DEM BASED REGISTRATION OF MULTI-SENSOR AIRBORNE POINT CLOUDS EXEMPLARY SHOWN ON A RIVER SIDE IN NON URBAN AREA

R. Boerner[1,*], Y. Xu [1] L. Hoegner[1], R. Baran[2], F. Steinbacher[2], U. Stilla[1]

[1]Photogrammetry and Remote Sensing, Technical University of Munich
80333 Munich, Germany - (richard.boerner, yusheng.xu, ludwig.hoegner, stilla)@tum.de
[2] SteinbacherConsult Ingenieurgesellschaft mbH & Co. KG
86356 Neusäß/ Augsburg, Germany - (r.baran, f.steinbacher)@ahm.co.at

Commission II, WG II/2

KEY WORDS: airborne point clouds, registration, 2.5D DEM, multi sensor

ABSTRACT:

This paper presents a method to register photogrammetric point clouds generated from optical images acquired by UAV and aerial LIDAR point clouds. Normally, the registration of two airborne scans of the same scene is solved by the use of control points and the direct registration using GNSS and INS information. However, the registration of multi-sensor point clouds without control points is more complicated and challenging. For the scene of non urban areas, the registration task gets even more complicated, because it is hard to extract sufficient geometric primitives from the building structures. For our proposed method, an outdoor scene is tested providing almost no man-made objects. Therefore, it is nearly impossible to search for planar objects and use them for registration. With no geometric primitives extracted, the proposed method utilizes the structure of the 2.5D DEM created from the ground points of the point cloud. Besides, instead of using control points or key points, the method automatic detect key planes from the 2.5D DEM as correspondences. These key planes are detected on a regular grid by the use of a predefined mask. To mark a DEM grid cell as key plane the histogram of sums of the angles between the center cell is used. Afterwards, similarity values between two key planes are calculated based on the histogram differences and a RANSAC based strategy is adopted to find corresponding key planes and estimate the transformation parameters. Experiments conducted in this paper indicate that it is feasible to register multi sensor point clouds with a big difference in their ground sampling distances with respect to the used cell size of the 2.5D DEM.

1. INTRODUCTION

Airborne Laser Scanning (ALS) is currently used to cover the 3D information of a large region (e.g., several km^2) in a reasonable time. Since there is a GNSS and INS used during the flight, the registration of ALS data can be done in a direct manner by using the GNSS and INS information to georeference the point cloud. Misalignments resulting from measure errors of the GNSS antenna and IMUs can be solved with a fine registration (e.g., ICP). In the case of using a UAV, multiple images are relative registered to each other with an aereotriangulation approach. A georeferencing of the adjusted image block is limited by the accuracy of the UAV mounted GNSS sensor. The resulting point cloud therefore needs to be georeferenced by using control points. Therefore, the registration is simple inside the same sensor system and not considered as an subject of research. Instead, the research question could be: how to register datasets of two different sensor systems without using any markers.

By comparing a photogrammetric point cloud captured with images taken from an UAV and a LIDAR point cloud taken from an airplane, two main differences are obvious: One is the big difference in the ground sampling distance of the two point clouds. The other difference is that photogrammetric point clouds have more noise and outliers than LIDAR point clouds. Both the difference in the ground sampling distance and the noise makes the point based registration of the different data difficult. There are some approaches in the field of point cloud registration which use geometric primitives like planes or congruent points to solve

the registration. This can be independent on ground sampling distances and noise but assumes that there are primitives inside the scene. However, this is not always the case, especially by considering non urban outdoor scenes which show no man made objects. For these scenes the landscape itself has to be used for registration which is done by the proposed approach by using the 2.5D DEM of the ground points.

The proposed approach aims for a combination of the advantages of UAV taken data and ALS data. The method is clustered in the following steps (see figure 1): the surface fitting, the detection of key planes, the matching of key planes, and the estimation of the transformation parameters. Possible applications would be the enrichment of 3D data with additional data with higher temporal and spatial resolution. Therefore, a registration approach is needed which is independent on the point cloud ground sampling distance or the sensor type used to capture the point cloud. Furthermore, in the case of capturing landscapes it is likely to get scenes without any man made objects and an approach is needed which uses the landscapes itself. The proposed method creates results for a coarse registration and the accuracy depends on the chosen cell size for surface reconstruction but the mentioned conditions above are fulfilled.

Section 2 gives an overview of the state of the art approaches in the context of point clouds registration. Section 3 explains the proposed method and section 4 shows the used data to challenge the proposed method. The results are shown in section 5 and the stated ideas for future work and applications are explained in more detail in section 6.

