

Article

Registration of Multi-Sensor Bathymetric Point Clouds in Rural Areas Using Point-to-Grid Distances

Richard Boerner ^{1,*}, Yusheng Xu ¹ , Ramona Baran ², Frank Steinbacher ², Ludwig Hoegner ¹ and Uwe Stilla ¹

¹ Photogrammetry and Remote Sensing, Technical University of Munich, 80333 Munich, Germany; yusheng.xu@tum.de (Y.X.); ludwig.hoegner@tum.de (L.H.); stilla@tum.de (U.S.)

² Steinbacher Consult Ingenieurgesellschaft GmbH & Co. KG, Neusaess, 86356 Augsburg, Germany; r.baran@steinbacher-consult.com (R.B.); f.steinbacher@steinbacher-consult.com (F.S.)

* Correspondence: richard.boerner@tum.de

Received: 20 March 2019; Accepted: 2 April 2019; Published: 5 April 2019



Abstract: This article proposes a method for registration of two different point clouds with different point densities and noise recorded by airborne sensors in rural areas. In particular, multi-sensor point clouds with different point densities are considered. The proposed method is marker-less and uses segmented ground areas for registration. Therefore, the proposed approach offers the possibility to fuse point clouds of different sensors in rural areas within an accuracy of fine registration. In general, such registration is solved with extensive use of control points. The source point cloud is used to calculate a DEM of the ground which is further used to calculate point to raster distances of all points of the target point cloud. Furthermore, each cell of the raster DEM gets a height variance, further addressed as reconstruction accuracy, by calculating the grid. An outlier removal based on a dynamic threshold of distances is used to gain more robustness against noise and small geometry variations. The transformation parameters are calculated with an iterative least-squares optimization of the distances weighted with respect to the reconstruction accuracies of the grid. Evaluations consider two flight campaigns of the Mangfall area in Bavaria, Germany, taken with different airborne LiDAR sensors with different point density. The accuracy of the proposed approach is evaluated on the whole flight strip of approximately eight square kilometers as well as on selected scenes in a closer look. For all scenes, it obtained an accuracy of rotation parameters below one tenth degrees and accuracy of translation parameters below the point spacing and chosen cell size of the raster. Furthermore, the possibility of registration of airborne LiDAR and photogrammetric point clouds from UAV taken images is shown with a similar result. The evaluation also shows the robustness of the approach in scenes where a classical iterative closest point (ICP) fails.

Keywords: airborne point clouds; registration; DEM; multi-sensor

1. Introduction

The task of registration is to transfer two data sets of the same scene captured on different times or with different sensors into the same coordinate frame. In the case of the registration of multi-sensor data, the use of geometric features is needed since the radiometry can be very different. Furthermore, in the case of 3D data linear structures, planar structures, or point correspondences are used as geometric features. However, these features have the disadvantage that using point correspondences as used in the iterative closest point (ICP) [1] algorithm is limited to point clouds with almost the same point distribution and density. Using linear or planar structures derived from the point clouds overcome the limitation of different point densities, but these are mainly limited to scenes with human-made structures. In rural areas, these solutions are not applicable in most cases.

This paper is an extension of a previously published approach for initial coarse registration [2]. Instead of the initial registration, the proposed method on this paper can be seen as a fine-registration method with higher accuracy than the previous one. This paper proposed another possibility to calculate the DEM and the variances of the DEM parameters. The calculation method can generate DEMs for large-scale data such as the whole flight strip and saves memory and computation time. The evaluation in the pilot study [3] was done only on a single scene which shows degrees of freedom along the rotation angle around the z-axis and the translation along the x-axis. To acquire knowledge about the performance of the proposed method in different scenarios and at different level of detail of the data, this paper shows a more complex evaluation of the method. Also, the scenes of the evaluation are chosen to fulfil all the degrees of freedom of the transformation parameters.

Therefore, the proposed approach uses a more general technique which does not rely neither on the point cloud density nor linear or planar structures. The registration approach is designed to be used for change detection in river areas so the registration must mostly deal with rural areas. The proposed approach assumes that the ground is in most of the whole scene and temporally stable. Therefore, the whole scene should be chosen to be large enough to exclude small zone changes or landslides. With this assumption the proposed approach uses the ground for registration of the data sets. The source point cloud is used to generate a DEM based on the classified ground points. The calculation of the transformation parameters is based on a least-squares optimization of point-to-grid distances between the source grid and each point of the target point cloud. An outlier removal is used to ensure robustness of the least-squares approach against noise and possible changes of the landscape.

Possible applications of the provided registration approach are the enhancement of 3D point clouds derived from aerial laser scans with a photogrammetric point cloud of higher point density derived from UAV data. Since the approach uses the geometry, it is not limited to one sensor and can also be used to compare point clouds captured from NIR-Cameras, RGB-Cameras, and LiDAR sensors. It could also be possible to compare LiDAR point clouds captured in different flight heights or with different beam divergence which both results in different ground-sampling distances.

Some state-of-the-art methods in the context of registration are briefly given in Section 2. The proposed method is described in Section 3 and the used data for evaluation is shown in Section 4. Experiments for the evaluation are described and evaluation results are given in Section 5, and a final discussion and conclusion is given in Section 6.

2. State of the Art

The widely used approach for point-based registration is the ICP algorithm, which minimizes point-to-point distances [1]. This algorithm as well as its variants [4,5] are very effective regarding accuracy but need an accurate initial alignment, and the matching process can be time-consuming. Algorithms which reduce the searching space for correspondences are based on key points which adopt classic key point descriptors of the photogrammetry and transfer their idea to the 3D case. There are SIFT-based [6,7], DoG-based [8] or FPFH key point descriptors [9]. There are also algorithms which search for intersection points or semantic points to consider these as features [10–12]. All these methods have the advantage that they create a good alignment of different point clouds. However, they have the disadvantage that they get into troubles when dealing with noisy data, large-scale data or point clouds with different levels of detail or point density.

There are also high-level geometric features which can theoretically increase the robustness of the matching process and overcome the disadvantages of different point densities and level of detail. These high-level primitives are mainly lines or planes. For line features spatial curves [13], 3D straight lines [14] or intersection lines of neighboring planes [15] can be found. For plane features, approaches are searching for plane correspondences [16–19]. However, the extraction of planar surfaces by model-fitting or region growing algorithms can be time-consuming [20,21]. Therefore, to simplify the data set, plenty of studies in the field of segmentation use a rasterized representation of volume elements (voxel). These voxels can be assigned attributes derived from their included points such as

planarity [21], last pulse property [22], or normal direction [23]. These attributes can further be used for supervoxelization or to define geometric constraints in voxel groups [24]. The approaches using voxels for solving registration problems are mainly located in the field of coarse registration and can be found using EGI features [20] or planar surfaces extracted from the raw point cloud [21].

While the point-based approaches are mostly located in the field of fine registration because of high accuracy, the high-level primitives can overcome problems in different point densities. Furthermore, the high-level primitives mostly focus on residential or other urban areas and therefore get into trouble when dealing with rural areas such as in bathymetric applications.

3. Method

Compared to the ICP [1], the proposed method does not assume point correspondences. Instead it minimizes distances of 3D points to the 2.5D raster of a DEM. Therefore, the proposed approach is independent of point density because of a geometric generalization of the source point cloud and considering the target point cloud at point level. The DEM is generated with the use of a voxel structure which filters ground points and creates information about the variance of the points coordinates inside the voxel. These variances are further used to get stochastic information about the DEM accuracy. The accuracies of each DEM cell are used in the registration to weight the observed distances and an outlier filtering is used to gain more robustness. This makes the proposed approach able to consider measurement accuracies or system information of the used scan system in contrast to ICP variants [4,5] or other geometric primitive-driven approaches [16–19,21]. Algorithm 1 shows the proposed approach. Detailed explanations for the distance minimizing approach are given in Section 3.1 and the included regularization for outlier filtering is explained in Section 3.2.

Algorithm 1: Algorithm overview.

```

Data: LiDAR and Photogrammetric point cloud
Result: transformation parameters
G = calculate DEM(LiDAR);
t = initial threshold;
M = initialize Model();
while iteration < maximal iteration do
  foreach point  $p \in$  photogrammetric point cloud do
    d = calc distance point to grid (G,p);
    fill histogram (d);
    if  $d < t$  then
      | update Model(M,G,p);
    end
  end
   $\Delta x$  = calculate Parameter update();
  if  $\Delta x < \epsilon$  then
    | return final parameters;
  else
    | update Parameters( $\Delta x$ );
  end
end

```

3.1. The Minimizing of Point to DEM Distances

The source data is used to generate a DEM based on the ground points as a first pre-processing step of the proposed approach. The classification of ground points is done with a voxel-based approach by considering the pulse echoes [22]. This approach creates a regular 3D grid to specify the voxels

and their location. Each point of the input point cloud is added to its corresponding voxel and the attributes of the points (number of returns, return number) are summed up for all points in the voxel in adaptive manner. A 3D point on the DEM is further defined by the 3D coordinates $(x; y; G(x, y))^T$. To calculate the height value of given x-y coordinates, bilinear interpolation of the point inside a single cell of the DEM is used (Equation (1)):

$$\begin{aligned} G(x, y) = & (1-s)(1-t) \cdot a_{i,j} \\ & + (1-s)t \cdot a_{i,j+1} \\ & + (1-t)s \cdot a_{i+1,j} \\ & + s \cdot t \cdot a_{i+1,j+1} \end{aligned} \quad (1)$$

with:

$s = x - x_i$: difference inside the range $[0, 1]$ of the x coordinate of the current point to the x coordinate of the anchor edge (x_i) on the grid.

$t = y - y_j$: analogue to s with the y coordinates of the current point.

a = heights of the grid cells edge points (i.e., estimated grid parameter).

i, j = coordinates of the anchor edge of the current grid cell, corresponding to the given x, y values.

In contrast to memory and time-consuming global optimization [3], this local interpolation estimates the height of each cell with inverse distance weighting. This allows calculation of large-scale grids in an adaptive manner and avoids the squared complexity of the global optimization.

