

## Image-based method for the pairwise registration of mobile laser scanning point clouds

Christodoulou, A.; Van Oosterom, P.

**DOI**

[10.5194/isprs-archives-XLII-4-93-2018](https://doi.org/10.5194/isprs-archives-XLII-4-93-2018)

**Publication date**

2018

**Document Version**

Final published version

**Published in**

International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences

**Citation (APA)**

Christodoulou, A., & Van Oosterom, P. (2018). Image-based method for the pairwise registration of mobile laser scanning point clouds. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 42(4), 161-168. <https://doi.org/10.5194/isprs-archives-XLII-4-93-2018>

**Important note**

To cite this publication, please use the final published version (if applicable).  
Please check the document version above.

**Copyright**

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

**Takedown policy**

Please contact us and provide details if you believe this document breaches copyrights.  
We will remove access to the work immediately and investigate your claim.

# IMAGE-BASED METHOD FOR THE PAIRWISE REGISTRATION OF MOBILE LASER SCANNING POINT CLOUDS

A. Christodoulou<sup>1</sup>, P. van Oosterom<sup>2</sup>

<sup>1</sup> TU Delft, MSc Geomatics, Delft, The Netherlands - A.Christodoulou-1@student.tudelft.nl

<sup>2</sup> TU Delft, GIS technology, Delft, The Netherlands - P.J.M.vanOosterom@tudelft.nl

Commission IV, WG IV/5

**KEY WORDS:** pairwise registration, relative registration, image registration, template matching, mobile scanning, point clouds, reliability, sub-pixel accuracy

## ABSTRACT:

In this paper, a method is proposed for solving relative translations of 3D point clouds collected by Mobile Laser Scanning (MLS) techniques. The proposed approach uses the attributes of the 3D points to generate and match 2D-projections, by employing a simple correlation technique instead of matching in 3D. As a result, the developed method depends more on the number of pixels in the 2D-projections and less on the number of points in the point clouds. This leads to a more cost-efficient method in contrast to 3D registration techniques. The method uses this benefit to provide redundant translation parameters for each point cloud pair. With the utilization of image-based evaluation criteria the reliable translation parameters are detected and only those are used to compute the final solution. Consequently, the confidence levels of each final estimation can be computed. In addition, an indication of robustness showing how many estimations were included for the computation of the final solution is included. It is shown that the method performs fast due to its simplicity especially when medium image resolution's such as 0.15m are used. Reliable matches can be produced even when the overlap of the point cloud sets is small or the initial offset large as long as the offsets are distinguishable in the projections. Furthermore, a technique is proposed to obtain capabilities for sub-pixel accuracy estimations, as the accuracy of the estimations is restricted to the grid cell size. The technique seems promising, but further improvement is necessary.

## 1. INTRODUCTION

### 1.1 Problem statement

Point cloud data is an important source of 3D spatial information, as vast amounts of highly dense 3D points can be collected with laser scanning techniques in a considerably short amount of time. Mobile Laser Scanning (MLS) techniques are used for the rapid and cost-effective recording of street view data. During a MLS process it is usually necessary to record a certain scene more than once to retrieve a complete representation of it, for instance at a road junction. In such cases, point clouds representing (part of) the same scene but retrieved at different epochs do not perfectly match. In contrast they tend to have offsets in the X, Y, Z coordinates (Figure 1).

This deviation between the point clouds is caused due to environment-dependending limitations of the Global Navigation Satellite Systems (GNSS) signals, which constitute the main source of positioning information. Namely, when signal multipathing or blockage occurs the navigation solution has poor quality or can even be unavailable. The positioning in such cases depends on an Inertial Measurement Unit (IMU), which calculates positions based on displacements from an initial known position (Levi and Judd, 1996). However this leads to the accumulation of potential encountered positioning errors.

The use of point cloud data retrieved from MLS techniques requires the integration of all the 3D scans collected at different times from different observation points in a common reference system, a process known as global registration (Sanchez et al., 2017). However, prior to that it is necessary to perform relative alignment of the overlapping point clouds to facilitate the

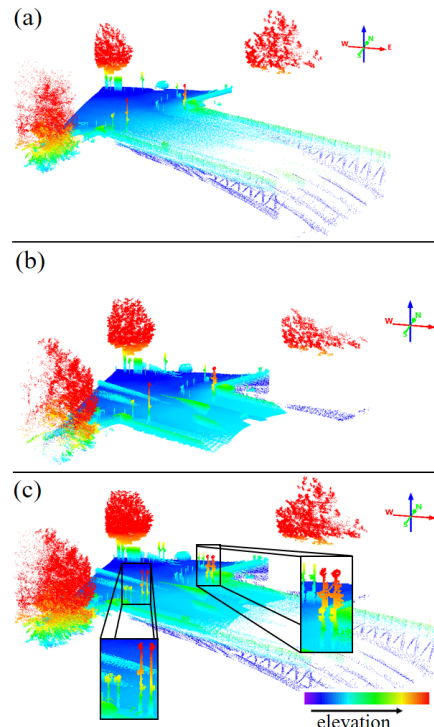


Figure 1. (a) and (b): Two point clouds representing part of the same scene but captured at different times are visualized from a diagonal point-of-view. (c): An overlay of (a) and (b), where the magnified parts highlight the offsets between (a) and (b).

construction of a single 3D point cloud model. Relative registration also known as local alignment or matching, refers to the estimation of the transformation parameters that are needed to match one point-cloud with another (Magnusson et al., 2007). Locally and ultimately globally registered point clouds can be used for several purposes such as object recognition, cultural heritage modeling, disaster management or even as alternatives to surveying processes such as coordinates extraction of cadastral parcels.