---

*Corresponding author
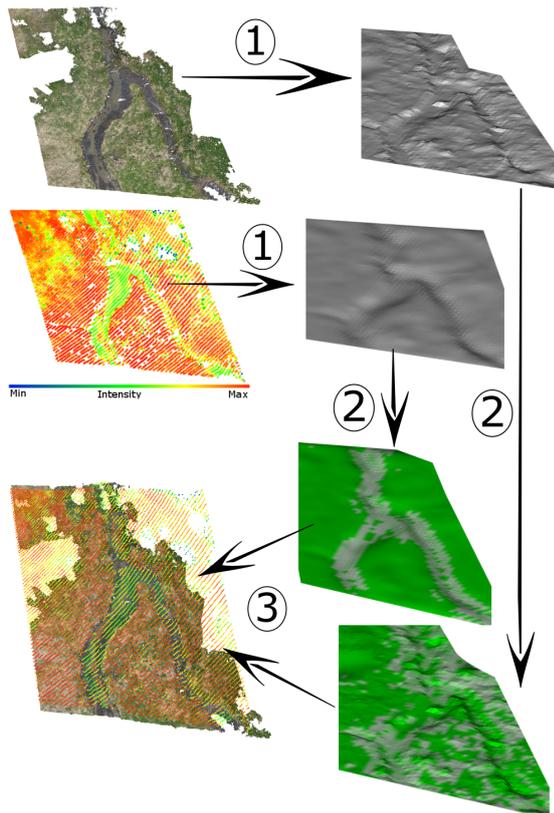
Figure 1. Scheme of the proposed approach. The arrows show the different steps for the registration, with start at the input point cloud: 1: surface reconstruction; 2: search for key planes; 3: the calculation of transformation parameters based on matched key planes and registration of input point clouds.

2. RELATED WORK

Currently, plenty of marker less registration approaches have been developed for aligning different scans by using geometric characteristics of given point clouds. According to the geometric features used, these registration methods can be grouped into two major categories: the point-based and the primitive-based approaches.

For the point-based approaches, their basic idea is to find corresponding point pairs in different scans. The Iterative Closest Point (ICP) and its variants minimize point-to-point distances in overlapping areas of two different point clouds (Besl and McKay, 1992; Habib et al., 2010). For decades, ICP-like methods have been proven to be effective in terms of accuracy in plenty of applications. However, the matching of corresponding points requires a time-consuming iterative process if no good initial alignment obtained. Besides, instead of using all the points as candidates selecting key points from the raw point cloud is also an alternative that can largely reduces the computation cost, for instance the SIFT key points (Weinmann et al., 2011), DoG key points (Theiler et al., 2014), FPFH key points (Weber et al., 2015), and semantic feature points (e.g., intersecting points) (Yang et al., 2016; Ge, 2017). Although all these methods are generally able to align point clouds, the point-based methods are sensitive to point density and noise and have problems in terms of efficiency when dealing with large-scale datasets.

For the primitive-based approaches, instead of using points, the geometric primitives formed by points (e.g., lines (Habib et al.,

2005), planes (Xiao et al., 2013), or surfaces (Ge and Wunderlich, 2016)) are adopted as geometric features for the registration. Theoretically, the use of higher level geometric features can increase the robustness of identifying corresponding feature pairs. Lines, planes, and curved surfaces are the representatives of geometric features. Large numbers of investigations using line features to register point clouds have been reported. For examples, the intersecting lines of neighbouring planes (Stamos and Leordeanu, 2003), 3D straight-lines (Habib et al., 2005) , and spatial curves (Yang and Zang, 2014) are used as matching primitives. Plane correspondences (Dold and Brenner, 2006; Von Hansen, 2006; Xiao et al., 2012) and surface correspondences (Ge and Wunderlich, 2016) are frequently used as geometric primitives for alignment as well. Compared with the point-based methods both line- or plane-based methods require abundant linear objects or smooth surfaces as candidates which largely depends on the content of scanned scenes. For scenes with only few buildings they may meet problems when finding appropriate candidate features of lines and planes and the accuracy of extracting lines and planes will affect the registration result at same time. Besides, for the plane-based methods the extraction of planar surfaces with region growing or model-fitting algorithm can be rather time consuming and unreliable, which largely limits the performance of the registration as well (Wang et al., 2016), (Xu et al., 2017). For efficiency matters the voxel structure is frequently used for processing point clouds. For segmentation plenty of studies have illustrated the advantages of using voxel structures: constructing a rasterized representation of points which simplifies the dataset and can overcome the uneven distributed point density. Nevertheless, selecting an appropriate resolution of voxels for the raw point cloud is a trade-off between the efficiency of processing and the preservation of details. Normally, the smaller the voxel, more details will be kept. With respect to the registration using voxel structures, in the work of (Wang et al., 2016) the EGI features formed by the clusters of voxels are utilized as correspondences for coarse registration, demonstrating effective results for aligning indoor scenes. Moreover, in related work (Xu et al., 2017) , planar surfaces are extracted from the voxel representation of the raw point cloud, largely increasing the process of extracting planar surfaces which also reveal a promising potential for coarse orientation of scans of urban scenes.

However, for the majority of the aforementioned primitive-based cases, they focus on residential or other urban areas which can provide sufficient geometric features like lines, planes, or intersecting points but when it comes to more general outdoor scenes like agricultural fields, forest or river valleys, it will be hard to find enough salient features. Therefore, how to generate reliable primitives for finding enough correspondences becomes a key point when registering point clouds of the large-scale general outdoor scene.

3. METHOD

The proposed method consists of three major steps, namely the surface fitting to generate a 2.5D surface on a regular grid. This surface consists of smaller planes which approximate the scanned area. The next steps of the proposed approach are the key plane detection (i.e. detection of surface planes which are considered as key features for the next steps), key plane matching, and the estimation of transformation parameters (see figure 1). Key planes are searched where the surface shows distinctive curvatures. For the detection of key planes, a defined neighbourhood on the regular 2.5D DEM is used. A histogram of the angles between nor-

mals is defined by using the normal vectors of the center cell and all neighboured cells. Afterwards, the histogram is used to mark a center cell as key plane if the histogram shows a step function with minimal step width. The matching of key planes uses a Ransac based approach and the calculation of a similarity value based on the histogram. Related work of the authors reached good results in estimating the rotation between two scenes (Xu et al., 2017). However, the estimation of translation shows some problems because no corresponding planes are used so far. Therefore, this paper focuses on the estimation of translation parameters but the proposed approach can be extended to estimate rotation parameters, too.