The weights of the weighted interpolation are calculated as inverse Euclidean distance (Equation (2)) with:

$$w_{k,l} = \frac{1}{(x_k - x_l)^2 + (y_k - y_l)^2 + z_k^2} \quad (2)$$

where:

x_k, y_k, z_k = coordinates of the point k .

$w_{k,l}$ = weight for point k to the cell's edge l .

x_l, y_l = coordinates of the cell's edge l .

The parameter on the raster can be calculated with the inverse weighting average (Equation (3)) as:

$$a_l = \frac{1}{\sum_{k \in N} w_{k,l}} \cdot \sum_{k \in N} w_{k,l} \cdot z_k \quad (3)$$

where:

a_l = grid parameter corresponding to the cell's edge l .

N = neighborhood of considered points.

This weighted average is chosen to get a moderate geometric representation of complex landscapes. There are other interpolation methods possible which use more complex mathematical representations, but they are extensive in computation at large scales and difficult to model considering complex natural scenes. Furthermore, by considering a variance of the height value, the resulting reconstruction accuracy ($stdZ_l$) of each parameter l is calculated as:

$$stdZ_l^2 = \frac{1}{\left(\sum_{k \in N} w_{k,l}\right)^2} \cdot \sum_{k \in N} w_{k,l}^2 \cdot \sigma_{z,k}^2 \quad (4)$$

The DEM is calculated based on the average of points inside small voxels. The voxel size is chosen as the expected GSD of the data set to hold as much detail as possible. Since the scene is captured with multiple flight strips, there is redundant information inside the voxel cells. The variance of the point coordinate is calculated as the variance of the average of the points inside a single voxel. The reconstruction accuracy (Equation (4)) will be further used to weight the distance of a single point to the DEM in the registration approach. Points in areas where the variance of height is high will get a lower weight and therefore have a lower impact on the registration. Such areas with higher variance are expected at regions below dense vegetation where fewer ground points are measured and at steep slopes. In case of steep slopes, the variance of the average points inside the voxel will be higher. Since the calculation of the DEM parameter and its variance is an equal weighted average of the points along the steep slope, the higher variances of these points will be transferred to a higher variance of the DEM parameter.

The parameters of the transformation from the target to the source system are the ones of the 3D similarity transformation given the point $\mathbf{p} = (x, y, z)^T$

$$\begin{pmatrix} x_t(\mathbf{p}) \\ y_t(\mathbf{p}) \\ z_t(\mathbf{p}) \end{pmatrix} = m \cdot \mathbf{R}(\alpha, \beta, \gamma) \begin{pmatrix} x - x_0 \\ y - y_0 \\ z - z_0 \end{pmatrix} \quad (5)$$

with:

x_0, y_0, z_0 = translation along x, y, z -axis.

x, y, z = coordinates of the point to be transformed.

m = scale.

$\mathbf{R}(\alpha, \beta, \gamma)$ = rotation with the Euler angles α, β, γ .

The parameters of the translation, rotation, and scale are estimated by using a least squares minimization of point-to-grid distances. For the least-squares approach the distance of each point of the target point cloud $\mathbf{p}_n = (x_n \ y_n \ z_n)^T$ to the DEM from the source point cloud is calculated by the observation function:

$$f_n = G(x_t(\mathbf{p}_n), y_t(\mathbf{p}_n)) - z_t(\mathbf{p}_n) \quad (6)$$

where the transformed coordinates are generated using Equation (5) with the use of \mathbf{p}_n . The transformation parameters are further estimated by the minimization of

$$\Omega = \sum_n \frac{1}{\sigma_n} f_n^2 \quad (7)$$

where σ_n is the variance of the observation n .

Solving the registration by the least-squares method gives an analytical formula (Equation (8)) for the transformation parameters:

$$\hat{\mathbf{x}} = \left(\mathbf{A}^T (\mathbf{BQB}^T)^{-1} \mathbf{A} \right)^{-1} \mathbf{A}^T (\mathbf{BQB}^T)^{-1} \mathbf{f} \quad (8)$$

with:

$\hat{\mathbf{x}}$ = vector of the parameters e.g., $(p_1, p_2, p_3, p_4, p_5, p_6)^T = (x_0, y_0, z_0, \alpha, \beta, \gamma)^T$.

\mathbf{A} = functional matrix (derivatives of the observation functions (Equation (6)) with respect to the

$$\text{parameters } p_i) = \begin{pmatrix} \frac{\delta f_1}{\delta p_1} & \dots & \frac{\delta f_1}{\delta p_m} \\ \dots & \dots & \dots \\ \frac{\delta f_n}{\delta p_1} & \dots & \frac{\delta f_n}{\delta p_m} \end{pmatrix}.$$

\mathbf{B} = matrix of conditional observations (derivatives of the observation functions (Equation (6)) with

$$\text{respect to the coordinates of } \mathbf{p}_n = \begin{pmatrix} \frac{\delta f_1}{\delta x_1} & \frac{\delta f_1}{\delta y_1} & \frac{\delta f_1}{\delta z_1} \\ \dots & \dots & \dots \\ \frac{\delta f_n}{\delta x_n} & \frac{\delta f_n}{\delta y_n} & \frac{\delta f_n}{\delta z_n} \end{pmatrix}.$$

\mathbf{Q} = covariance matrix of observations (i.e., stochastic model).

\mathbf{f} = vector of f_n for each point.

For simplification, it is assumed that there are no correlations between the 3D points of the target cloud or between the cells of the source grid. By ignoring possible correlations, the accuracy of the final registration may be lower, but because of the diagonal weighting matrix, no observation matrix is needed to be saved. Therefore, this kind of simplification saves computation time and reduces the needed memory. Notably, by considering point clouds with several 100 million to even billions of points, memory saving gets very important.

With no correlations, the multiplication of \mathbf{BQB}^T becomes a diagonal matrix with scalar values for a single point (Equation (9)). For a single observation point \mathbf{p}_n it becomes

$$bqb_n = \sum_{j=1}^3 \left(\frac{\delta f_n}{\delta x_j} \right)^2 \sigma_{n;j} + \sigma_{z;n} \quad (9)$$

with $\sigma_{n;j}$ as the variance of the j -th coordinate of point n and $\sigma_{z;n}$ as the variance of the height from point n on the DEM. The latter variance is calculated using the reconstruction accuracy from Equation (4). Therefore, observations inside grid cells with low reconstruction accuracy are considered with lower weight.

Because of the non-linearity of the functional model, an iterative process is needed to solve the derivatives. Some initial values for the transformation parameters are fine-tuned in each iteration step. This iterative process stops if either a maximum number of iterations is reached or if the parameter update is small enough to be considered as convergence.

3.2. Outlier Removal

The proposed approach uses outlier removal to enhance robustness against noise. Since the grid is generated based on ground points, outliers can also be classified directly as non-ground points. This outlier removal becomes very important when working with photogrammetric point clouds where ground filtering can be very complex.

Figure 1 shows a scheme for the calculated dynamic threshold for outlier removal during the iterative registration process. Each point where distance is below the threshold is considered to be an outlier and therefore not taken as an observation. At every step of the iteration, the threshold is recomputed with a histogram-based approach. The histogram is built with all distances of the points of the target point cloud. There are some assumptions for the definition of the threshold: First, the initial values are considered to be good enough to create an overlap of ground points, and second, ground points should be the major part of the points inside the target data. These assumptions result in the expectation that the highest bin of the histogram can be found at lower bin IDs (i.e., on the left side of the histogram). Therefore, the threshold is defined by the bin with a percentage below the highest bin and on the right side of the highest bin. Figure 1 perfectly shows the determination of the bin ID of the threshold (t) in the histogram of distances. This bin ID is used to transfer threshold t to a distance value. All points with distances above the distance value corresponding to threshold t are handled as outliers and therefore not considered for registration. The values are the number of points corresponding to the distance, which is discretized with the corresponding bin. This outlier detection is done at each iteration step of the registration. At each iteration, the histogram is calculated again with the current transformation parameters. Since most of the points will have a lower distance to the DEM after each iteration, the highest bin in the histogram will move to the left. Therefore, the threshold for outlier removal will be lower in the next iteration too. The chosen value for the fixed percentage should be

small enough to include most of the expected ground points. Since the ground points are expected at the highest bin, a small percentage value will mark the bin where most points close to the ground are included.

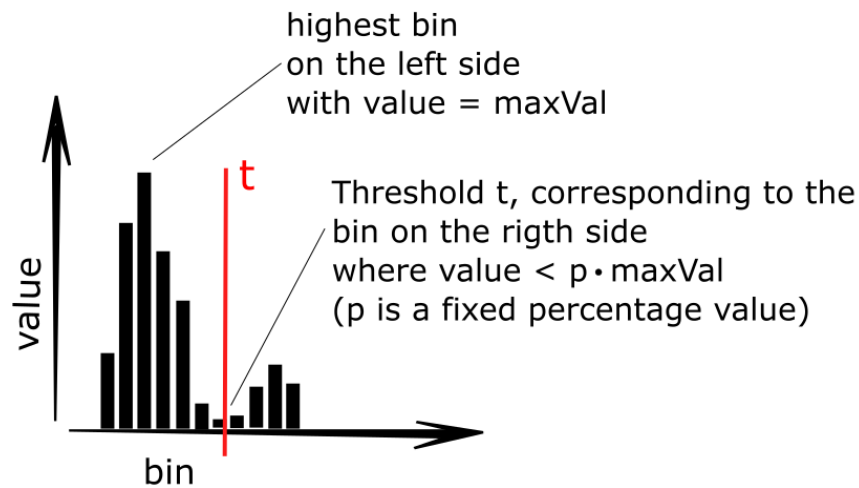


Figure 1. Scheme of threshold definition: The bins discretize the point-to-grid distances. Values are the number of points corresponding to one bin. The threshold t is searched on the right side of the bin with maximal value.