## 1.2 Paper objectives

The main focus of this work is to assess the extend to which it is attainable to solve the 3D relative registration challenge by creating projections from the point clouds and aligning in 2D. Only translation parameters are estimated, neither rotation nor scaling are considered as translation errors may be quite large in case of limited GNSS visibility, in comparison with rotation and scaling errors. Particularly, there might be rotation errors in Z axis of the mobile platform around the  $\frac{1}{10}$  of a degree. This happens when the driver of the recording vehicle takes successively the same direction in turns for some period of time. With regard to scaling, when performing MLS for a long time without strong GNSS reception in theory, the scale factor of the point clouds will not be exactly equal to 1.

A method is proposed with which the quality of an alignment between a pair of scans could be automatically assessed. In addition an approach is proposed to gain translations of sub-pixel accuracy as the matching is performed between imagery. The quality evaluation method is independent of highly accurate ground truth data such as reference points that could be used for comparison. Enormous amounts of reference points would be required for large scale projects so as to judge the quality of relative registrations.

In summary, the proposed image-based method for the registration of mobile laser scanning point clouds consists of several important aspects:

- Creation of 2D projections so that the 3D point clouds are best described.
- Computation of the transformation parameters that align relatively 3D overlapping point cloud pairs by matching 2D imagery.
- Automated determination of the final results' quality.
- Development of a sub-pixel accuracy technique to improve the drawback of discrete grid cell size.

## 1.3 Motivation

Some of the advantages of reducing the dimensionality of the problem, relate to the limitations faced by the Iterative Closest Point (ICP), which is commonly used for pairwise registration of point clouds in 3D. It requires computationally expensive and extensive search of point correspondences between the point clouds (Godin et al., 1994). Also, ICP-based methods perform poorly when points in one scan do not have correspondences in the other (Pomerleau et al., 2013), which is very common in MLS data (Figure 1a and 1b). Furthermore, this algorithm delivers incorrect results if the initial position of the point clouds is not close to the required matching position, or simply put the offset is large (Shetty, 2017), (Sanchez et al., 2017)). This can be the case in

MLS data when streets separate dense blocks of high structures or when tall trees are present, as then the GNSS reception is highly restricted. Another advantage of reducing to 2D but not related to ICP, is that the computation time when employing 2D registrations processes can be decreased. The reason is that the method becomes then less dependant on the (large) number of points, and more on the number of grid cells.

## 2. RELATED WORK

### 2.1 3D local point cloud registration

3D relative registration techniques may be preferred as they do not require compressing data into discrete grid cells. Many researches employ a variant of the Iterative Closest Point (ICP) (Besl and McKay, 1992) algorithm, the most commonly used algorithm for the local registration of 3D overlapping scans in 3D (Byun et al., 2017). The algorithm establishes iteratively correspondences between the points of two point cloud sets and computes the spatial distances between them. It terminates when the sum of the spatial distances between the correspondences is minimum (Sanchez et al., 2017).

The Iterative Closest Compatible Point (ICCP) algorithm (Godin et al., 1994), reduces the search space of points correspondences by finding firstly those that are compatible according to their intensity value and then the one that lays at a minimum distance. Despite this, the compatible points are recomputed at each iteration which is a costly operation, and ICCP, like ICP, estimates correct results when most of the points in the one set have a correspondence in the other. Additionally, even when the intensity value of two points is compatible and their distance is minimum, it is not guaranteed that the correspondence is correct. If the offset between the 2 scans is large, then other compatible points may be closer. The trimmed-ICP algorithm (Chetverikov et al., 2002) sorts the square distances between the points and minimizes their sum by iteratively excluding a number of extreme values. Thus, it can eliminate faulty estimated correspondences. However, it requires knowledge regarding the overlap of the scans, so as to know how many point correspondences to trim. Iterative Closest Point using Invariant Features (Sharp et al., 2002) improves the selection of point correspondences by extracting quantities of the point clouds that are invariant to rotations and translations, such as the points' curvature. The correspondences are computed between points of the detected features like in an ICP approach. The results are sufficient for coarse registration purposes, which means that the scans do not completely match, but shift closer to the matching position (Sharp et al., 2002).

### 2.2 2D local point cloud registration

A 2D point cloud registration technique proposes the creation of bearing angle images from 3D point clouds (Lin et al., 2017). This type of images is used as it can highlight the depth discontinuities and direction changes of a scene (Scaramuzza et al., 2007), useful properties for 2D matching. A 2D feature-based matching method is used to find corresponding pixels between an image pair. As a result, for each pixel correspondence it is possible to retrieve an equivalent set of corresponding 3D points. These are used in a least squares approximation to derive the transformation parameters. Due to the 2D matching, the computation cost is significantly less than ICP. However, it is shown that the precision of the method is not better than that of generalized-ICP, which is a plane to plane ICP registration (Segal et al., 2009), because sometimes outlying correspondences are included.

### 2.3 Quality of local registration

The quality of relatively registered data was evaluated in a research by assessing whether the overlapping region after the alignment represents the same physical surface (Huber and Hebert, 2003). Two measures of surface consistency were used. The Euclidean distance between corresponding points of two overlapping point clouds and the angle between the normal vectors of corresponding points. A surface is considered consistent when these two measures are less than a threshold. In another work, similarly to the Euclidean distance, a mean-square error (MSE) was computed after correspondences were found (King et al., 2005). The matching solution is accepted as soon as the MSE is close to the approximate noise of the sensor. In addition, a 'non-randomness score' is used to detect the random alignments or mis-alignments due to repetitive structures, such as a row of windows in a building. This score results from the matching of distinctive structures detected on the two point clouds. If the transformation parameters are close to the first registration estimation, then this accounts as a proof that the estimated solution is non-random.