3.1 Surface fitting

The proposed approach works on a regular grid surface which is fitted on the input point cloud with a least squares approach. One key feature of this surface is the generalisation of different resolutions of input point clouds. Even if the calculation of the grid parameters needs more processing time, the regular grid has several advantages in contrast to the Delaunay triangulation: It can be estimated with smoothing regularization and gets therefore less noisy. The output of this least squares estimation is shown in comparison of a Delaunay triangulation in figure 2. The given coordinates of a 3D point can be translated very fast to a grid cell by calculating the index of the raster cell in the regular grid. Therefore, the estimation of the point to grid distance is faster as with a Delaunay mesh. Furthermore, the regular grid is more appropriate to work with defined masks like in an image.

The surface is reconstructed with a least squares estimation where the height components of each edge of the regular grid cells are the estimated parameters. The input point cloud is the point cloud of the ground points which are filtered on the ALS data with the author's voxel based approach (Boerner et al., 2017). In the case of the UAV data the ground points are segmented manually to avoid errors because of misclassification. The least squares estimation of the surface is done like described in Förstner and Wrobel (2016)[758-760]. For each point of the input point cloud the observation to the least squares estimation is the bilinear interpolation of the parameters:

$$
\begin{aligned}
z = &(1-s)(1-t) \cdot a_{i,j} \\
&+ (1-s)t \cdot a_{i,j+1} \\
&+ (1-t)s \cdot a_{i+1,j} \\
&+ s \cdot t \cdot a_{i+1,j+1}
\end{aligned} \tag{1}
$$

with:
z = z-coordinate of the current point on the grid
s = $x - x_i$: difference of the x coordinate of the current point to the x coordinate of the anchor edge(x_i) on the grid inside the range [0,1]
t = $y - y_j$: analog to s with the y coordinates of the current point
a = parameters of the grid (i.e. the z coordinate of each edge of the current grid cell)
i,j = coordinates of the anchor edge of the current grid cell

The smoothing regularization is defined with the fictitious observations:

$$0 = \delta_{i,i} = a_{i-1,j} - 2a_{i,j} + a_{i+1,j} \tag{2}$$

$$0 = \delta_{j,j} = a_{i,j-1} - 2a_{i,j} + a_{i,j+1} \tag{3}$$

$$0 = \delta_{i,j} = a_{i-1,j-1} - a_{i-1,j} - a_{j,i-1} + a_{i,j} \tag{4}$$

The surface is reconstructed using the optimization of:

$$\Omega = \frac{1}{\sigma_n^2} \sum_m (l_m - z_m)^2 + \frac{1}{\sigma_\delta^2} \sum_{i,j} \delta_{i,i}^2 + 2\delta_{i,j}^2 + \delta_{j,j}^2 \tag{5}$$

with:
l_m = z-coordinate of observation point m
z_m = z-coordinate of point m on the grid
σ_n^2 = variance of observation point
σ_δ^2 = variance of the regularisation term
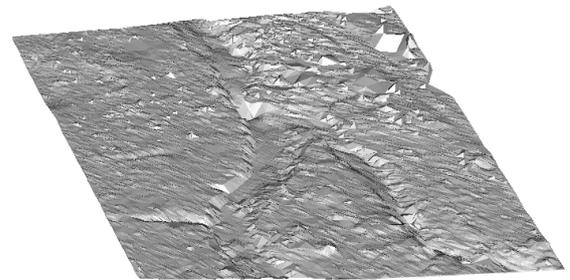
(a)
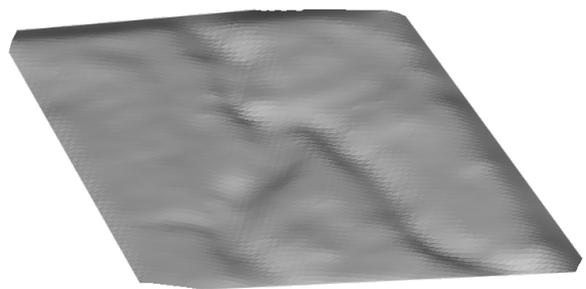
(b)

Figure 2. Comparison of a Delaunay triangulation (a, done with the software Cloud Compare[1]) and the reconstruction of a regular 2.5D DEM (b)

3.2 Key plane detection

The decision whether a center cell should be marked as a key plane is based on the following consideration: If the local neighbourhood represents a planar region the angles between the normals of grid cells (equation 3.2) will be near zero. If the local

---

[1]http://www.danielgm.net/cc/ (visited on 11.01.2018)

neighbourhood is non planar the normals point in different directions. In case of the local neighbourhood representing a cone there will be a continuous change of angle directions up to a local maximum. The key planes are marked in areas where the local neighbourhood shows the characteristics of a non planar region and in the best case of a cone-like region. The local neighbourhood has to be chosen large enough to include enough cells for the statistical evaluation of the normal directions. The descriptor becomes rotation invariant by choosing the differences of the normal directions in the local neighbourhood instead of the normal vectors themselves.