4. Data

The proposed approach is evaluated based on the data of two LiDAR flight campaigns along the Mangfall area in Bavaria, Germany. The area shows a river valley with a lake (Tegernsee) in the south and a highway (A8) in the north. An overview of the landscape as the processed DEM is shown in Figure 2. The flight campaigns are done in December 2012 and in April 2017. Both datasets consist of an airborne LiDAR bathymetric point cloud which is captured from an airplane and used for testing the registration approach. Therefore, the ground layer is considered to be continuous without gaps in the water areas. Furthermore, the registration must deal with the possibility of changed geometry. The whole area of these two flight campaigns has a length of nearly 17 km and a width of more than 500 m. The river itself and therefore the water area has a width of several meters. Possible influences of the registration due to the rivers morphodynamic character are considered to be an outlier and filtered with the proposed approach. During the flight campaign in 2012 the scanning system Riegl VQ820G was used, and during 2017 the system Riegl VQ880G was used. These two different scan systems differ in the amount of possible maximal recognized pulses along a single laser beam as well as in the ground-sampling distance using the same flight height. Inside a single flight strip the ground-sampling distance of the VQ820G system is about 0.5 m along the scan line as well as between the scan lines. For the VQ880G system, the ground sampling along a single scan line is smaller and holds 0.1 m. The maximum amount of recognizable pulses along a single laser beam differ between 4 for VQ820G and 8 for VQ880G. Since the data is co-registered with a direct georeferencing using the GNSS and IMU of the airplane, the data from the flight campaigns is used as the primary data for testing the registration accuracy and correctness of the proposed approach.

There is also a photogrammetric point cloud derived from a UAV flight campaign used for co-registration which shows the possibility of registering multi-sensor data. The UAV flight was done in 2016 and captured images of a sidearm of the river Mangfall. This UAV flight campaign covers a small part of the airplane data set and shows a flat landscape with a resulting GSD is 6 cm. The water area can be considered as standing water with no critical changes. The scale is estimated using control points. The LiDAR and the photogrammetric point clouds show big differences in point

density. While the photogrammetric point cloud shows a ground-sampling distance of about 6 cm, the LiDAR point cloud shows a GSD of about 10 cm along the scan line and 50 cm between two scan lines. Furthermore, there are big differences in detail due to a laser footprint of approximately 4 dm and in noise due to the photogrammetric calculation. These differences make a registration challenging and allow an evaluation of the proposed method for complex co-registration tasks.

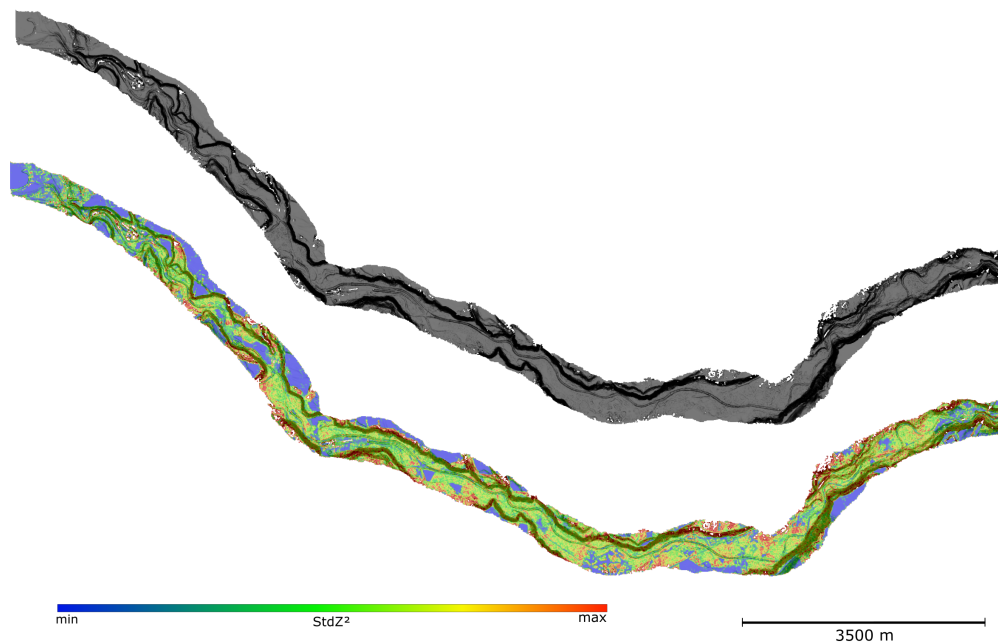


Figure 2. The whole flight strip shown as the processed DEM. Shown as shaded topography (**top**) and the variance of the estimated parameters (**bottom**).

5. Results

5.1. Description of Experiments

The proposed method is evaluated on multiple scenes of the dataset. There is an evaluation on the whole flight strip as well as on closer examination to the scenes (Figure 3). The scenes are chosen depending on the planarity of the landscape, the occurrence of buildings, and the vegetation coverage. The first scene shows a few buildings in the valley of the Mangfall river. The surface itself shows a typical river valley with mountains to both sides. Furthermore, the ground is covered almost completely with vegetation. The second scene shows a small village in a more structured valley compared to the first scene. There is more bare ground on the top of the hill side. The highway bridge was successfully filtered as non-ground by the ground segmentation. While the first two scenes show few buildings, the third one shows a rural area with no human-made objects. The ground is covered nearly entirely with vegetation and hills are unequally distributed along the valley. The surface itself shows a planar valley with a few cliffs at the borders of the scene. The fourth scene shows the most urbanized area at the village Gmund which is the largest village covered in the data set. The surface shows a curved valley with a planar bare ground on top of the mountains, and sparse vegetation cover. The fifth and final scene is similar to the fourth, but it shows more vegetation and more prominent buildings.

These scenes are chosen to evaluate the registration approach for different scenarios. Results are compared with a classical ICP. The challenge consists of the natural scene widely covered with vegetation. Further tests include the influence of different point cloud resolutions which takes an essential role by aiming for multi-sensor applications. For the experiments with different resolutions, the target point cloud is voxelized with different resolutions of the voxel structure. Therefore,

the coarser resolutions also determine a lower level of detail. Beneath the resolution of the target point cloud also the resolution of the source grid is evaluated with respect to its influence on the final registration accuracy.

To evaluate the accuracy of the proposed approach the same scene is registered multiple times with different initial values. The initial values are chosen randomly and within a predefined interval to simulate different sensor outputs used for georeferencing. The use of already geo-referenced data makes the evaluation of correct transformation parameters possible. Since only the initial values are changed, the correct transformation is the one with $x_0 = y_0 = z_0 = 0$ m and $\alpha = \beta = \gamma = 0$ degrees.

To evaluate the influence of initial values, the registration is tested with different initial values which are chosen manually. Furthermore, this empirical test is used to define an interval of initial values where the registration problem can be solved.

The values where no convergence is visible, or where the calculation failed are declared as the values outside of the interval.

Finally, the possibility of multi-sensor registration is shown on the output of the proposed method with the UAV data.

5.2. Results of Experiments

The test for the initial values shows that suitable transformation parameters can be calculated with an initial translation with 20 m error and a rotation with 2 degrees error. To get a visual result, all scenes are registered with the initial values: $x_0 = -17.9$ m, $y_0 = 15.5$ m, $z_0 = 15.1$ m, $r_x = 1.6$ degrees, $r_y = -1.5$ degrees, $r_z = 1.6$ degrees. The initial transformation corresponding to these values and the resulting registration are shown in Figures 3–8.

The results for the presented scenes indicate that the approach can solve the registration problem with different kinds of landscapes. The registration looks suitable for large-scale landscapes (whole flight strip) as well as for close-range scenes. While the whole flight strip shows a mostly rural area, the proposed method shows that no human-made structure is needed for registration.

By evaluating the influence of different resolutions of the source grid, it turns out that the resolution of the grid influences the correctness of the resulting registration but will need fewer iteration steps for convergence at coarser resolutions.

The shown data has differences in smoothness, point density, and details. These differences are the main reasons why the registration is challenging for a classical ICP approach. To compare the output of the proposed method to the one of an ICP approach, the ICP implemented in CloudCompare was used, as input point clouds are the point clouds of the ground points considered. The same initial values as described above are used, and the results of the whole flight strip for both methods are shown on chosen scenes in Figures 9–11.

The main differences are visible in the building footprints. The buildings are filtered in both point clouds and therefore the footprints should match each other. Each scene of the results of the proposed algorithm shows clear footprints but not each scene of the ICP results. While Figure 10 looks quite similar, Figure 11 shows big differences. Also, for the close-range scenes, the result is the same. Therefore, the proposed approach can return a suitable registration while the ICP gets into trouble.

The registration was performed multiple times (20 calculations) as described in Section 5.1 to test the accuracy and correctness of the proposed approach. Each random chosen initial transformation is inside an interval of 8 m in translation and 2 degrees of rotation around zero. The final resulting correctness is shown as average and the accuracy as root mean square error (RMSE) in Tables 1 and 2. These tables also show the influence of different GSD of the target point cloud.