### 2.4 Sub-pixel accuracy

As mentioned, the relative transformation parameters between point clouds are estimated in this project by matching imagery. As a result, the parameters can have at maximum the accuracy of the images' grid cell size. Techniques to retrieve values of sub-pixel accuracy, or in other words, accuracy higher than that of the pixel, are explored. A second order polynomial is fitted on the values of some pixels of interest (Zhang et al., 2009). Figure 2 illustrates three pixels in a row  $P_1$ ,  $P$ ,  $P_2$  as point entities. It can be seen that the polynomial fitting operates like an interpolation method as values of sub-pixel accuracy can be retrieved. For instance, information is acquired regarding the highest value between the area that the three pixels cover. This is the value  $W_{max}$  of the point  $P_{max}$  and not the value  $W$  of the point  $P$ , as one would think by examining the discrete pixel values. In another work, the fitting of a 1D Gaussian function is discussed (Naidu and Fisher, 1991) and applied with the same concept as explained for the polynomial fitting.

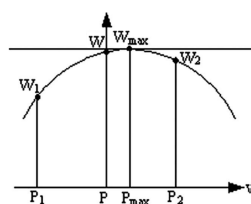


Figure 2. Interpolation of pixel values with a second order polynomial.  $P_1$ ,  $P$ ,  $P_2$  illustrate the locations of three pixels as points placed in the middle of the pixels.  $W_1$ ,  $W$  and  $W_2$  are the corresponding pixel values.

## 3. METHODOLOGY

### 3.1 Pre-processing

The proposed method assumes that the 3D mobile laser scanned data are stored in tiles, for instance of 50 meters in X axis by 50 meters in Y axis. Each tile of 3D points is simply mentioned in the rest of the paper as a 'point cloud set'. The procedure begins with the computation of the surfaces' normal vectors on each 3D

point of a point cloud set by using the Principal Component Analysis (PCA) method (Shlens, 2014). For the calculation of the normal vectors at each point, the neighbouring points are considered so as to compute the vector that is perpendicular to the surface fitted by the points' neighbourhood. In order too make the normal vectors direction consistent, the trajectory points of the moving vehicle are used as vantage points. These are used to orientate the normal vectors towards the laser scanner direction.

The next step is to pair the point cloud-sets. Despite the fact that the positioning of the 3D data could be degraded due to the lack of GNSS reception, it is still possible to detect the overlapping point cloud sets from their coordinates by assuming a buffer zone. As soon as the overlapping point clouds are known, all the possible and unique combinations of pairs of point clouds sets are formed.

### 3.2 From 3D to 2D

For every point cloud pair the combined minimum and maximum X, Y, and Z coordinate is computed. These are meant to be used as the boundaries for the two 2D projections generated from the two corresponding point cloud sets in a pair. In such a way, it is feasible to spot and compute the positioning offsets in the created imagery. Next, each point cloud set is projected on three planes resembling a top and two orthogonal views. Namely, we project the point clouds sets in the XY-plane (Figure 3), in the XZ-plane (Figure 4) and in the YZ-plane (Figure 5).

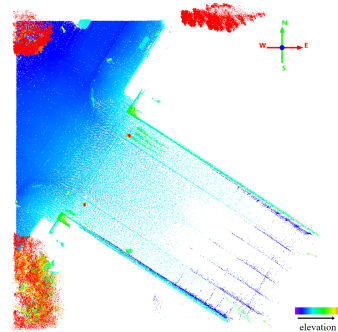


Figure 3. Projection of Figure 1a on the XY-plane.

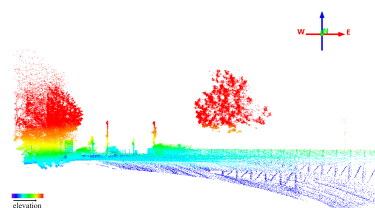


Figure 4. Projection of Figure 1a on the XZ-plane.

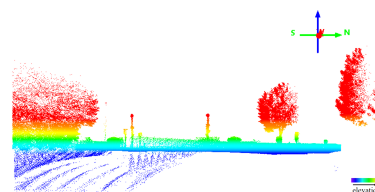


Figure 5. Projection of Figure 1a on the YZ-plane.

According to the plane we project, the relevant combined boundaries of the two point cloud sets are utilized (see Section 3.1). For



example, for the creation of the XY-plane we use the boundaries in X and Y. These, along with a specified grid cell size, are used to define the number of the grid cells and their edges. Subsequently, the 2D-coordinates of the points are used to spatially bin the points into 2D grid cells. Thereafter, attributes of the points are used as the information illustrated on the 2D-projections. We use points' characteristics that could best describe the 3D information in 2D. Specifically, these are the density of the points within the grid cells, the intensity, the depth, the gradient of the intensity, the gradient of the depth and the calculated normal vectors of the points. As a result, a set of image pairs is produced for every point cloud pair.

Each grid cell of a density image simply illustrates the total amount of points that fall in the 2D cell. The density is considered an important characteristic because long and thick features, such as walls or the ground, will be represented with high amounts of points (i.e. bright pixels). Thus these features can be highly distinguishable in a certain pair of projections. However, one object captured in two overlapping point cloud sets may be represented with less (or more) dense points, according to the recording vehicle's position. Nevertheless, even if the laser scanner captured a scene from a long distance, a wall will still be represented by more points than a pole. To minimize the effect of having different densities of points in the two point cloud sets, the density values in the projections are normalized up to the same value. Examples of density imagery in the three projections are given in Figure 6.

