

EFFICIENT ESTIMATION OF 3D SHIFTS BETWEEN POINT CLOUDS USING LOW-FREQUENCY COMPONENTS OF PHASE CORRELATION

R. Huang, Y. Xu*, L. Hoegner, U. Stilla

Photogrammetry and Remote Sensing, Technische Universität München, 80333 Munich, Germany
- (rong.huang, yusheng.xu, ludwig.hoegner, stilla)@tum.de

Commission II, ICWG II/III

KEY WORDS: Point cloud, 3D shifts estimation, Low frequency components, Robust phase Correlation

ABSTRACT:

Registration of multiple point clouds acquired via terrestrial laser scanning (TLS) is usually compulsory to obtain the scanned data covering a whole urban scene. However, the automated processing of aligning multiple scans is still a concern because of the complex urban environment. To this end, we propose a fast and sturdy estimation of 3D shifts between point clouds by an automated marker-free process using global features, converting translation measurement between two point clouds in the space domain to the frequency domain and estimating the phase difference. By using the low-frequency components from the normalized cross-power spectrum, accurate 3D shifts are calculated by solving parameters in the linear equation representing phase difference angles, with the help of a robust estimator. The results of experiments using TLS datasets of different scenes show that the proposed approach is both practical and efficient. In particular, the proposed approach can achieve results with a translation error of less than about 1.0 m on test datasets.

1. INTRODUCTION

Point Clouds acquired via Terrestrial Laser Scanning (TLS) are often used in numerous fields such as urban planning, transportation management, 3D modeling, virtual reality, development of digital surface models and environment monitoring (Vosselman and Maas, 2010) in recent years. The point cloud has been suggested as the most appropriate data source for the sake of 3D mapping in large-scale urban scenes because measured 3D points can provide directly spatial coordinates of measured surfaces, which simplifies the following surface reconstruction procedure. In order to cover a large urban scene in the mapping task, however, it is generally required to scan the scene with a laser system more than once, particularly in view of the occlusions of street building observations and scanner location (Yang et al., 2016, Dong et al., 2018). Therefore, a co-alignment of point clouds is ideally mandatory before point clouds are applied, in order to ensure that the entire scene is included (Dong et al., 2020).

In general, a standard registration method must address two critical issues for points clouds acquired with a TLS of calibrated sensors, namely, the extraction of geometric features and identification of associated features (Habib et al., 2010). Theoretically, if the respective feature pairs are known, the transition parameters between the reference frames can be well calculated (three shifts and three rotations, with no adjustments to the scale). However, the occlusion caused by complex urban environments and the low overlap rate between scans are challenging for the automatic registration of point clouds.

Areas with small overlaps and poor data quality can lead to a failed key points selection and geometric feature extraction. For instance, outliers and missing points can greatly influence the reliability of the extracted features. The points acquired in the TLS data sets themselves display enormous variations in density and decrease quadratically with increasing observation distances

(Hackel et al., 2016), which may also affect the characteristics of extracted features (Xu et al., 2017). In addition, due to similar shapes and street scenes, the complex and same layout of city scenes will cause the corresponding feature pairs to mismatch. On the other side, given the fact that the registration of point clouds is computational-intensive, the efficiency is also an essential factor. This should be taken into account when dealing with large data sets.

To solve these concerns, we present an automatic marker-free solution for grossly aligning two point clouds having 3D shifts via global features that are generated and aligned in the frequency domain. Features using global information are considered to be robust to the adverse effects of non-overlaps and unequal distributed densities of points. In addition, by converting 3D points into discrete 3D signals in the frequency domain, the high-frequency components corresponding to noise and outliers can be separated and rejected. Compared to methods that use points or geometric primitives (e.g., lines or planes) as registration features, the global feature generated from the entire point cloud considers all the underlying information of the observed scene, which actually results in limited degrees of freedom (DOF) of the geometric information (Xu et al., 2017) and at the same time enhancing the reliability. Moreover, in the frequency domain, the ill-posed registration task can be converted to a straightforward estimation of phase difference (Tong et al., 2015b), so that we can achieve a good balance between the efficiency and effectiveness for coarse but robust alignment of various 3D shifts.