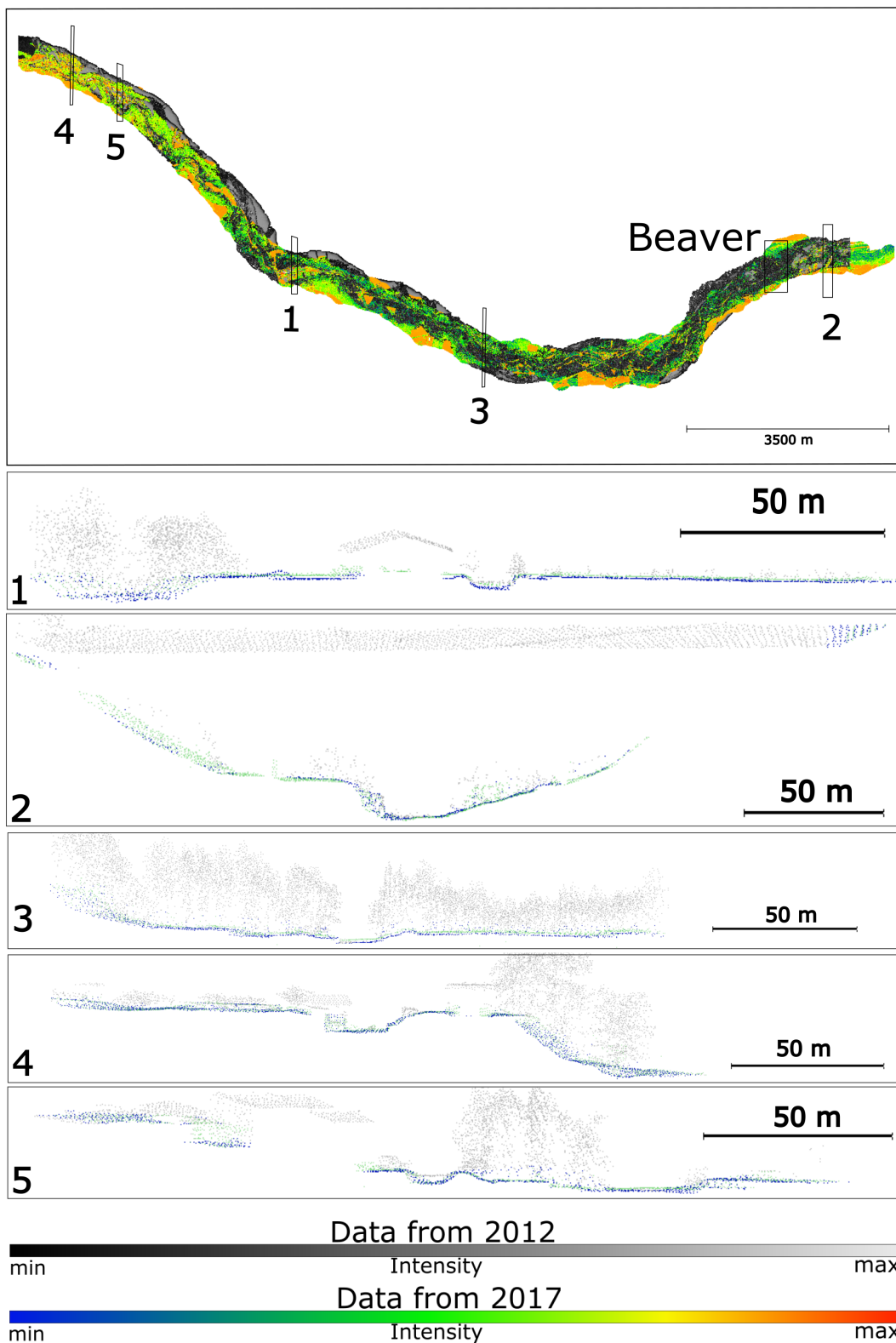


Figure 3. Result of the proposed registration approach shown on the whole flight strip. The marked areas show the location of the profile views of the selected scenes. Shown data are the data of the whole flight strip downsampled to 1 m GSD. The profile views show the data of 2012 and 2017. Ground points from the data of 2012 are colored in blue, non-ground points from 2012 are colored in gray, and ground points of 2017 are colored in green.

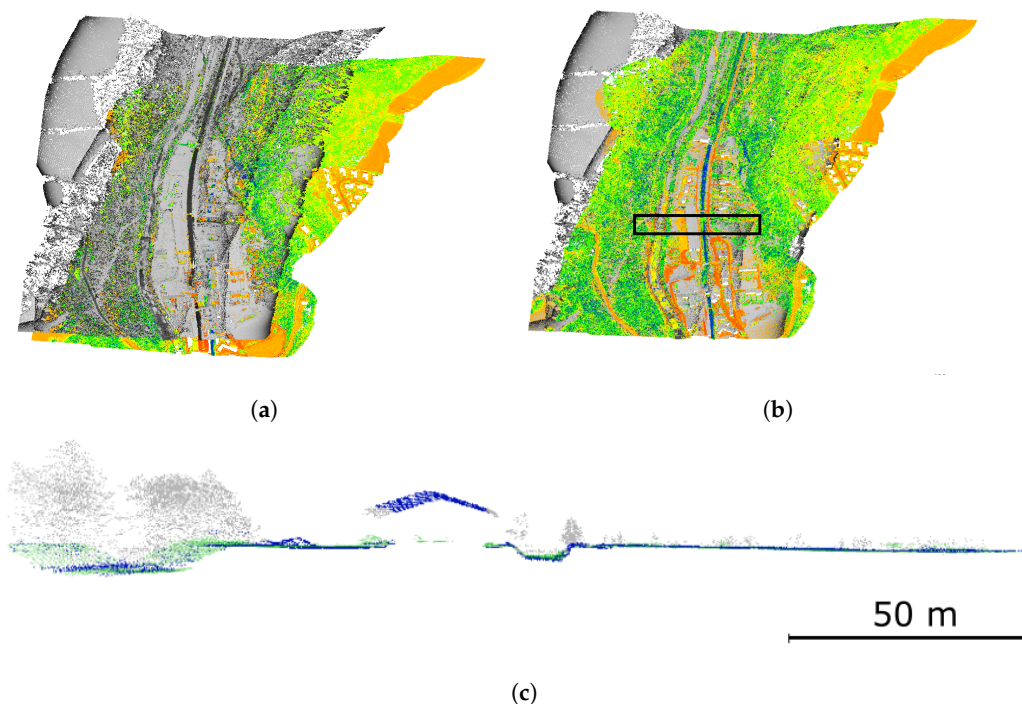


Figure 4. Result of the proposed approach shown on the first scene. (a) Overview of the first scene with initial transformation. (b) Overview of the first scene with registration result. (c) Profile view of the marked area in (b). Shown data are the original data of 2012 and the data of 2017 downsampled to 0.5 m GSD. The color representation is the same as in Figure 3.

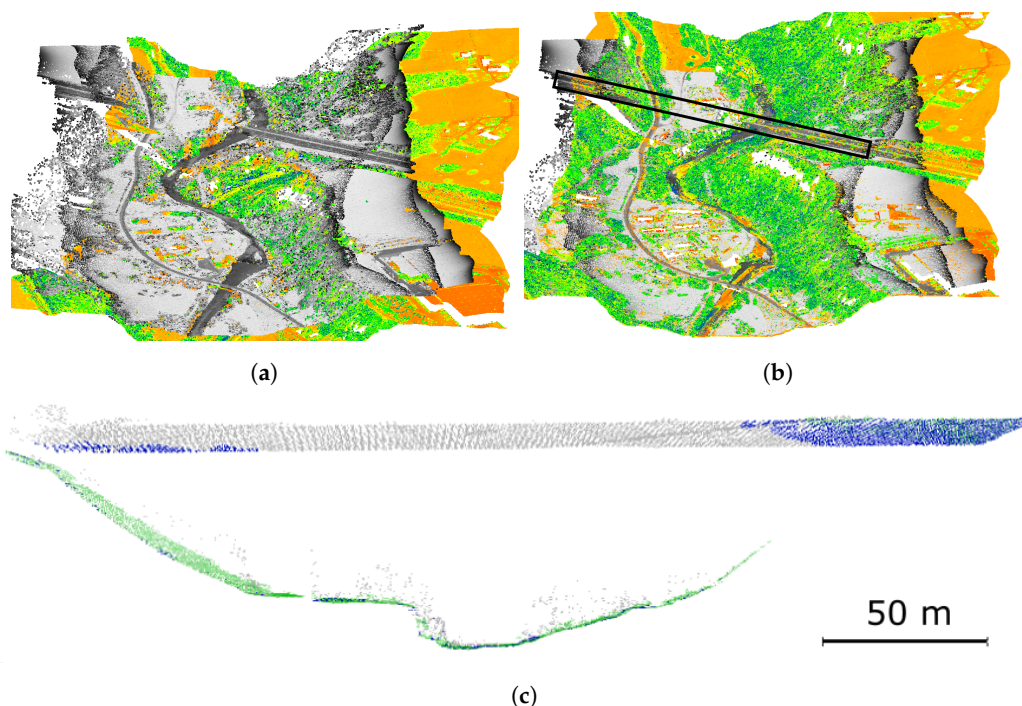


Figure 5. Result of the proposed approach shown on the second scene. (a) Overview of the second Scene with initial transformation. (b) Overview of the second Scene with registration result. (c) Profile view of the marked area in (b). Shown data are the original data of 2012 and the data of 2017 downsampled to 0.5 m GSD. The color representation is the same as in Figure 3.

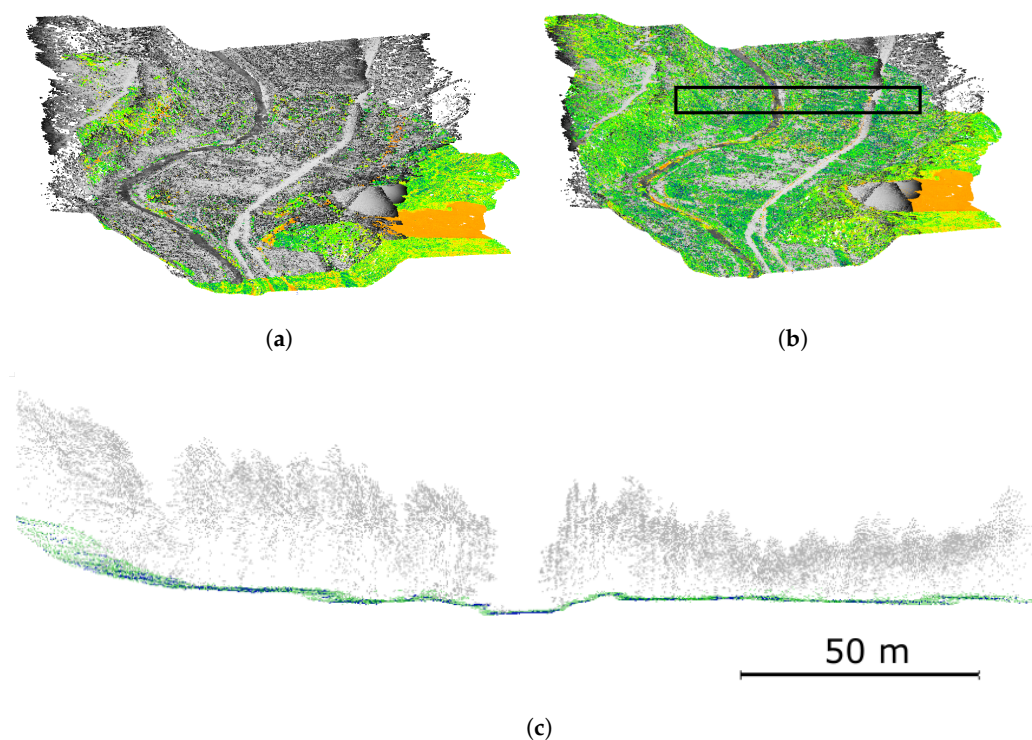


Figure 6. Result of the proposed approach shown on the third scene. (a) Overview of the third scene with initial transformation. (b) Overview of the third scene with registration result. (c) Profile view of the marked area in (b). Shown data are the original data of 2012 and the data of 2017 downsampled to 0.5 m GSD. The color representation is the same as in Figure 3.