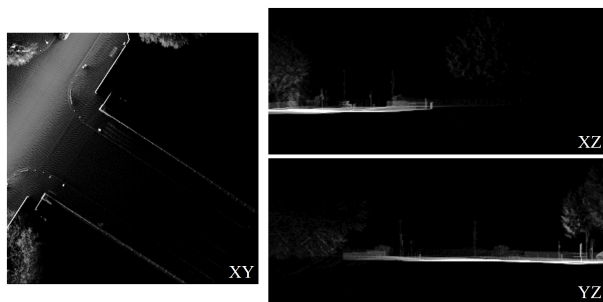


Figure 6. The three images produced from the point cloud set of Figure 1a illustrating density values.

The intensity of each 3D point, which is the strength of each received laser signal is also used to create a second type of images. By using the points' intensity, features like painted lines on the roads become visible, as it can be observed at the upper left corner of an XY-plane shown in Figure 7.

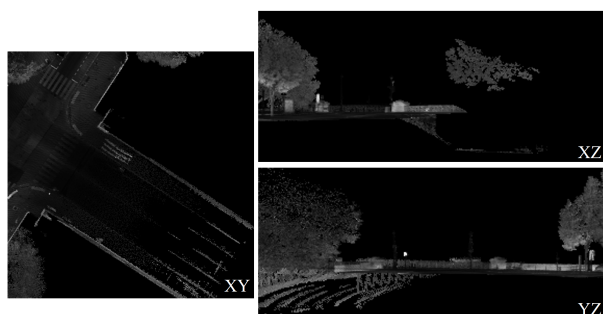


Figure 7. The three images produced from the point cloud set of Figure 1a illustrating intensity values.

Because a 2D grid cell can possibly include points of different intensities, the mean of the intensities is computed to output a single intensity value per grid cell.

Corresponding objects in a point cloud pair may be represented by points of different intensity, due to the different position of the recording vehicle. Thus, the intensity images are used for the construction of images where the pixels' values depict the gradient of the intensity. The gradient images can be useful for highlighting corners and edges. These will not differ in two projections resulted from point clouds sets recorded from different observation points (Figure 8).

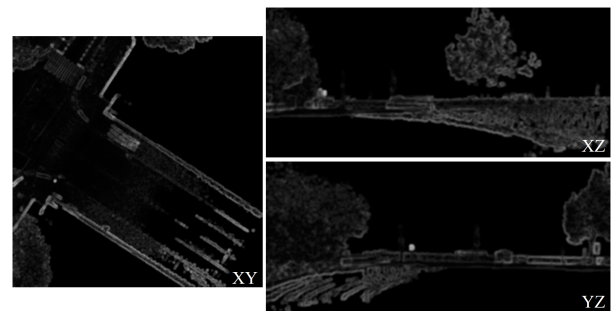


Figure 8. The three images produced from the point cloud set of Figure 1a illustrating gradient of intensity values.

Depth projections depict how far the objects of an image are with respect to each projected plane's view point. Particularly, the 2D grid cells contain the coordinates of the 3rd dimension. For example, the grid cells of an XY-plane illustrate the Z-values of the points. In that case, the cells which contain points with higher elevation will be brighter, and the cells which contain points with lower elevation will be darker. Similarly the depth values are assigned on the other two planes (Figure 9).

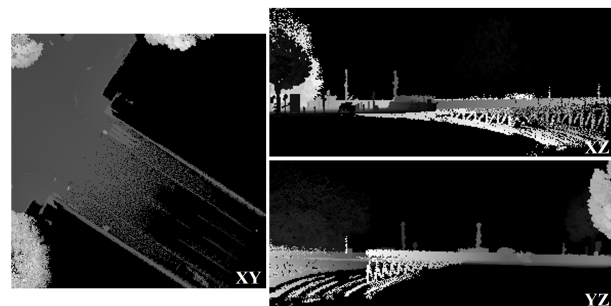


Figure 9. The three images produced from the point cloud set of Figure 1a illustrating depth values.

As the depth projections represent the value of a coordinate, corresponding objects in a point cloud pair may be represented by points of different depth due to the positioning error. For that reason, gradient of depth images are created (Figure 10).

Finally, the grid cells of the 2D projections are filled with the values of the normal vectors. These images contain values in 3 colour channels, one for each direction of the vector. The normal vectors are used to create imagery illustrating the orientation of the features' surfaces (Figure 11). Furthermore, 2D projections are constructed that illustrate the vector's value only in X, only in Y and only in Z direction. By doing this, the surfaces' orientation are strengthened. As example, Figure 12 shows the values of the normal vectors in Z direction.

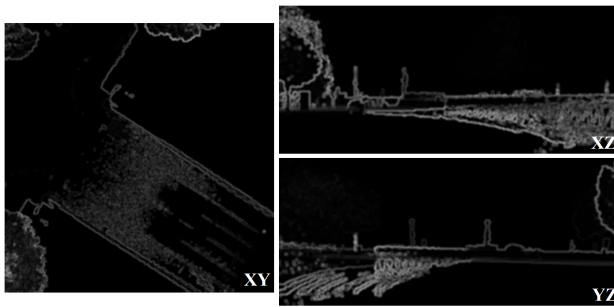


Figure 10. The three images produced from the point cloud set of Figure 1a illustrating gradient of depth values.