2. LITERATURE REVIEW

To date, a considerable number of studies have offered solutions for aligning various point clouds through the use of geometric features. Such registration methods can generally be categorized according to the kind of features used into three main categories: points-based, primitive-based, and global-based approaches.

*Corresponding author

2.1 Point based registration

For the point-based approaches, the basic concept is to find corresponding point pairs from different point clouds. These found pairs of points will be used to establish the transformation between the local coordinate frames of source and target point clouds. For example, Iterative Closest Point (ICP) and its improved versions search for associated points on the basis of minimizing point-by-point distances between the various point clouds (Besl and McKay, 1992, Habib et al., 2010, Al-Durgham and Habib, 2013). Likewise, 4-point congruent systems (4PCS) and its enhanced versions are also representative approaches for finding corresponding points, using particular and constant sets of conglomerate points with affine invariant ratios as the constraint (Aiger et al., 2008, Mellado et al., 2014, Theiler et al., 2014). Either the ICP- or 4PCS-based methods have proven to be of good performance in many applications. Nonetheless, a time-consuming iterative process is required when a good initial alignment is not achieved. Thus, instead of selecting all the points as candidates, only utilizing key points is a solution that greatly reduces the cost of computation. SIFT key points (Böhm and Becker, 2007, Weinmann et al., 2011), DoG key points (Theiler et al., 2014), the PFH and FPFH key points (Weber et al., 2015), virtual intersections (Theiler and Schindler, 2012), and the points of predefined semantics (Yang et al., 2016, Ge, 2017) are examples. While they are all typically capable of aligning point clouds, the point-oriented approaches are susceptible to point densities, noise, and problems when dealing with large-scale datasets in terms of efficiency.

2.2 Primitive based registration

For the primitive-based method, rather than points, geometric primitives (e.g. lines, curves, planes, or surfaces) created from points are extracted as the description of geometrical characteristics serving the registration. The application of high-level geometric features can increase the accuracy of the identification of the appropriate characteristic pairs. The representatives of geometric primitives include lines, planes, and surfaces. A vast number of articles have been published using lines or planes as features to align point clouds. For example, lines created by intersecting planes (Stamos and Leordeanu, 2003), 3D straight-lines (Habib et al., 2005, Al-Durgham and Habib, 2013), and spatial curves (Yang and Zang, 2014) are used as line primitives for matching. In comparison, planes (Dold and Brenner, 2006, Von Hansen, 2006, Xu et al., 2017, Xu et al., 2019, Chen et al., 2019) and surfaces (Ge and Wunderlich, 2016) are also used as basic corresponding primitives for alignment. In contrast with point-based methods, either line-based or plane-based methods require an abundance of linear objects or smooth surfaces as candidates, which depends mainly on the content and layout of scanned scenes. In a scene of only a few houses, problems can occur when we try to find suitable candidate characteristics. Using voxel structure to represent points is an alternative in the context of the registration task as well. (Wang et al., n.d.) has recently proposed a method using correspondences of EGI features from voxel clusters which shows efficient results when aligning indoor scenes. Moreover, in the work of (Xu et al., 2017, Xu et al., 2019), the voxel structure is combined with plane primitives for fast and coarse registration between point clouds.

2.3 Global feature based registration

For both of the above registration methods, they all use local information from the point clouds itself or from primitive clusters. In fact, registration can also take the whole point cloud with

global features. In the NDT algorithm (Magnusson et al., 2007) built for 2D space position and mapping and expanded into 3D space, all points have been transformed into a normal distribution so as to force the point cloud alignment. Point density at the global scale is also used for alignment in coherent point-drift (Myronenko and Song, 2010) and kernel affinity correlation (Tsin and Kanade, 2004), respectively. (Dong et al., 2018) presents a global feature vector for the registration of multiple point clouds without knowing the viewing orders or locations. In addition, 3D point clouds of as-built buildings are decoupled into 1D histograms and 2D images for registration in the frequency domain (Huang et al., 2019), which correspond to the matching of low-frequency components at a global scale in the space domain. Theoretically, features using global information can provide a more robust and comprehensive description of the 3D geometry from points (Huang et al., 2020). Thus, global feature-based registration methods are more reliable but less explicit than those based on local features, but it normally requires a large overlapping ratio between scans. Otherwise, the global features may have significant differences.