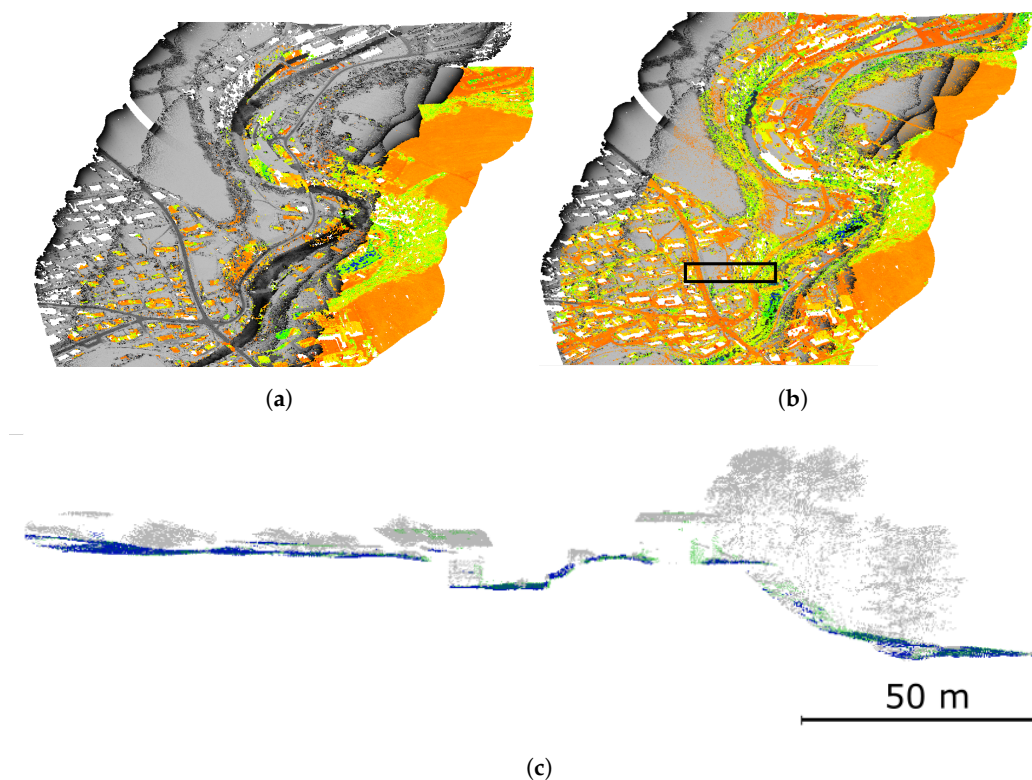


Figure 7. Result of the proposed approach shown on the fourth scene. (a) Overview of the fourth scene with initial transformation. (b) Overview of the fourth scene with registration result. (c) Profile view of the marked area in (b). Shown data are the original data of 2012 and the data of 2017 downsampled to 0.5 m GSD. The color representation is the same as in Figure 3.

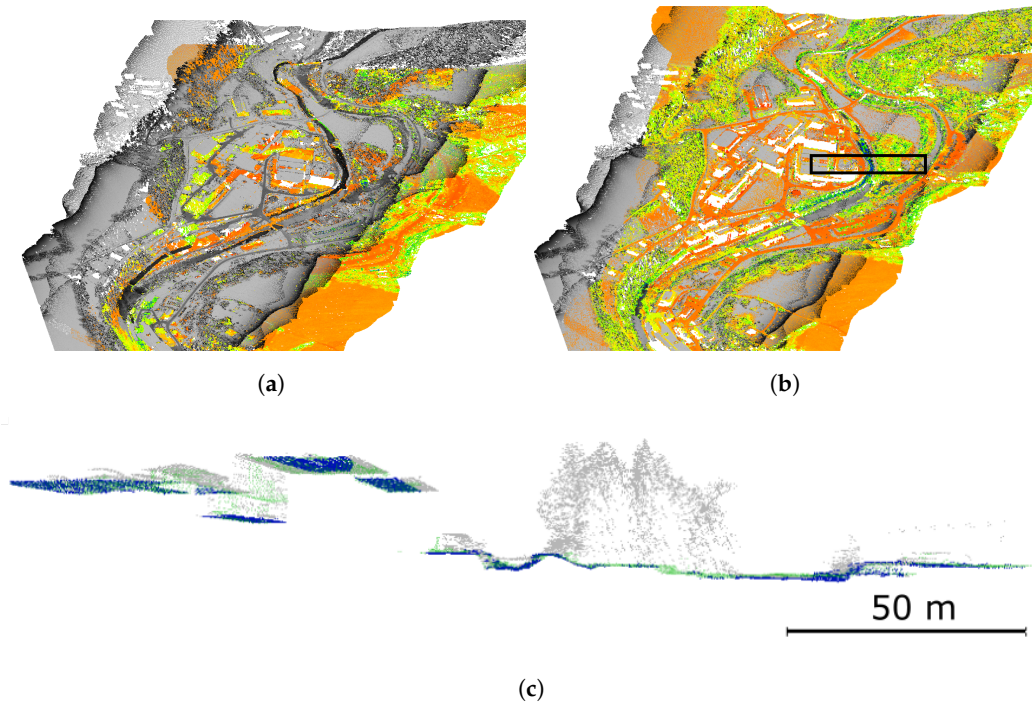


Figure 8. Result of the proposed approach shown on the fifth scene. (a) Overview of the fifth scene with initial transformation. (b) Overview of the fifth scene with registration result. (c) Profile view of the marked area in (b). Shown data are the original data of 2012 and the data of 2017 downsampled to 0.5 m GSD. The color representation is the same as in Figure 3.

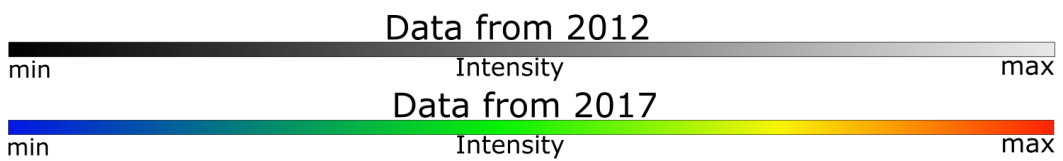
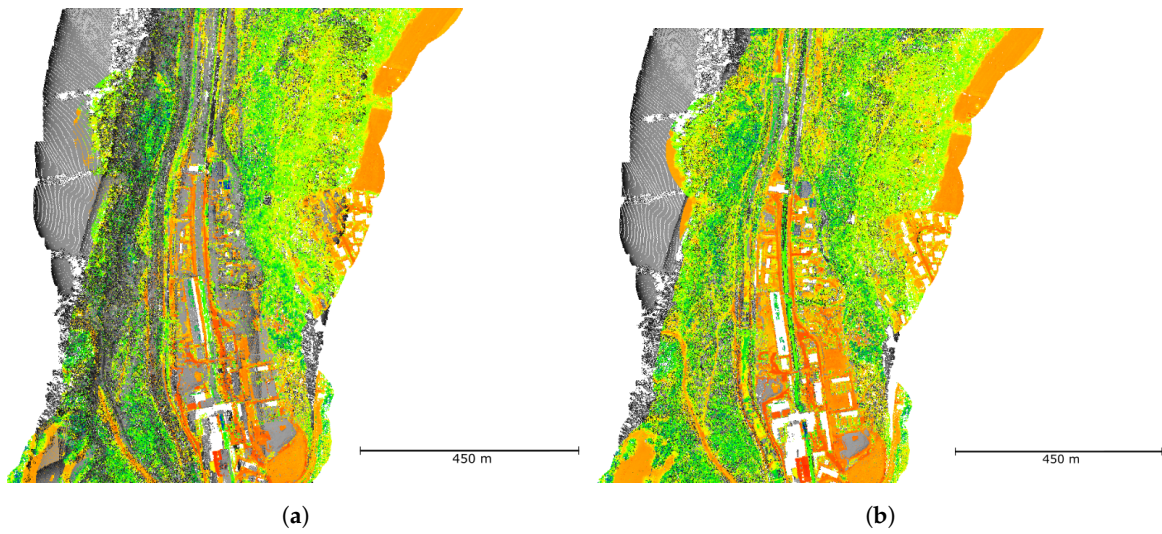


Figure 9. Comparison of the registration results using ICP and the proposed approach shown on the whole flight strip. (a) Result of ICP. (b) Result of the proposed method. Shown data are the date of the whole flight strip downsampled to 1 m GSD.