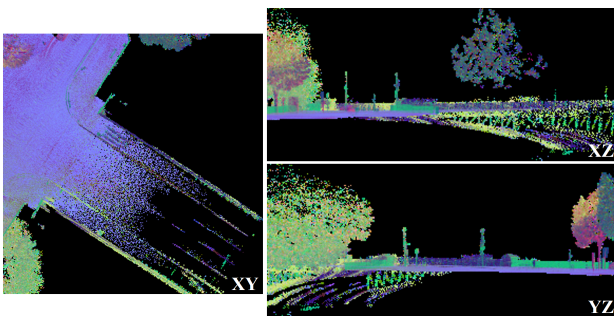


Figure 11. The three images produced from the point cloud set of Figure 1a illustrating normal vector values.

The images that show the normal vectors' at the Z direction illustrate brighter values when a normal vector points towards the Z direction. In contrast, they illustrate darker values when the surfaces are not perpendicular to the Z direction.

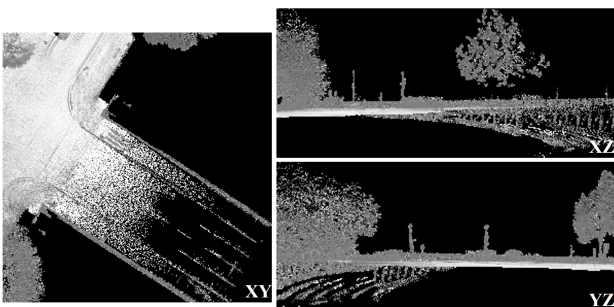


Figure 12. The three images produced from the point cloud set of Figure 1a illustrating the Z values of the normal vectors.

Nine types of images are used to represent each point cloud set. For each image type a set of an three projection is produced. Overall, 27 image pairs are constructed for each point cloud pair (Figure 13).

### 3.3 Image registration

An image registration technique is applied to match produced images that contain common visual information (Gaidhane et al., 2014) in order to retrieve the translation parameters. A template matching method, which determines the location of a template image (smaller sample) within a reference image (larger sample) is used (Sarvaiya et al., 2009). The implemented method is based

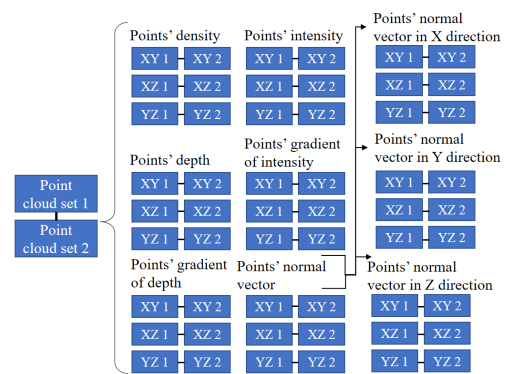


Figure 13. The 27 produced image pairs from each overlapping point cloud pair.

on a simple cross correlation statistical analysis of the brightness values of the two images. Particularly, the template image shifts over every possible location of the reference image, pixel by pixel. At every location, a degree of similarity between the two images is calculated with the cross correlation approach. The similarity degree equals to the sum of all the multiplications between the corresponding pixels of the template and the reference image (Ding et al., 2001). The best match is the location where the highest similarity value is found. Despite the fact that a cross correlation method is sensitive to intensity differences between the two images (Sarvaiya et al., 2009), it is used in this project because high correlations can still be obtained, as long as the values of the pixels follow the same pattern (Ding et al., 2001). Additionally, the method performs successfully even when small rotation and scaling is presented (Sarvaiya et al., 2009).

The template matching technique is implemented twice, the first time one image of a pair is considered to be the reference and next the other. There are two reasons for that. Firstly, for the current application there is no knowledge with regard to which point cloud set in a pair has correct absolute position, or even if any of them has. Second, the expectation is that the same translation parameters will be retrieved from both matchings once with negative and once with positive sign. As mentioned, the template matching technique requires that the reference image considers a larger sample. Thus a border of pixels with zero value is added to the reference image. The amount of extra pixels is determined according the maximum expected translation error.

The so-called score map, is the output of the image registration method. It is a 2D array that contains the similarity values computed between the template and reference image at all the possible overlay positions. Its size equals to  $W - w + 1, H - h + 1$  in pixels, where  $W$  and  $H$  the width and the height of the reference image, and  $w$  and  $h$  the width and the height of the template image, correspondingly. An example of the resulted score map's size according to some specific parameters is given in Table 1. A score map with these parameters is shown in Figure 14. As the matching of each image pair is applied twice with swapped roles between the reference and the template image, one of the two resulted score maps is flipped horizontally and vertically. Then it is superimposed onto the other score map to retrieve the final score map. The next step is the detection of the location of the maximum correlation value in the score map, shown with red in Figure 14. This is the location of the template image into the reference image, namely the 2D translation between the images.

| Image                | W (m) | H(m)  | W (pix) | H (pix) |
|----------------------|-------|-------|---------|---------|
| Image 1              | 50    | 50    | 1000    | 1000    |
| Border               | 10    | 10    | 200     | 200     |
| Image 1 as reference | 60    | 60    | 1200    | 1200    |
| Image 2              | 50    | 50    | 1000    | 1000    |
| Score map            | 10.05 | 10.05 | 201     | 201     |

Table 1. Parameters that define the size of a score map resulted from images' of 50m by 50m with grid cell size 0.05m and border of 5m in each side.