$$\cos(\alpha) = \vec{nc} \cdot \vec{nn} \qquad (6)$$

with:
$\vec{nc}$ = normal vector in the center
$\vec{nn}$ = normal vector for a neighbour
α = angle between center normal and the neighbour normal

Each angle inside the mask (eq.) is used to calculate a histogram of sums which is used as the mask descriptor. The histogram is defined with:

$$h(i) = h(i-1) + \frac{\texttt{count}(\alpha_i)}{n} \qquad (7)$$

$$\alpha_i = i \cdot \frac{180 \text{ deg}}{bins} \qquad (8)$$

with:
i = index of the histogram bin starting at zero
$h()$ = the value of the histogram bin
$\texttt{count}()$ = count of angles which are round off to the given value
n = count of halve cells inside the mask
$bins$ = count of bins in the histogram

A schematic visualization of the descriptor mask and the histogram is shown in figure 3 where figure 3a shows a non planar neighbourhood with its corresponding angles (3c) and the resulting histogram (3e) and figure 3b a planar neighbourhood with normals (3d) and histogram (3f).

If the local neighbourhood represents a very flat region the histogram shows a flat curve. If the local neighbourhood is non flat the histogram shows a step function. If the local neighbourhood represents a cone the histogram will sum up from zero to one in a small step. In other words, the latter case is represented by a histogram with a step width of a low bin count. Key planes should be found where the local neighbourhood shows strong curvatures like for example in a cone-like structure. Therefore, searching for key planes means to search for histograms which have a small step width, excluding a step width of zero (in case of a planar neighbourhood). The detection uses two thresholds which filter the cone like histograms from others: A two side threshold to define minimal and maximal values in the histogram and a one side threshold to define the bit count between the minimal and maximal value. If the bit count is fulfilled with respect to the chosen minimal and maximal values, the center cell is marked as key cell.

(a)   (b)

(c)   (d)

(e)   (f)
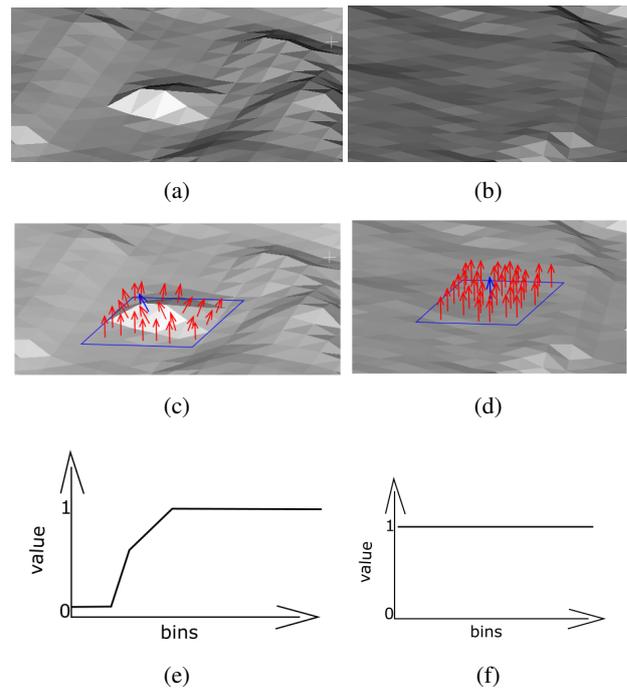
Figure 3. Principle of the key plane descriptor in comparison of a cone like area (a,c,e) and a plane like area (b,d,f), top (a,b): the area itself, middle (c,d): shematic drawing descriptor mask (blue) with the center normal (blue) and the neighbour normals (red), bottom (e,f): the histogram of angles with normed relative sums

3.3 Key Plane matching

The matching of key planes uses the descriptor defined in section 3.2 to find corresponding key planes for the coregistration of two DEMs.

$$s = 1 - \frac{1}{n \cdot \texttt{max}(|\vec{h1} - \vec{h2}|)} \sum_i |h1(i) - h2(i)| \qquad (9)$$

with:
s = similarity inside the range [0,1]
n = count of bins,
      have to be the same for both histograms
$\vec{h1}, \vec{h2}$ = the whole histograms

To avoid a division by zero, the maximum value is set to one if all histogram values are equal. This similarity is normed to be between zero and one, with a similarity of one showing that the two histograms show the same graph. Several neighbourhoods could look the same, therefore the output of a similarity match has to be filtered to exclude false matchings. Such a filtering is done by firstly conducting all matches with a similarity above a threshold as a match and filter these fully connected matches with a RANSAC approach. To speed up the RANSAC process, an initial position is given which is used to conduct only matches within this initial model as potential fully connected matches. This matching process is summarized with:

1. Conduct key planes as match where $s(\vec{h1}, \vec{h2}) > t_s$ and $\vec{x}_1 - [\mathbf{R}|\mathbf{t}]\,\vec{x}_2 < acc$, where acc is the accuracy threshold for the initial model and t_s is the similarity threshold and x_1, x_2 are the center points of the halve cell.