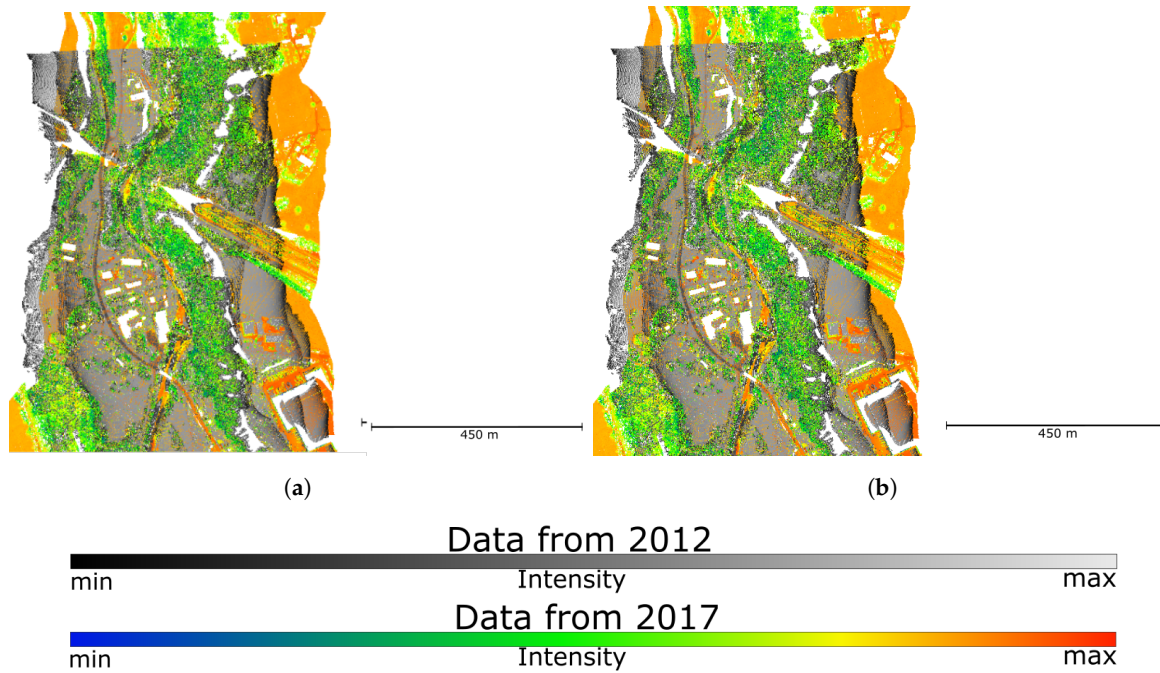


Figure 10. Comparison of the registration results using ICP and the proposed approach shown on the second scene. (a) Result of ICP. (b) Result of the proposed method. Shown data are the date of the whole flight strip downsampled to 1 m GSD.

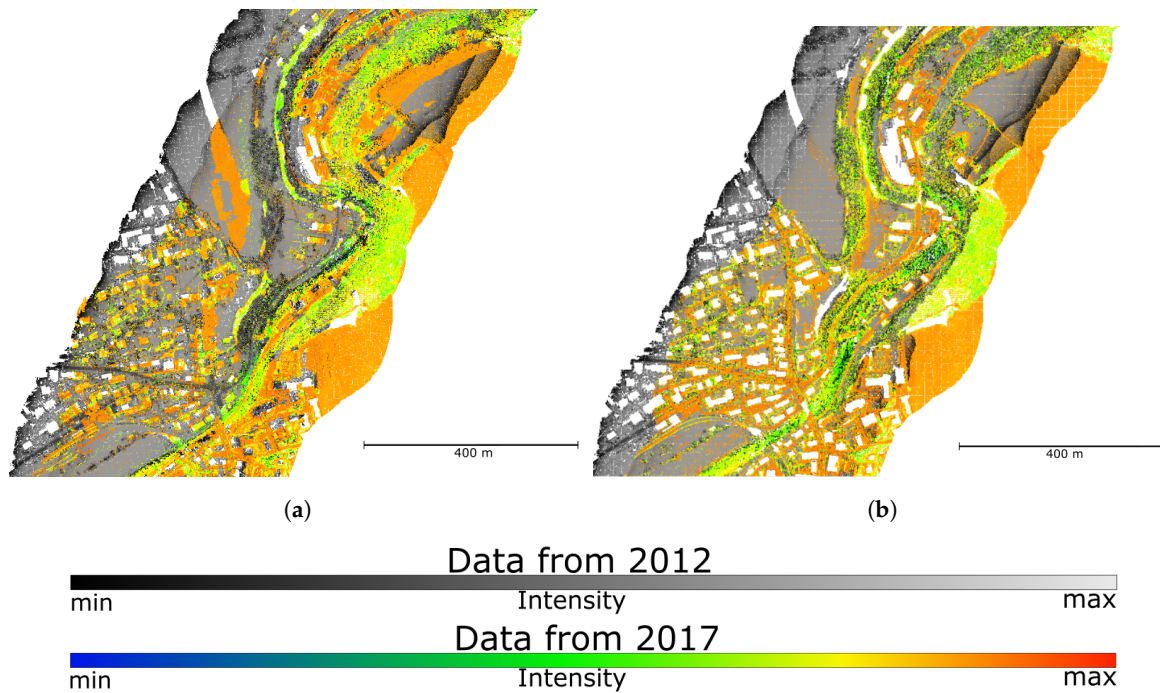


Figure 11. Comparison of the registration results using ICP and the proposed approach shown on the fourth scene. (a) Result of ICP. (b) Result of the proposed method. Shown data are the date of the whole flight strip downsampled to 1 m GSD.

Table 1. Results of the scenes with Point Cloud GSD: 0.5m and grid size: 1 m .

Scene	x		y		z	
	Average [m]	RMSE [m]	Average [m]	RMSE [m]	Average [m]	RMSE [m]
0.5 m						
1	0.11	0.28	−0.10	0.19	−0.324	0.012
2	0.13	0.33	0.37	0.35	−0.269	0.060
3	0.39	0.64	0.19	0.58	−0.315	0.033
4	−0.11	0.40	0.20	0.084	−0.137	0.007
5	0.22	0.44	0.23	0.17	−0.271	0.024
w	−1.268	0.77	0.48	0.04	0.094	0.022
1.0 m						
1	0.009	0.14	−0.11	0.11	−0.282	0.009
2	−0.11	0.26	0.65	0.12	−0.22	0.080
3	0.019	0.53	0.05	0.40	−0.252	0.033
4	−0.18	0.41	0.12	0.039	−0.056	0.15
5	0.002	0.20	0.13	0.12	−0.270	0.013
w	−1.320	0.02	0.026	0.0018	0.778	0.001
2.0 m						
1	−0.027	0.201	−0.101	0.046	−0.267	0.011
2	−0.030	0.379	0.622	0.225	−0.247	0.062
3	0.108	0.476	−0.005	0.333	−0.248	0.054
4	−0.310	0.318	0.153	0.054	−0.084	0.117
5	0.068	0.203	0.156	0.101	−0.274	0.013
w	−1.239	0.011	−0.198	0.002	1.195	0.001

Table 2. Results of the scenes with Point Cloud GSD: 0.5m and grid size: 1 m .

Scene	rx		ry		rz	
	Average [deg]	RMSE [deg]	Average [deg]	RMSE [deg]	Average [deg]	RMSE [deg]
0.5 m						
1	4.0×10^{-4}	5.5×10^{-4}	-1.1×10^{-3}	2.3×10^{-3}	1.7×10^{-2}	5.6×10^{-3}
2	1.3×10^{-2}	1.9×10^{-3}	-1.1×10^{-2}	6.9×10^{-3}	-1.6×10^{-2}	1.5×10^{-2}
3	-3.9×10^{-3}	2.2×10^{-3}	-8.1×10^{-3}	3.1×10^{-3}	-4.4×10^{-3}	1.1×10^{-2}
4	-1.1×10^{-2}	2.2×10^{-3}	-4.0×10^{-3}	2.4×10^{-3}	-2.4×10^{-2}	4.6×10^{-3}
5	-2.2×10^{-5}	7.1×10^{-4}	-7.1×10^{-4}	2.5×10^{-3}	-2.5×10^{-3}	9.4×10^{-3}
1.0 m						
1	2.0×10^{-3}	6.5×10^{-4}	2.2×10^{-3}	1.2×10^{-3}	1.6×10^{-2}	2.6×10^{-3}
2	1.3×10^{-2}	1.7×10^{-3}	-8.4×10^{-3}	9.7×10^{-3}	-2.8×10^{-2}	6.2×10^{-3}
3	-2.6×10^{-3}	2.2×10^{-3}	-5.8×10^{-3}	3.1×10^{-3}	-3.6×10^{-3}	1.7×10^{-2}
4	-1.2×10^{-2}	4.45×10^{-3}	-1.9×10^{-3}	3.0×10^{-3}	-2.1×10^{-2}	5.7×10^{-3}
5	-7.6×10^{-5}	1.7×10^{-4}	-9.2×10^{-4}	1.6×10^{-3}	4.6×10^{-4}	8.1×10^{-3}
w	1.1×10^{-3}	9.2×10^{-6}	1.1×10^{-2}	1.5×10^{-5}	5.7×10^{-4}	2.1×10^{-5}
2.0 m						
1	3.1×10^{-3}	5.7×10^{-4}	2.4×10^{-3}	1.8×10^{-3}	1.4×10^{-2}	3.9×10^{-3}
2	1.4×10^{-2}	2.0×10^{-3}	-1.0×10^{-2}	6.9×10^{-3}	-2.7×10^{-2}	8.2×10^{-3}
3	-2.1×10^{-3}	2.5×10^{-3}	-4.6×10^{-3}	4.4×10^{-3}	8.2×10^{-4}	5.2×10^{-3}
4	-1.2×10^{-2}	3.4×10^{-3}	-3.4×10^{-3}	2.9×10^{-3}	-2.3×10^{-2}	4.1×10^{-3}
5	3.0×10^{-5}	2.4×10^{-4}	-1.4×10^{-3}	1.6×10^{-3}	3.4×10^{-4}	6.7×10^{-3}
w	3.6×10^{-4}	3.5×10^{-6}	1.1×10^{-2}	5.5×10^{-6}	1.5×10^{-3}	1.8×10^{-5}

The averages, as well as the RMSE shown in the tables, have the same dimensions, which indicates the influence of different point densities to be low. Therefore, the proposed approach can

solve registration problems where the point densities between the data sets differs. Furthermore, a higher difference in the point densities shows a vanishing influence on the registration accuracy. Figure 12 shows the comparison of taking the same data and taking the datasets with 0.5 m sampling.



Figure 12. Diagrams of the correctness of calculated transformation parameters. Panels (a) and (b) show the transformation evaluated on the same data with translation parameters (a) and rotation parameters (b). Panels (c) and (d) show the transformation evaluated on the data corresponding to a GSD of 0.5 m with translation parameters (c) and rotation parameters (d). It should be named that there is a change of the grid cell size between the single scenes and the whole data. Single scenes are evaluated using a cell size of about 1 m and the whole data using a cell size of about 4 m.

The graphs have only a few differences. Possible reasons are different geometries in the case of multi-temporal data. While usage of the same point cloud shows the internal accuracy and correctness of the registration, the usage of the captured data shows the suitability of the proposed approach when dealing with multi-temporal data. The difference of translation parameters between different scenes is on dm level. By considering a grid size of 1 m (or 4 m) and a point distance of 0.5 m (or 1 m) the influence of different landscape forms is below the point or grid sampling.