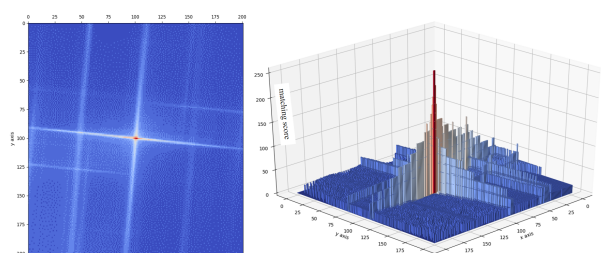


Figure 14. Left: An example of a score map image resulted from image registration (with the parameters described in Table 1). Right: For better understanding of the score map's concept the score map is also visualized in 3D.

### 3.4 Optimal Solution and Reliability Evaluation

For the quality evaluation of the relative transformations we adapt and use theoretical knowledge about the reliability and precision of spatial data from the 'Adjustment theory' mainly developed (Baarda, 1968) and applied (Sweco Nederland B.V., 2016) for the processing and quality control of survey data. Thorough relevant information will be given in a future thesis work.

To apply the theoretical knowledge we mainly use image processing techniques. For every image pair of a point cloud pair, three evaluation criteria are applied on the resulted score maps. The highest similarity values of the score maps computed with two different methods must be the same. Also, there should only be one highest similarity value and if there are more there must be a significant difference between the values of the highest two. Lastly the highest peak's value must be higher than a threshold. Subsequently, only the reliable score maps (which pass all the criteria) are used for the computation of the final translation parameters that match a point cloud pair. The optimum solution is the average of the accepted solutions. It is accompanied with its standard deviation and the amount of image pairs eventually used for the computation of the mean, as an indication of robustness.

### 3.5 Sub-pixel Accuracy

The accuracy of the results with the proposed method can not be better than the accuracy of the images' grid cell size. Thus an interpolation method is proposed with which translation parameters of sub-pixel accuracy can be retrieved. The method is applied on the score maps and particularly on the highest similarity pixel value and its neighboring pixels. We assume that if we would be aware of how the values in a score map are distributed, then the type of distribution could be used for interpolation. To gain insight into this, we create 3D visualizations of some score maps resulted from high resolution images, like the example in Figure 14. The conclusion is that the correspondence values of

many score maps seem to be distributed with a 2D Gaussian or a 2D Laplacian distribution. According to the input point clouds the distribution of the similarity values in the score maps varies. Therefore, the selection of single distribution will not be the perfect solution. Nevertheless, we examine the suitability of a Gaussian distribution as the light incident on the laser scanner sensor is distributed nearly normally. Particularly, a least squares adjustment method is applied to find the optimal 2D elliptical Gaussian fit. Then the location of the highest amplitude of the 2D Gaussian surface is used as the 2D match location.

## 4. IMPLEMENTATION

The proposed method was implemented in Python 2.7. For the experiments a computer with Core(TM) i7 processor, CPU power 2.7GHz and 16GB RAM was used. No multi-processing or any other technique that could minimize the time execution was used to this point. Point cloud data from the West Paris, France and Sciedam, The Netherlands were provided by the company Cyclo-Media Technology B.V. In total 57 LAS files of point cloud tiles of 50m x 50m were processed. The overlapping point clouds pairs were 181. This, multiplied by 27 types of images gives a total of 4887 image registrations.

## 5. RESULTS

### 5.1 Execution Time

Approximately 51, 55, 77 and 257 minutes needed for the creation and registration of 4887 image pairs with image cell size of 0.5m, 0.3m, 0.15m and 0.05m correspondingly. Each image type refers to the reading of the 3D points, their binning into grid cells, the creation of the corresponding image and the application of the correctness evaluation criteria on the resulted score maps. The depth images include also the creation of the depth gradient images. The intensity images include in addition the retrieval of the intensity values from the LAS files and the generation of the intensity gradient images. The registration of the images is for the 0.5m, 0.3m and 0.15m grid cells faster than the creation of every image. The image registration of images with cell size 0.05m is more than 6.5 times slower than the registration with grid cell size 0.15m. The detection of the optimal solution is almost constant for all the cells sizes as it is not depended on the amount of cells, but on the amount of image pairs.

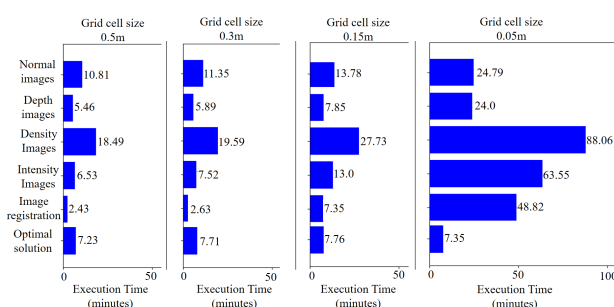


Figure 15. Execution time of algorithm's main steps when using grid cells size of 0.5m, 0.3m, 0.15m and 0.05m.

### 5.2 Case 1: Big overlap, changes in the scene

The registration result of the example point clouds shown in Figure 1c, is shown in Figure 16. Many points of the two point cloud



sets shown in Figure 1 do not have a correspondence in the other point cloud. It can also be seen in Figure 1b that there are changes in the scene. These are the points illustrated with light blue colour on the overlapping part. Despite these, the overlap between the two is quite big and this has influenced positively the registration as shown in Figure 16. The standard deviations of the final solutions and the amount of image types used for their calculation are given. All the image types passed the evaluation criteria. However, for the X and Y translations not all the image types gave the same results which is shown in their standard deviations.