2.4 Our contributions

The major contribution of this work is twofold, including: (1) A 3D phase correlation-based method is developed for matching point clouds with 3D shifts, converting the problem of measuring 3D translations to the estimation of the phase difference between to 3D tensors; (2) A RANSAC-based strategy is applied to fit the low-frequency components of the correlated cross-spectrum, which can avoid the high-frequency noise caused by the details and non-overlapping area of point clouds.

3. METHOD

Our proposed method consists of three steps: discretization of point clouds, phase correlation and frequency components selection, and robust estimation of 3D shifts. In Fig. 1, the workflow is illustrated, showing the core steps of involved methods and sample results. A detailed explanation of each step will be given in the following subsections.

3.1 Discretization of point clouds

The discretization of point clouds is a resampling of 3D data, which is designed to convert two original point clouds with 3D shifts into two 3D signals with mainly phase differences. To achieve this, two major steps are implemented, namely, the voxelization and binarization and the rough alignment of bounding boxes.

3.1.1 Voxelization and binarization The voxelization is to resample unevenly distributed points, as well as the 3D space, into a regularly sampled discrete 3D grid. The point cloud is depicted by centers of all cells (i.e., voxels) in this 3D grid. It is noteworthy that unlike the voxelization steps in other previous studies like (Xu et al., 2018, Wang et al., n.d., Gehrung et al., 2016), which voxelized only the point cloud, here, the entire 3D space covering the point cloud will be voxelized, and centers of all these voxels will be used as input in the phase correlation later. For assigning attributes to all voxels, we conduct a binarization annotating voxels with binary values, namely zero or one. The annotating binary value for a voxel is identified in accordance with the occupancy of this voxel. In other words, if there are points fallen into a voxel, this voxel will be annotated with value

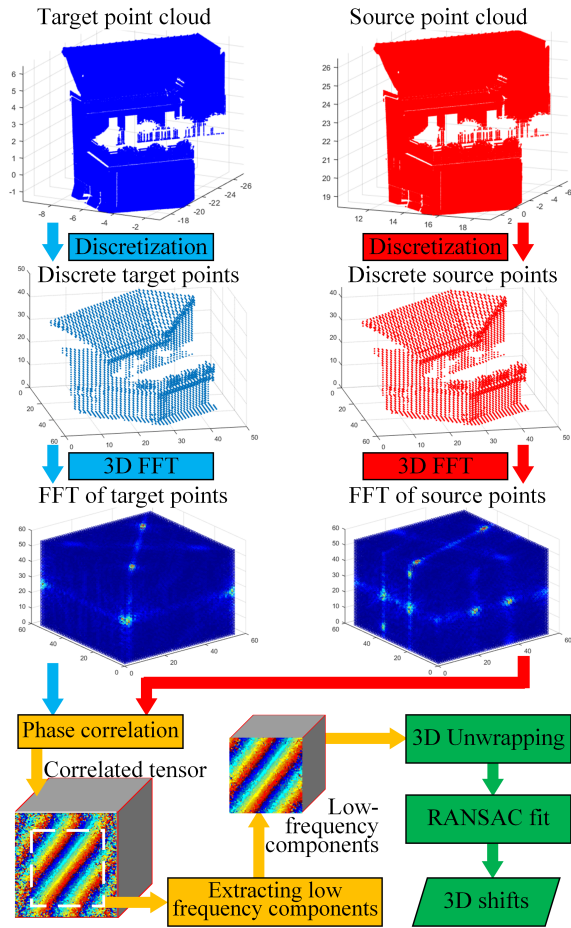


Figure 1. Workflow of the proposed 3D shifts estimation method.

one. In contrast, if a voxel is an empty one, it will be annotated with value zero. Their centers will represent the annotated voxels with assigned labels as input for further processing. In this way, a regularly sampled 3D position inside the bounding box of the given point cloud will be assigned with 3D coordinates and a binary label. After the voxelization and binarization, the original point cloud has been converted to a regularly sampled discrete 3D signal with two attributes for each element. In Fig. 2, a top-viewed illustration shows the voxelization and binarization of 3D point clouds.