The influence of changing the sampling distance is shown for selected scenes in the Figure 13. There are different GSD used for the target point cloud as well as a different grid size for the source grid.

As already seen in Tables 1 and 2, different ground-sampling distances have a vanishing influence. The main accuracy limitations of the proposed approach results from the chosen grid size of the source model. By changing the grid size to 4 m (or to 10 m) there is a difference in the transformation parameters of about one order of magnitude. Since the grid size also determines the smoothness of the resulting grid, this result is not surprising. A coarser grid size results in a smoother surface and thus eliminates more details of the landscape which is used for registration.

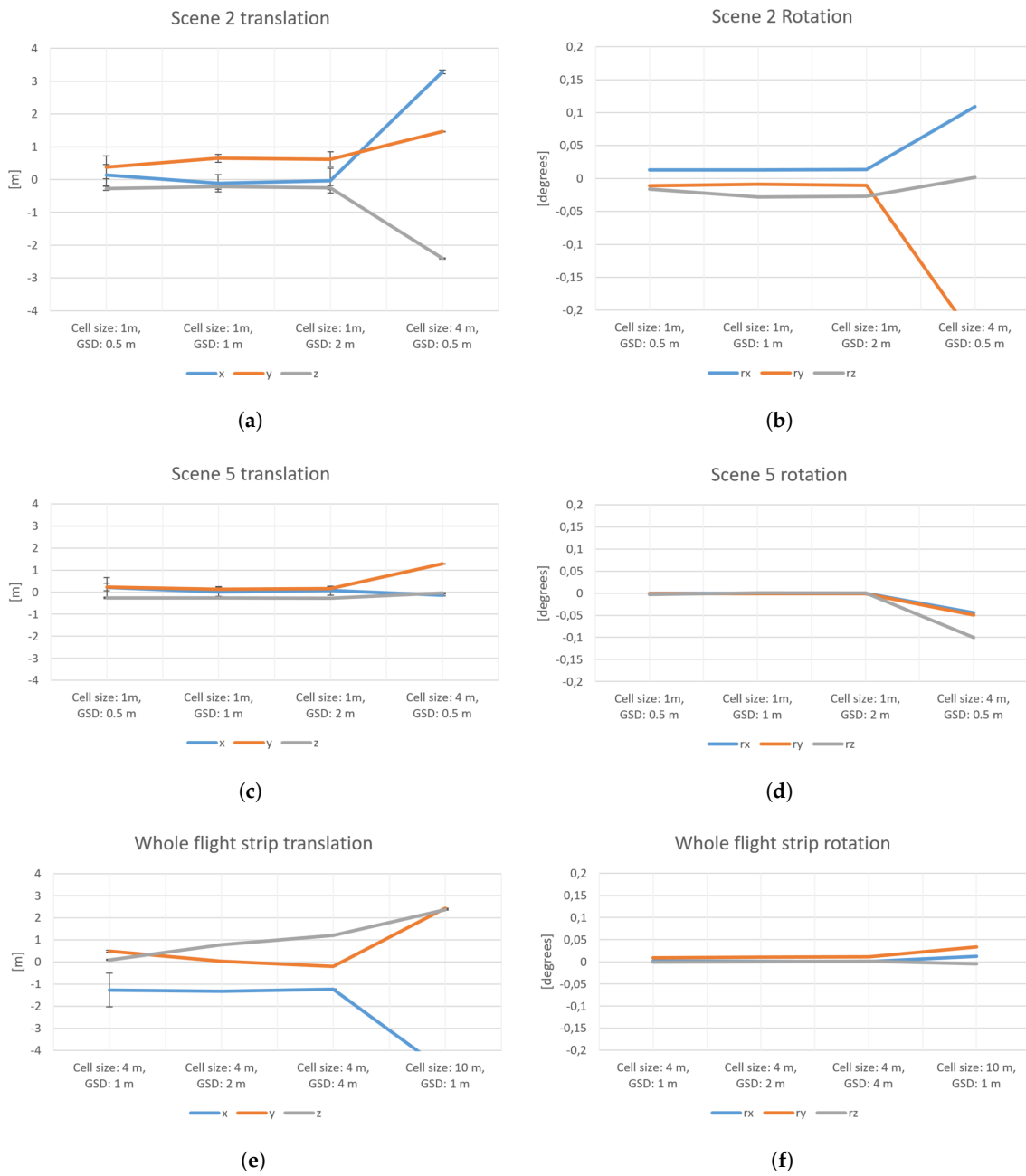
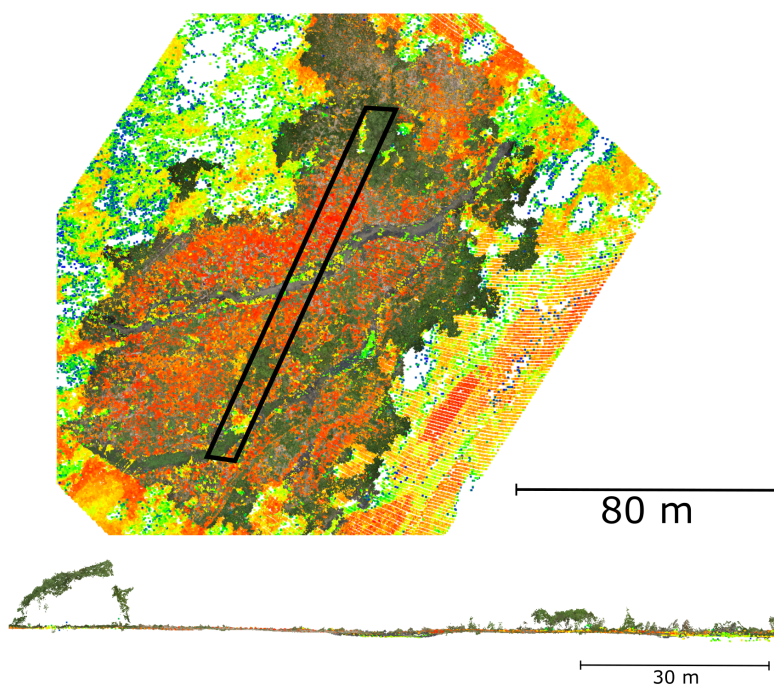


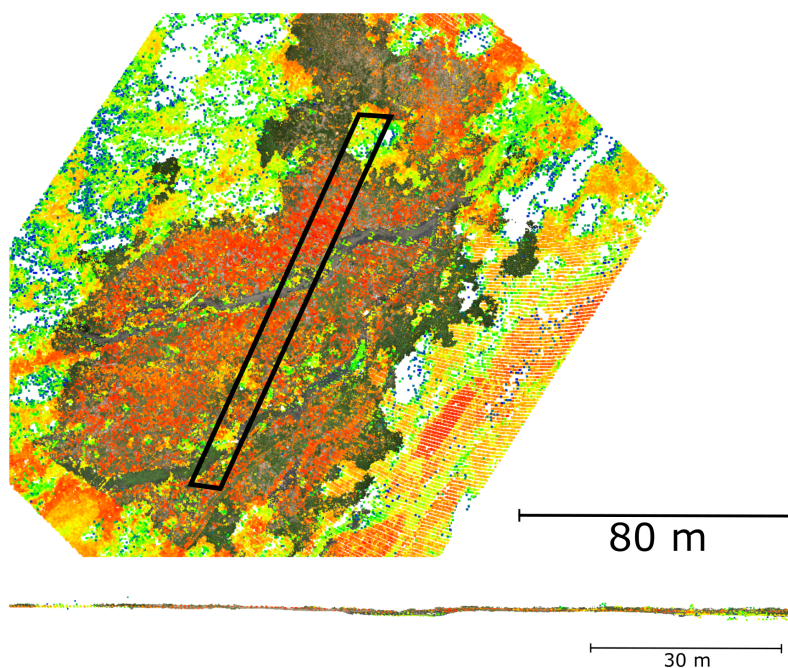
Figure 13. Diagrams of the correctness of evaluated parameters with the influence of different point distances. Shown are the scenes 2 (a,b), 5 (c,d) and for the whole flight strip (e,f) with the translation parameters on the left side and rotation parameters on the right side.

The registration result of the proposed LiDAR data and UAV data is shown in Figure 14. This figure also shows the influence of the proposed outlier removal in the marked areas.

The proposed approach can solve complex registration tasks as in the case of multi-sensor and/or multi-temporal registration. The marked areas indicate that the outlier removal successfully filters noise and furthermore works as binary classification.



(a)



(b)

Figure 14. Registration of the multi-sensor data and influence of the outlier removal. Shown are a photogrammetric point cloud, generated from UAV taken images and the ground points of the LiDAR point cloud of the flight campaign of 2017. (a) Registration result of the proposed method shown on the unfiltered data as top down view. (b) Registration shown on the filtered data as top down view. The marked areas locate the profile views which show the outlier removal in where successful vegetation points inside the photogrammetric data are filtered.

6. Discussion

6.1. Discussion of Results and Comparison with Other Work

The main advantage of the proposed method is the independence of different point densities and the level of noise between the data. In comparison to the classical ICP [1], it is demonstrated that the proposed method can solve the registration where the ICP gets into trouble. Other registration methods that aim for independency of different point densities use geometric primitives such as planes [16–19] or lines [14,15].

ICP variants gain the highest accuracies but are also limited to the GSD of the input point cloud. Fine-registration approaches using geometric primitives can benefit from the mathematical representation of the primitive, which reduces noise and overcomes the dependence on the point density. However, the primitive approaches are limited to the accuracy of the generalized primitive and the appearance of this primitives in the scene. The limitation on the accuracy of the calculated DEM as a primitive also exists in the proposed approach but it is shown that the approach can be used in urban scenes as well as in rural scenes and does not assume the appearance of planes or lines.

It is also shown that the proposed method can be used for multi-sensor registration, especially for registration of LiDAR and photogrammetric point clouds.