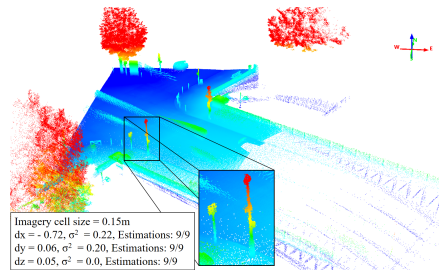


Figure 16. Relative transformation result of the example shown in Figure 1 with imagery of resolution 0.15m.

### 5.3 Case 2: Small overlap and different densities

Many points from the point cloud of Figure 17b do not have a correspondence in the point cloud of Figure 17a and their densities are very different. The two point clouds are superimposed in Figure 17c showing their offsets before registration. The magnified part highlights the small offset in Z between the two. The registration result is shown in Figure 17d. The optimum Z offset given from imagery with cell size 0.15m, is 0.20m, the overall standard deviation 0.0m and the image types' results included in the final estimations were 4 out of 9.

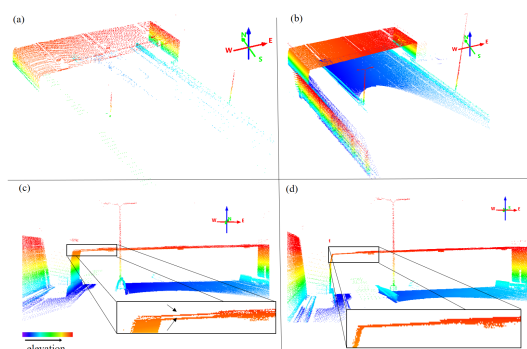


Figure 17. a) and b) are two overlapping point cloud sets. c) The point clouds are superimposed before registration. d) Shows the result of the registration with imagery of resolution 0.15m.

### 5.4 Case 3: Point clouds with large offset

Two point clouds with large offset in Z, and their registration result are illustrated in Figure 18. It can be seen at the magnified part of Figure 18d that there is still a small offset in Z (and X) after the registration. The Z offset with bin width 0.05m is 1.83m, the standard deviation 0.15m and the redundancy number 8/9. The Z offset with bin width 0.15m is 1.85m, the standard deviation is 0.13 and the redundancy number 9/9.

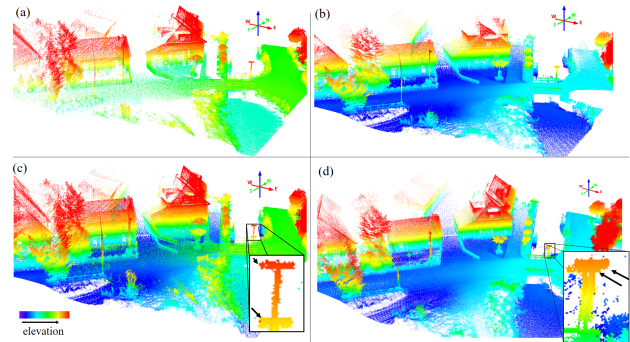


Figure 18. a) and b) are two overlapping point cloud sets. c) The a) and b) are superimposed. d) The result of the registration with imagery resolution 0.05m.

### 5.5 Sub-pixel accuracy

We examine to which extent the transformation parameters resulted from: a) registering imagery of low resolution and applying the proposed sub-pixel accuracy method, could approach the transformation parameters resulted b) from registering imagery of high resolution. The grid cell size of the low resolution imagery used for the experiments was 0.2m, and 0.05m of the high resolution imagery. Figure 19a indicates with box plots the absolute differences in X, Y and Z between the two approaches. Four large outliers are encountered. For visualization purposes the Figure 19a is magnified as shown in Figure 19b. The sub-pixel approach is successful for dx, dy and dz that are closer to zero. More than 25% of the dx, dy and dz are equal to or less than 0.05m. More than 50% of the dx, dy and dz are smaller than the accuracy of the low resolution images (0.2m). This indicates that the accuracy of the corresponding translation parameters was enhanced.

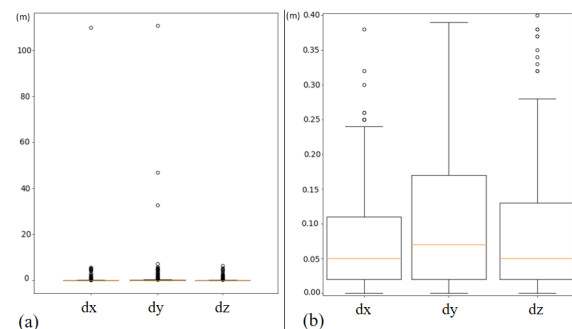


Figure 19. Differences between the translations resulted from the registration and the sub-pixel accuracy method of images with low resolution and the translations resulted from the registration of images with high resolution 0.05m

## 6. CONCLUSIONS

In this paper a method is proposed for relative point cloud registration by using the attributes of the points to create and match imagery. The execution time of the algorithm is barely dependant on the number of points in the tiles. It is also shown that the cross correlation technique used for the image registration performs very fast due to its simplicity. The method benefits from this and estimates redundant solutions which contribute to the determination of the confidence levels of the results. Some of the

results were investigated. Even when points in one scan do not have correspondence in the other, it is possible to produce a correct match if there is some overlap between them. Also, when there are large offsets between the scans, the algorithm converges to a solution that is close to a perfect one. The standard deviations of the results are not always small showing that the created imagery from the point clouds could be improved to produce more robust registrations. For example, only the 'mean' of the points' attributes that are binned into a cell is used to produce a single pixel value. Other statistical measures should be evaluated, too. Another reason for the large standard deviations could be that the evaluation criteria of the score maps do not reject all the bad results of. Thus, more sophisticated techniques must be researched to evaluate the strength of the similarity values in the score maps. Further future work would be the integration of the proposed relative registration method in a global registration. Then it would also be necessary to determine if the absolute positioning of any of the overlapping point clouds can be trusted more than the other. The error retrieved from the relative registration could be distributed between the overlapping scans to support the global registration.

