipolar lines do not really exist since each line composes an image with its own position and attitude. But by projecting a point of the template image to different heights and back projecting onto the patch image the trajectory of the corresponding point and the epipolar curve may be defined. For rectified images the epipolar curve, if it is not too long, can be approximated by a line. Two different approaches have been adopted for geometrically constrained matching. The first method intersects the template image ray with height planes with a step of 1 m, backprojects these object point to each patch image and fits a straight line to these points.. The second method finds an approximate height by an initial forward intersection (pairwise, using the template and each patch image), the search range in image space is transformed to height search range on the ray of the template and the points that are selected for matching are back projections of 3D points along this ray, within the height search range, and with a height step that corresponds to one pixel step in image space. Both methods give similar results for the matched points, but the second method is faster than the first one.

In the tests up to now, 3 PAN lines have been used. As template image the one with the least possible occlusions, compared to the remaining ones i.e. in this case the nadir, has been used. The forward and backward channels can be matched either separately or simultaneously with the nadir. In the first case the matching is performed for backward and nadir and the extracted 3D points are back projected onto the forward channel and used as initials approximations for the matching. In the second case, backward and forward are matched simultaneously. In both cases a quality control procedure checks each ray, and the final 3D point coordinates are calculated by forward intersection using only the good rays. We are currently investigating the integration of red or green and near-infrared channels in the matching process. The RGB channels and the NIR are positioned on the focal plane on either side of the nadir, and thus can be combined pairwise with the backward and forward channels to model surfaces that are occluded in the other channels. Among the RGB channels, red or green are preferable due to the better sensitivity of the CCD in these spectral ranges.

3.5 Quality control

During run time, quality measures are calculated and used for elimination of blunders and false matches. The similarity measure, the 2nd best similarity score and its distance to 1st one, the size of search window, the change of similarity measure between the 3 patch sizes, the change of position between patch sizes, the angle of dominant edge direction with the epipolar line, the residuals from 3D forward intersection, the change of final point position from the starting position of the largest patch, the approximate standard deviations for x and y pixel coordinates are the most important quality criteria calculated. The quality criteria are combined according to possible occurring errors. E.g. in case of an occlusion, the cross-correlation coefficient would be small and the change of the similarity measure would be generally decreasing from the largest to the smaller patch. Thresholds for the quality criteria are derived by a statistical analysis of their values in the given pyramid level. We will investigate a fuzzy approach for threshold calculation, and relative weights when combining together multiple quality criteria to derive a single quality measure. All quality criteria are used for each image ray individually, to detect problems occurring in individual images, e.g. occlusions. Weak rays are excluded and final 3D intersection is computed using only the good rays.

Regarding multiple solutions, along the epipolar lines, the multi-patch approach may not detect such problems. Thus, additionally the consistency of height in the local neighbourhood is checked to detect and eliminate spike errors by a median-type filter.

Least squares matching is used after blunder elimination with the above quality crieria. Successful matches are rechecked with harder quality criteria resulting from the least squares approach, such as number of iterations and change of shape parameters of the patch.

4. RESULTS

The matching algorithm has been used for DSM generation in the previously mentioned datasets. Out of these datasets, the Level 1 PAN images were processed and DSM's have been extracted. Figure 3 shows a portion of the results for the Waldkirch data for the 0th pyramid level and matching without constraints using doublets. The part of the area that has been used for DSM extraction was 2000 x 1000 pixels. The number of points that have been matched in the lower level of the last doublet is about 165,000 and after the quality control 130,000. The percentage of successful matches was around 80%. The irregularly distributed match points have been used to interpolate a regular grid with 0.5 m spacing which was hill-shaded. The match points were also used to interpolate contours with an interval of 2 m. The results were controlled visually by overlaying contours on stereo images using SOCET SET and by inspecting individual points. The major problems encountered were with repetitive fine texture (multiple solutions), e.g. in agricultural fields.

We have applied the algorithm also on Level 0 images where matching can be more problematic, due to the perturbations of the aircraft and the non perfect camera stabilization that influence the produced raw images. As the results were good, the developed algorithm can be also used in aerial triangulation for extraction of tie points or determina-