2. Take a match at random

3. Calculate transformation (for testing in the context of this paper its the difference of the center points of halve cells)

4. Count matches which fulfil $\vec{x}_1 - [\mathbf{R}|\mathbf{t}]\,\vec{x}_2 < matchD$, where $matchD$ is the threshold which decides if a match should be considered or not

5. If current count is higher than current best: adjust best model

6. After iteration n reached: return best model

4. DATA

As testing site, a part of approximately 3500 m^2 of the river Mangfall in Bavaria of Germany was chosen (see figure 4). The UAV flight was done on May 2016 and the ALS data was captured on April 2017. For the ALS data a Riegl VQ 880G scanning system was used and the point cloud have a ground sampling distance of approximately 40 cm between two scan lines and approximately 10 cm between two points in the same scan line. The UAV taken images were captured using a Sony NEX-7 and a Falcon UAV. The photogrammetric point cloud generated from the UAV taken images was calculated using the software Pix4D[2] and the resulting point cloud has a ground sampling distance of about 6 cm.

The test site shows no man made objects like buildings, only the landscape and some vegetation. The landscape itself is flat which makes the marker less registration challenging. The 2.5D DEM was calculated with a cell size of 80 cm. Figure 5 shows the used two point clouds. The resulting 2.5D DEM are shown in figure 6.

In comparison of the two point clouds (i.e., photogrammetric and ALS), it is obvious that the photogrammetric point cloud has a higher resolution but contaminated with more noise and outliers.

Figure 4. The Mangfall area, in the background: the map of Germany from Google maps, in the marked area: green marked area: the testing side located in the Mangfall area.

As described above, the test side is a multi temporal data of a river area. Therefore, there are two aspects which make a registration complicated. Firstly, the riverside can have changed over the time which creates more outliers by searching for correspondences. Then, the influence of different light reflection characteristics has to be corrected for the river ground points. Since the two point clouds are showing the same riverside which shows no visible changes, the aspect of possible changes is ignored. The ALS data, which is acquired via a bathymetric scanning system,

[2]`https://pix4d.com/` (visited on 11.01.2018)
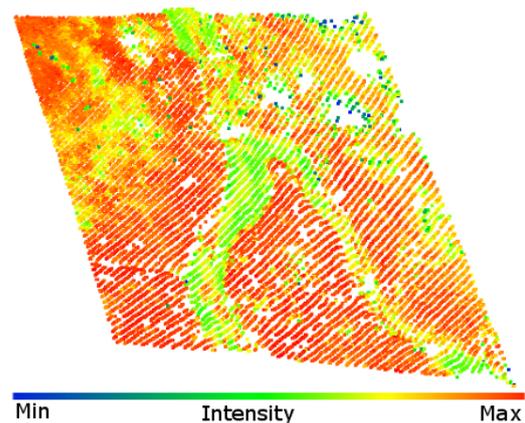
(a)

(b)

Figure 5. The used point cloud data, a) ALS data b) photogrammetric point cloud generated from the UAV taken images.

shows only single returns in the river area, because of low water depth and suspendent matter in the water. These single returns are considered as water points and therefore not corrected considering different light reflection characteristics. This consideration is also closer to the photogrammetric point cloud which shows only water points in the river area. These considerations may create errors in the final output but the correction of these errors is not a subject of this paper.

5. RESULTS

The first question to answer is what cell size should be used to reconstruct the surface. If the cell size is chosen to be smaller, the calculation of the parameters takes more computation time and memory. However, if the size of cells is chosen to be bigger, the level of details gets lower and the naturally flat landscape will result in an even more flat surface representation. Furthermore, since the matching of key planes is limited to a halve cell the resulting registration is expected to be limited to a halve cell accuracy, too. As shown in section 4 the photogrammetric point cloud consist of more noise than the ALS one. Therefore, for choosing the cell size the focus will be on the photogrammetric point cloud.

(a)

(b)

(a)

(b)

(c)

(d)

Figure 7. The photogrammetric point cloud with lower
resolution(a: 40cm, c: 80 cm) and the corresponding DEMs (b,d)
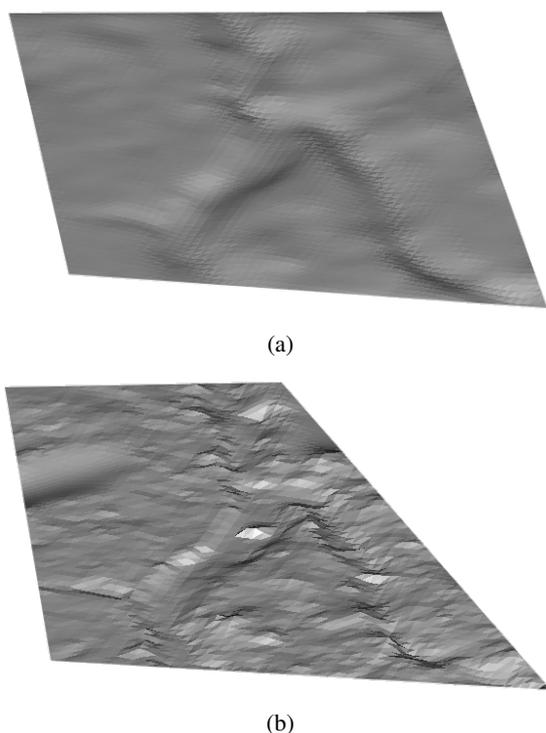
Figure 6. Corresponding 2.5D DEM to the shown point cloud
data (see figure 5) a) DEM for the ALS data, b) DEM for the
photogrammetric point cloud.

5.1 Choosing the cell size

There are two different ways to change the resolution for the
DEM in the case of photogrammetric point clouds. The first
is changing the point density for the input point cloud and the
second is changing the cell size for the DEM. Both variants are
shown in the figures 7 and 8. Figure 7 illustrates the influence
of the point density to the surface reconstruction. In comparison
figure 8 shows the influence of the chosen cell size for the surface
reconstruction.