Evaluations show that the proposed approach reaches accuracies at a level below the chosen cell size and below the GSD target data. In concrete numbers, with a GSD of about 5 dm and a cell size of about 10 dm, a translation accuracy of several decimeters can be reached. The calculation of rotation parameters is less than 0.05 degrees. Since the proposed approach reaches accuracies below the GSD of the input point cloud, it is declared as a fine registration which is comparable with other fine-registration approaches [1,4,5,19] and can be used to optimize the results of coarse registration approaches such as [21].

6.2. Limitation and Future Direction

The primary influence of the final registration accuracy depends on the chosen cell size and neighborhood for smoothing of the grid. Evaluations show that the accuracy of the translation parameters are mainly influenced by the chosen DEM resolution. Choosing a coarser resolution of the DEM results in a higher smoothing of the landscape and will capture less detail. This lower amount of detail influences the accuracy of the translation in a negative way. The accuracy of the rotation parameters is also influenced by the chosen cell size but, compared to the translation, the cell size shows a vanishing influence on the rotation parameters. Therefore, the accuracy of the proposed approach is mainly limited to the used DEM as reference. Another limitation is caused by the assumption of temporal stable ground which limits the use case of the proposed method to registration of large-scale data without large change (e.g., landslides).

The multi-sensor aspect should be further addressed in the future. In particular, the possibility to register LiDAR and photogrammetric point clouds based on NIR images is very promising and could be a topic for further investigations. The primary usage of the proposed method in a future application will be the registration of airborne and UAV-based data to enhance airborne point clouds with higher spatial resolution and higher temporal resolution in the context of a change detection. Further future applications aim for change detection of river ground points. Such change detection needs statistical information about the transformed points to distinguish between real changes and measurement or registration errors. The outlier removal based on the histogram threshold will be improved in the future considering a confidence interval based on the standard deviation of the used sensor.

7. Conclusions

This paper proposed an approach to solve the registration of two different point clouds. This approach is marker-free and uses the landscape of the captured scene for registration by focusing on segmented ground points. To calculate the transformation parameters, the proposed approach uses

a minimization of point-to-grid distances with a regular grid of a DEM based on the segmented ground points. The DEM is generalized with a local optimization which offers the possibility to calculate large-scale DEMs. This possibility to use the proposed approach in large-scale scenarios is shown by the registration of the whole flight strip of two flight campaigns, which captures an area of eight square kilometers. To evaluate the proposed approach, this paper uses five different scenes, which differ in the appearance of manmade structures and the form of the landscape and fulfil all degrees of freedom of the transformation parameters. The multi-sensor aspect is shown using LiDAR and photogrammetric points clouds as well as using LiDAR point clouds captured with different scan systems.

A dynamic threshold is used for outlier removal, which makes the registration more robust and performs a ground segmentation during the registration process. This outlier removal also filters slightly different geometry caused by morphological changes. Furthermore, the proposed method works if most of the ground can be seen as temporally stable in the case of multi-temporal data.

The evaluation shows that the proposed approach reaches an accuracy and correctness of translation parameters below the GSD of the target point cloud and below the chosen cell size of the DEM. The advantage of the proposed method is the considering of statistic information along the whole pipeline starting at the raster generation and ending at the statistical model of the final registration. The statistical model for registration uses the information about the reconstruction accuracy and therefore can directly be used to mask areas which should not be considered, or be considered with low weight for registration (e.g., water areas) by setting the cell reconstruction variance to a high value at these areas.

Since the DEM can be calculated with different interpolation methods, the sort of interpolation that creates the best results considering different scenarios, or if different scenarios need different interpolation methods, should be evaluated in the future. Other future work could evaluate more multi-sensor data sets. A registration of NIR images captured from an UAV flight and data captured from an airplane is very promising. Water areas have a high absorption rate in the near-infrared (NIR) channel and can therefore be easily marked in NIR images. Therefore, the recognition of water areas could be enhanced by a registration of NIR images and the reference data set. Future work in the direction of change detection of the river ground will benefit from a data set of co-registered NIR data and airborne LiDAR data.

Author Contributions: Richard Boerner: methodology, writing, validation, software; Yusheng Xu and Ludwig Hoegner and Uwe Stilla: writing—review and editing; Ramona Baran and Frank Steinbacher: Data acquisition; Ludwig Hoegner and Uwe Stilla: supervision, funding acquisition.

Funding: This research was funded by “Bayerische Forschungsstiftung”, project “Schritthaltende 3D-Rekonstruktion und -Analyse (AZ-1184-15)”, subproject “Änderungsdetektion in Punktwolken”. This work was supported by the German Research Foundation (DFG) and the Technical University of Munich within the funding program Open Access Publishing.

Acknowledgments: The authors would like to thank our project partner Steinbacher Consult for providing the ALS dataset.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

MDPI	Multidisciplinary Digital Publishing Institute
GSD	Ground-Sampling Distance
DEM	Digital Elevation Model

References

1. Besl, P.J.; McKay, N.D. Method for registration of 3-D shapes. In *Robotics-DL Tentative*; International Society for Optics and Photonics: Bellingham, WA, USA, 1992; pp. 586–606.
2. Boerner, R.; Xu, Y.; Hoegner, L.; Baran, R.; Steinbacher, F.; Stilla, U. DEM based registration of multi-sensor airborne point clouds exemplary shown on a river side in non urban area. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2018**, *XLII-2*, 109–116. [[CrossRef](#)]
3. Boerner, R.; Xu, Y.; Hoegner, L.; Stilla, U. Registration of UAV DATA and ALS data using point to DEM distances for bathymetric change detection. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2018**, *42*, 51–58. [[CrossRef](#)]
4. Habib, A.; Datchev, I.; Bang, K. A comparative analysis of two approaches for multiple-surface registration of irregular point clouds. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2010**, *38*, 6.
5. Al-Durgham, M.; Habib, A. A framework for the registration and segmentation of heterogeneous lidar data. *Photogramm. Eng. Remote Sens.* **2013**, *79*, 135–145. [[CrossRef](#)]
6. Böhm, J.; Becker, S. Automatic marker-free registration of terrestrial laser scans using reflectance. In Proceedings of the 8th Conference on Optical 3D Measurement Techniques, Zurich, Switzerland, 9–12 July 2007; pp. 338–344.
7. Weinmann, M.; Weinmann, M.; Hinz, S.; Jutzi, B. Fast and automatic image-based registration of TLS data. *ISPRS J. Photogramm. Remote Sens.* **2011**, *66*, S62–S70. [[CrossRef](#)]
8. Theiler, P.W.; Wegner, J.D.; Schindler, K. Keypoint-based 4-Points Congruent Sets-Automated marker-less registration of laser scans. *ISPRS J. Photogramm. Remote Sens.* **2014**, *96*, 149–163. [[CrossRef](#)]
9. Weber, T.; Hänsch, R.; Hellwich, O. Automatic registration of unordered point clouds acquired by Kinect sensors using an overlap heuristic. *ISPRS J. Photogramm. Remote Sens.* **2015**, *102*, 96–109. [[CrossRef](#)]
10. Theiler, P.; Schindler, K. Automatic registration of terrestrial laser scanner point clouds using natural planar surfaces. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2012**, *3*, 173–178. [[CrossRef](#)]
11. Yang, B.; Dong, Z.; Liang, F.; Liu, Y. Automatic registration of large-scale urban scene point clouds based on semantic feature points. *ISPRS J. Photogramm. Remote Sens.* **2016**, *113*, 43–58. [[CrossRef](#)]
12. Ge, X. Automatic markerless registration of point clouds with semantic-keypoint-based 4-points congruent sets. *ISPRS J. Photogramm. Remote Sens.* **2017**, *130*, 344–357. [[CrossRef](#)]
13. Yang, B.; Zang, Y. Automated registration of dense terrestrial laser-scanning point clouds using curves. *ISPRS J. Photogramm. Remote Sens.* **2014**, *95*, 109–121. [[CrossRef](#)]
14. Habib, A.; Ghanma, M.; Morgan, M.; Al-Ruzouq, R. Photogrammetric and LiDAR data registration using linear features. *Photogramm. Eng. Remote Sens.* **2005**, *71*, 699–707. [[CrossRef](#)]
15. Stamos, I.; Leordeanu, M. Automated feature-based range registration of urban scenes of large scale. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Madison, WI, USA, 18–20 June 2003; Volume 2.
16. Dold, C.; Brenner, C. Registration of terrestrial laser scanning data using planar patches and image data. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2006**, *36*, 78–83.
17. Von Hansen, W. Robust automatic marker-free registration of terrestrial scan data. *Proc. Photogramm. Comput. Vis.* **2006**, *36*, 105–110.
18. Xiao, J.; Adler, B.; Zhang, H. 3D point cloud registration based on planar surfaces. In Proceedings of the 2012 IEEE Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI), Hamburg, Germany, 13–15 September 2012; pp. 40–45.
19. Hebel, M.; Arens, M.; Stilla, U. Change detection in urban areas by object-based analysis and on-the-fly comparison of multi-view ALS data. *ISPRS J. Photogramm. Remote Sens.* **2013**, *86*, 52–64. [[CrossRef](#)]
20. Wang, J.; Lindenbergh, R.; Shen, Y.; Menenti, M. Coarse point cloud registration by egi matching of voxel clusters. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2016**, *III*, 97–103. [[CrossRef](#)]
21. Xu, Y.; Boerner, R.; Yao, W.; Hoegner, L.; Stilla, U. Automated coarse registration of point clouds in 3d urban scenes using voxel based plane constraint. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2017**, *IV-2/W4*, 185–191. [[CrossRef](#)]
22. Boerner, R.; Hoegner, L.; Stilla, U. Voxel based segmentation of large airborne topobathymetric lidar data. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2017**, *XLII-1/W1*, 107–114. [[CrossRef](#)]

23. Xu, Y.; Tuttas, S.; Hoegner, L.; Stilla, U. Voxel-based segmentation of 3D point clouds from construction sites using a probabilistic connectivity model. *Pattern Recognit. Lett.* **2018**, *102*, 67–74. [[CrossRef](#)]
24. Shirowzhan, S.; Sepasgozar, S.M.; Li, H.; Trinder, J. Spatial compactness metrics and Constrained Voxel Automata development for analyzing 3D densification and applying to point clouds: A synthetic review. *Autom. Constr.* **2018**, *96*, 236–249. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).