tion of a coarse DSM (latter can provide better approximations for tie point area selection and tie point matching). In this case constraints are less strict than in Level 1 images since the orientation data have not been adjusted with aerial triangulation yet, and the GPS/INS data are not perfect. In addition, the quality control procedure can be stricter than the one used when matching Level 1 images, resulting in less, but however sufficient in number and distribution, matched points.
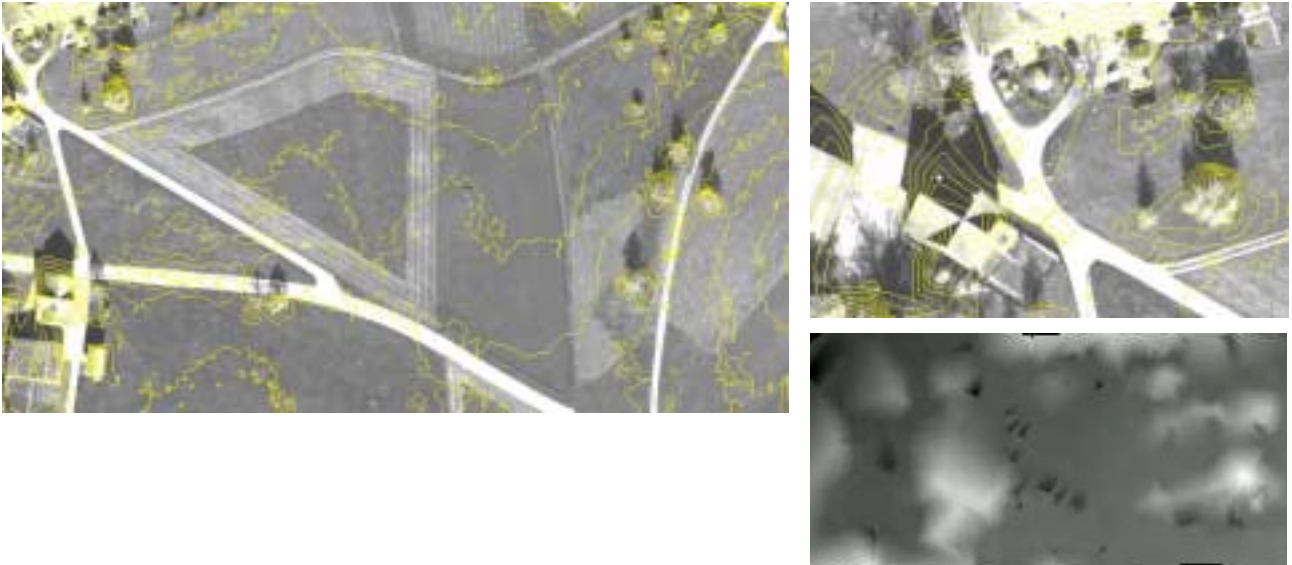

Figure 3. Portion of test area in Waldkirch (CH). Left: Level 1 nadir channel overlaid with contours with 2 m interval. Right top: zoomed region from the same area. Right bottom: DSM image

5. CONCLUSIONS AND FUTURE WORK

ADS40 is a new camera system, which is still being fine-tuned. Thus, there are not so many data available, and even less data in regions with different land cover and relief and accurate reference DSMs for testing our matching methods. This was one of the main obstacles in our previous work and we hope that the next period significant improvements will be done in this direction. Thus, it is planned to quantitatively compare various matching methods and side aspect variants. Furthermore, the quality control procedure, probably the most critical aspect of matching, will be further developed and tested. Further work will also include reduction of DSM to DTM using also NIR information for the elimination of trees.

ACKNOWLEDGEMENTS

REFERENCES

Baltsavias, E.P., 1991. Multiphoto geometrically constrained matching. Ph.D. dissertation, Institute of Geodesy and Photogrammetry, ETH Zurich, Report No. 49.

Baltsavias, E., Pateraki, M., Zhang, L., 2001. Radiometric and Geometric Evaluation of IKONOS Geo Images and their use for 3D building modelling. Proc. Joint ISPRS Workshop "High Resolution Mapping from Space 2001", Hannover, Germany, 19-21 September (on CD-ROM). Available also at http://www.ipi.uni-hannover.de/html/publikationen/2001/isprs_cd.pdf (accessed 28 June 2002).

Reulke, R., Franke, K., Pomierski., T., Schönermark, M., Tornow, C., Wiest, L., 2000. Target related multispectral and true colour optimization of the colour channels of the LH Systems ADS40. Int'l Archives of Photogrammetry and Remote Sensing, Vol. 33, Part B1, pp. 244-250.

Sandau, R., Braunecker B., Driescher, H., Eckardt, A., Hilbbert, S., Hutton, J., Kirchhofer, W., Lithopoulos, E., Reulke, R., Wicki, S., 2000. Design principles of the LH Systems ADS40 airborne digital sensor. Int'l Archives of Photogrammetry and Remote Sensing, Vol. 33, Part B1, pp. 258-265.

Tempelmann, U., Börner, A., Chaplin, B., Hinsken, L., Mykhalevych, B., Miller, S., Recke, U., Reulke, R., Uebbing, R., 2000. Photogrammetric Software for the LH Systems ADS40 Airborne Digital Sensor. Int'l Archives of Photogrammetry and Remote Sensing, Vol. 33, Part B2, 552-559.