The maximal resolution is a ground sampling distance of 6 cm but
there was also two additional outputs created, one with a ground
sampling distance of 40 cm which is corresponding to the sam-
pling distance between two scan lines of the ALS data, the other
with a sampling distance of 80 cm to show the influence of sparse
point clouds. A comparison of the output DTM shows that the
noise as well as the level of details is lower by using a higher
resolution. Furthermore, the riverside gets smaller when using a
higher resolution. The riverside is the only area which gives the
structure in the otherwise flat landscape. Therefore, a resolution
of 80 cm seems to be to big to create suitable key planes for this
scene. Thus, the chosen point cloud resolution are the maximal
resolution of 6 cm and 40 cm.

The cell size for the 2.5D DEM is chosen to be the same for
the ALS data and the photogrammetric point cloud. There are
three different cell sizes considered, one to be 0.8 m which holds
enough details and results in a suitable accuracy for a coarse reg-
istration. The other is a cell size of 1.6 m to evaluate the effect of
doubling the cell size. First results are generated by using the cell
size of 0.8m because this cell size on one hand eliminate noise
but keeps small surface structures. The cell size of 1.6m is then
used for comparison.
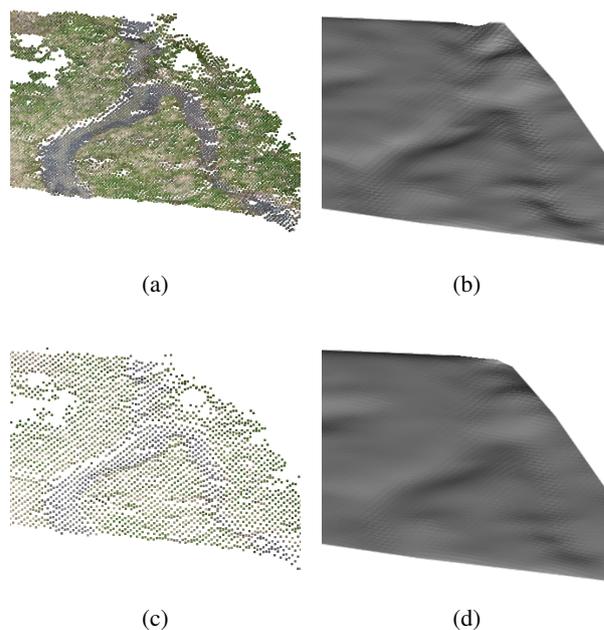
(a)

(b)

(c)

(d)
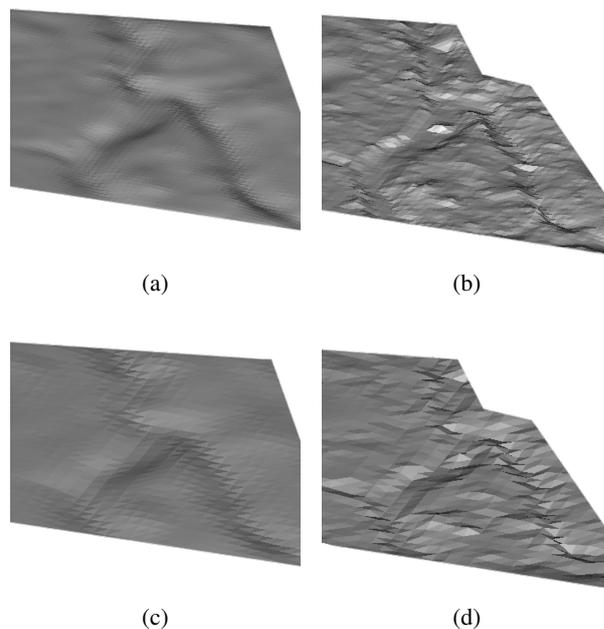
Figure 8. DEMs with different cell sizes top: cell size of 0.8 m
a) DEM for the ALS data and b) DEM for the photogrammetric
point cloud, bottom: cell size of 1.6 m c) DEM for the ALS data
and d) DEM for the photogrammetric point cloud

By comparing the two ways of changing the resolution it turns out
that the changing of the sampling distance have a higher impact
of changing noise and details as the changing of the cell size for
the surface reconstruction.

5.2 Detected key planes

The decision to consider a cell as key plane depends on two val-
ues. First the size of the mask and second the maximum step
size of a key plane histogram. The latter is set to a fixed value

of 10 bins inside a histogram of 100 bins. This is a value which gives some room for noise but is not to big to create to much key planes. The size of the mask is tested with different values on the different resolution DEMs. The best mask size is chosen with a visual interpretation of the results. There are more normals taken inside the mask when using a higher mask size. Therefore, also more normals with a different direction than the center normal are considered. In case of a flat area, considering more normals results in a higher potential to create a step in the otherwise flat graph. Therefore, a higher mask size detects more key planes in the area of the riverside which shows the most changes of the surface normals. Figure 9 shows the detected key planes for the two point clouds. The used mask size for the descriptor was chosen to be 5 m x 5 m for the cell size of 0.8 m. For the cell size of 1.6 m a mask size of 16 m x 16 m was used, a smaller one creates no suitable count of key planes.