Moreover, an interpolation method that employs an elliptical 2D Gaussian function to retrieve transformation parameters of sub-pixel accuracy is proposed. The method seems promising as for many translation parameters higher accuracy is reached. However, there are many cases where the method is not successful. Two conclusion are drawn. First, the Gaussian fitting is good but not the optimal option. Second, possibly the values in some score maps where the 2D Gaussian is fitted do not form a 2D peak. The improvement of the method is taken into account for future work.

## ACKNOWLEDGEMENTS

This project held in cooperation with the company CycloMedia Technology B.V. Specifically, the authors would like to thank Peter Joosten and Berry van Someren for the valuable knowledge, suggestions and feedback they provided, and Arjen Swart and Ravi Peters for the fruitful discussions.

## REFERENCES

- Baarda, W., 1968. *A testing procedure for use in geodetic networks. New Series, vol 2, no 5. Netherlands Geodetic Commission, Publications on Geodesy New Series. Delft, The Netherlands. Kanaalweg 4, Rijkscommissie voor Geodesie.*
- Besl, P. J. and McKay, N. D., 1992. A method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 14(2), pp. 239–256.
- Byun, S., Jung, K., Im, S. and Chang, M., 2017. Registration of 3d scan data using image reprojection. *International Journal of Precision Engineering and Manufacturing* 18(9), pp. 1221–1229.
- Chetverikov, D., Svirko, D., Stepanov, D. and Krsek, P., 2002. The Trimmed Iterative Closest Point algorithm. In: *Object recognition supported by user interaction for service robots*, Vol. 3, pp. 545–548.
- Ding, L., Goshtasby, A. and Satter, M., 2001. Volume image registration by template matching. *Image and Vision Computing* 19(12), pp. 821–832.
- Gaidhane, V. H., Hote, Y. V. and Singh, V., 2014. A computationally efficient approach for template matching-based image registration. *Sadhana* 39(2), pp. 317–331.
- Godin, G., Rioux, M. and Baribeau, R., 1994. Three-dimensional registration using range and intensity information. In: *Proc. SPIE 2350, Videometrics III*.
- Huber, D. F. and Hebert, M., 2003. Fully automatic registration of multiple 3d data sets. *Image and Vision Computing* 21(7), pp. 637 – 650.
- King, B. J., Malisiewicz, T., Stewart, C. V. and Radke, R. J., 2005. Registration of multiple range scans as a location recognition problem: hypothesis generation, refinement and verification. In: *Fifth International Conference on 3-D Digital Imaging and Modeling (3DIM'05)*, pp. 180–187.
- Levi, R. W. and Judd, T., 1996. Dead reckoning navigational system using accelerometer to measure foot impacts.
- Lin, C.-C., Tai, Y.-C., Lee, J.-J. and Chen, Y.-S., 2017. A novel point cloud registration using 2d image features. *EURASIP Journal on Advances in Signal Processing* 2017(1), pp. 5.
- Magnusson, M., Lilienthal, A. and Duckett, T., 2007. Scan registration for autonomous mining vehicles using 3d-NDT. *Journal of Field Robotics* 24(10), pp. 803–827.
- Naidu, D. K. and Fisher, R. B., 1991. A Comparative Analysis of Algorithms for Determining the Peak Position of a Stripe to Sub-pixel Accuracy. In: *BMVC91*, Springer, London, pp. 217–225.
- Pomerleau, F., Colas, F., Siegwart, R. and Magnenat, S., 2013. Comparing ICP variants on real-world data sets. *Autonomous Robots* 34(3), pp. 133–148.
- Sanchez, J., Denis, F., Checchin, P., Dupont, F. and Trassoudaine, L., 2017. Global Registration of 3d LiDAR Point Clouds Based on Scene Features: Application to Structured Environments. *Remote Sensing* 9(10), pp. 1014.
- Sarvaiya, J. N., Patnaik, S. and Bombaywala, S., 2009. Image Registration by Template Matching Using Normalized Cross-Correlation. In: *2009 International Conference on Advances in Computing, Control, and Telecommunication Technologies*, pp. 819–822.
- Scaramuzza, D., Harati, A. and Siegwart, R., 2007. Extrinsic self calibration of a camera and a 3d laser range finder from natural scenes. In: *2007 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 4164–4169.
- Segal, A. V., Haehnel, D. and Thrun, S., 2009. Generalized-ICP. In: *Proceedings of Robotics: Science and Systems*, Seattle.
- Sharp, G. C., Lee, S. W. and Wehe, D. K., 2002. ICP registration using invariant features. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24(1), pp. 90–102.
- Shetty, A., 2017. GPS-LiDAR sensor fusion aided by 3D city models for UAVs. PhD thesis, University of Illinois at Urbana-Champaign.
- Shlens, J., 2014. A Tutorial on Principal Component Analysis. *arXiv preprint arXiv:1404.1100*.
- Sweco Nederland B.V., 2016. MOVE3 Software User Manual, Version 4.4. <https://move3software.com/downloads/> (18 March 2016).
- Zhang, G., Lei, M. and Liu, X., 2009. Novel template matching method with sub-pixel accuracy based on correlation and Fourier-Mellin transform. *Optical Engineering* 48(5), pp. 057001.