(a)   (b)

(c)   (d)

(e)   (f)
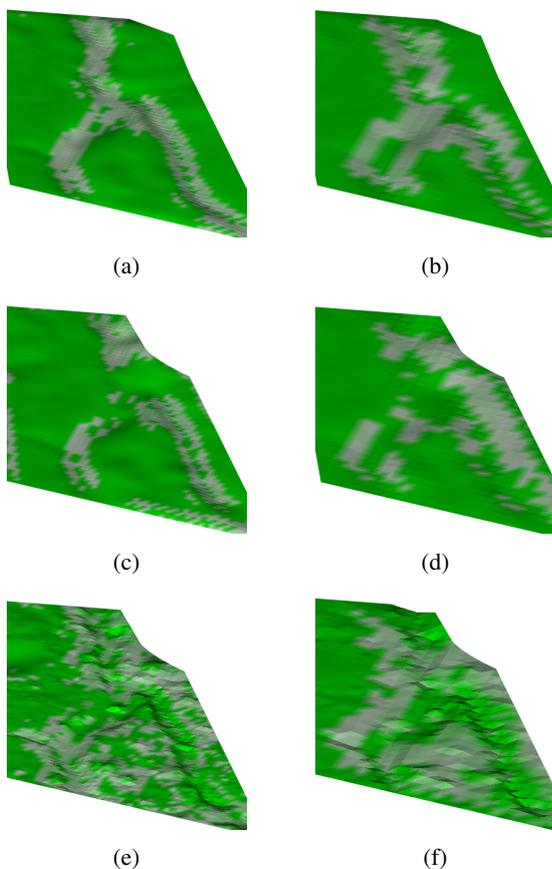
Figure 9. Detected key planes on the DEMs with different resoultions, key planes are drawn in gray. First row: DEMs for the ALS data a: cell size of 0.8 m b: cell size of 1.6 m; middle row: DEMs for the photogrammetric point cloud with a resolution of 40 cm with cell size of 0.8 m (c) and 1.6 m (d); last row: photogrammetric point cloud with a resoultion of 6 cm with cell size of 0.8 m (e) and 1.6 m (f)

By looking on the photogrammetric point cloud, it is obvious that noise also creates key planes. The reconstructed surface of the potogrammetric point cloud shows more hills than the ALS surface. The surface normals changes in the area around these hills and therefore they create some key planes. In both cases (the ALS point cloud and the photogrammetric one) there are lots of detected key planes in the area of the river side which have potential to create true matches.

5.3 Registration

(a)

(b)
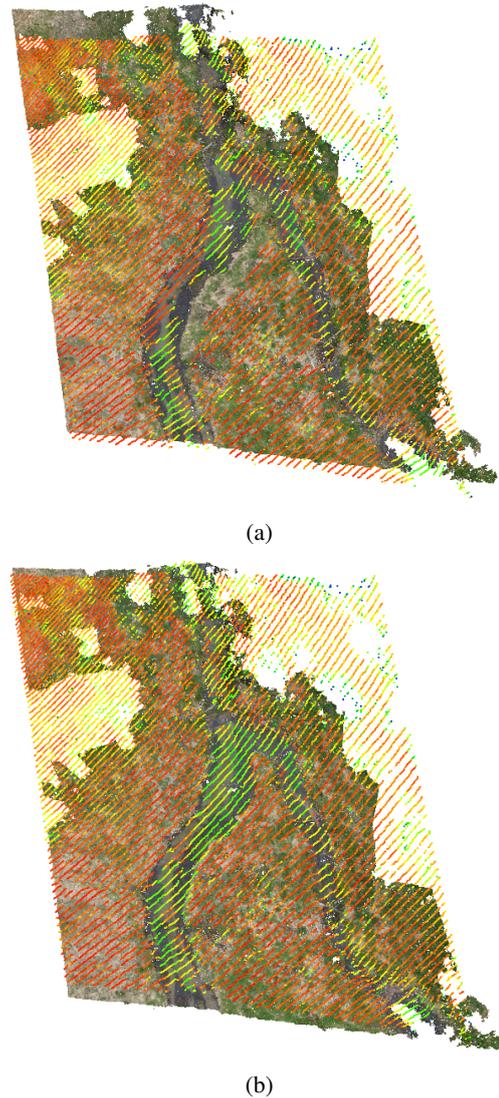
Figure 10. Registration output, shown are the overlayed two point clouds a: output of the proposed method with left misalignement of 3 m b) ground truth.

The Registration output is evaluated in two ways. First by the use of already registered data sets which are for example two photogrammetric point clouds with different resolution. And second by the use of the multi sensor dataset which was manually registered to create ground truth. The initial values for the transformation are set to be around 25 m false and the accuracy threshold was set to be 50 m. This values simulate the measurement of a UAV mounted GNSS antenna with low accuracy.

Using two photogrammetric point clouds with different resolutions shows that the calculated translation have a left error of about 1 m. Considering the used cell size of 80 cm and the fact that the key planes are matched with halve cell accuracy this resulting transformation is expectable. Using the cell size of 1.6 m results in a misalignment of up to 3 m. Furthermore, this shows that the proposed method is independent of high difference in the point cloud resolutions. In comparison with the use of multisensor data the left transformation error is bigger. Considering the ground truth for the registration of the ALS to the photogrammetric point cloud the misalignment is about 3 m instead of the

1 m mentioned above. This result is also shown in figure 10. One reason for a lower accuracy in this case could be the mentioned aspects in section 4. Another reason could be the influence of the noise of the photogrammetric point cloud. But the use of a lower resolution of the photogrammetric point cloud with lower noise creates the same result. Therefore, the misalignement seems to be more connected to the difference of the reconstructed surface cells. Therefore, the proposed method seems to have the following pros and cons: It seems to be robust against noise and point cloud resolutions, works with the landscape itself and is therefore independent of man made objects and works with multi sensor data as well. At the other hand the accuracy is directly limited by the used cell size for the surface reconstruction.

6. SUMMARY

This paper shows an approach for coarse registration between point clouds from different sensors of non urban scenes. For this a regular DEM is used and a descriptor for key cells on this DEM is defined. All potential similar descriptors are matched and the final correspondences are searched with a RANSAC based approach. With the proposed approach it is possible to register two point clouds with an accuracy depending on the cell size of the 2.5D DEM. However, this accuracy needs to be improved for fine registration approaches, and the proposed approach works also well for landscapes which would be hard to register without markers.

Future work could be to adapt the Förstner operator to the cell descriptor instead of using the histogram and evaluate if a more complex descriptor creates better results. After the coarse registration using the key-cells, the correspondences for all cells in both DEMs are estimated with a nearest neighbour approach and the fine registration result could be calculated by minimizing the cell distances. Another future adaptation would be to use a total least square approach to register a point cloud to a DEM by minimizing point to DEM distances. This could provide a multi sensor registration in real time by considering each new measured point for the calculation of transformation parameters. Possible applications of such real time coarse registration would be for examples an online densification or online detection of changes of a reference ALS data.

ACKNOWLEDGEMENTS

REFERENCES

Besl, P. J. and McKay, N. D., 1992. Method for registration of 3-d shapes. In: *Robotics-DL tentative*, International Society for Optics and Photonics, pp. 586–606.

Boerner, R., Hoegner, L. and Stilla, U., 2017. Voxel based segmentation of large airborne topobathymetric lidar data. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XLII-1/W1, pp. 107–114.

Dold, C. and Brenner, C., 2006. Registration of terrestrial laser scanning data using planar patches and image data. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 36(5), pp. 78–83.

Förstner, W. and Wrobel, B. P., 2016. *Photogrammetric Computer Vision*. Vol. 11, Springer.

Ge, X., 2017. Automatic markerless registration of point clouds with semantic-keypoint-based 4-points congruent sets. *ISPRS Journal of Photogrammetry and Remote Sensing* 130, pp. 344–357.

Ge, X. and Wunderlich, T., 2016. Surface-based matching of 3d point clouds with variable coordinates in source and target system. *ISPRS Journal of Photogrammetry and Remote Sensing* 111, pp. 1–12.

Habib, A., Detchev, I. and Bang, K., 2010. A comparative analysis of two approaches for multiple-surface registration of irregular point clouds. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*.

Habib, A., Ghanma, M., Morgan, M. and Al-Ruzouq, R., 2005. Photogrammetric and lidar data registration using linear features. *Photogrammetric Engineering & Remote Sensing* 71(6), pp. 699–707.

Stamos, I. and Leordeanu, M., 2003. Automated feature-based range registration of urban scenes of large scale. In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol. 2, IEEE, pp. II–Ii.

Theiler, P. W., Wegner, J. D. and Schindler, K., 2014. Keypoint-based 4-points congruent sets–automated marker-less registration of laser scans. *ISPRS Journal of Photogrammetry and Remote Sensing* 96, pp. 149–163.

Von Hansen, W., 2006. Robust automatic marker-free registration of terrestrial scan data. *Proceedings of Photogrammetric Computer Vision* 36, pp. 105–110.

Wang, J., Lindenbergh, R., Shen, Y. and Menenti, M., 2016. Coarse point cloud registration by egi matching of voxel clusters. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* pp. 97–103.

Weber, T., Hänsch, R. and Hellwich, O., 2015. Automatic registration of unordered point clouds acquired by kinect sensors using an overlap heuristic. *ISPRS Journal of Photogrammetry and Remote Sensing* 102, pp. 96–109.

Weinmann, M., Weinmann, M., Hinz, S. and Jutzi, B., 2011. Fast and automatic image-based registration of tls data. *ISPRS Journal of Photogrammetry and Remote Sensing* 66(6), pp. S62–S70.

Xiao, J., Adler, B. and Zhang, H., 2012. 3d point cloud registration based on planar surfaces. In: *Multisensor Fusion and Integration for Intelligent Systems (MFI), 2012 IEEE Conference on*, IEEE, pp. 40–45.

Xiao, J., Adler, B., Zhang, J. and Zhang, H., 2013. Planar segment based three-dimensional point cloud registration in outdoor environments. *Journal of Field Robotics* 30(4), pp. 552–582.

Xu, Y., Boerner, R., Yao, W., Hoegner, L. and Stilla, U., 2017. Automated coarse registration of point clouds in 3d urban scenes using voxel based plane constraint. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* IV-2/W4, pp. 185–191.

Yang, B. and Zang, Y., 2014. Automated registration of dense terrestrial laser-scanning point clouds using curves. *ISPRS Journal of Photogrammetry and Remote Sensing* 95, pp. 109–121.

Yang, B., Dong, Z., Liang, F. and Liu, Y., 2016. Automatic registration of large-scale urban scene point clouds based on semantic feature points. *ISPRS Journal of Photogrammetry and Remote Sensing* 113, pp. 43–58.