3.1.2 Rough alignment using bounding boxes After the voxelization and binarization of point clouds, the translation difference (i.e., X_0 , Y_0 , and Z_0) between center coordinates of bounding boxes of source and target point clouds will be calculated. This calculated translation is regarded as a rough alignment between point clouds, which can reduce large phase angle differences in the normalized cross-power spectrum.

3.2 Phase correlation and frequency components selection

Once the point cloud is converted from raw 3D points into regularly sampled 3D signals, a 3D phase correlation will be conducted by the use of 3D fast Fourier transformation. On the basis of the phase correlation result, the low-frequency components will be extracted and then used for estimating accurate 3D shifts.

3.2.1 3D fast fourier transformation Before introducing the 3D phase correlation method used in this paper, a short introduction to phase correlation will be given. Compared with some

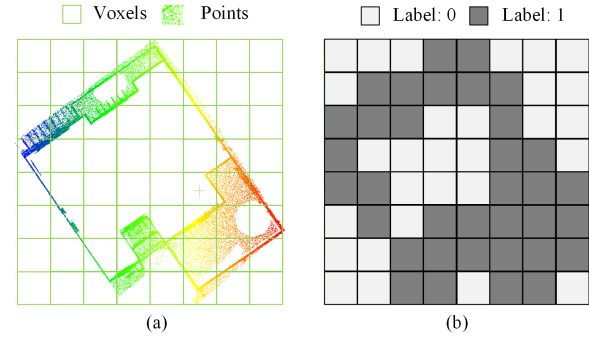


Figure 2. A 2D illustration of voxelization and binarization. For simplifying the sketches, we show top views of the results as examples. (a) Voxelization of the point cloud (top view). (b) Binarization of the voxels (top view).

other commonly used correlation-based methods, phase correlation tends to be more accurate and efficient. The general idea of phase correlation is that any translation between two relevant signals (i.e., 3D discrete points) in the spatial domain can be represented as a phase shift in the frequency domain. Assume that two point clouds are related to each other by shifts in x -, y -, and z - directions denoted as δx , δy , and δz , respectively, and the relation can be presented by:

$$s(x, y, z) = r(x - \delta x, y - \delta y, z - \delta z) \quad (1)$$

where $s(x, y, z)$ and $r(x, y, z)$ represent the two point clouds in spatial domain. Then, a 3D fast Fourier transformation (FFT) can be conducted on these two discrete point clouds to transform them into the frequency domain.

$$S(u, v, w) = FFT(s(x, y, z)) \quad (2)$$

$$R(u, v, w) = FFT(r(x, y, z)) \quad (3)$$

in which $S(u, v, w)$ and $R(u, v, w)$ are the corresponding Fourier transforms of $s(x, y, z)$ and $r(x, y, z)$.

3.2.2 3D phase correlation Afterward, the phase correlation between $S(u, v, w)$ and $R(u, v, w)$ is carried out. After the correlation, the relation between the Fourier transformed point clouds can be written as:

$$S(u, v, w) = R(u, v, w)e^{-i(u\delta x + v\delta y + w\delta z)} \quad (4)$$

The normalized cross spectrum can be represented as:

$$Q(u, v, w) = \frac{S(u, v, w)R^*(u, v, w)}{|S(u, v, w)R^*(u, v, w)|} = e^{-i(u\delta x + v\delta y + w\delta z)} \quad (5)$$

in which R^* is the complex conjugate of R , and the magnitude of Q is normalized to 1. The inverse Fourier transform (IFT) of $Q(u, v, w)$ is a Dirac delta function centered on $(\delta x, \delta y, \delta z)$. Thus, the translation can be estimated by finding the peak coordinates of this function. However, this solution can only provide results in the accuracy of integer voxels, which is not precise enough for our demand and highly depends on the resolution of voxels. To tackle this problem, instead of finding the peak in the IFT of the normalized cross-power spectrum, we provide a solution for identifying sub-voxel shifts directly in the frequency domain. Since the phase different angle is a linear function of the shift parameters defined by:

$$\angle Q(u, v, w) = u\delta x + v\delta y + w\delta z, \quad (6)$$

As seen from this equation, we can easily find that the representation of the phase difference angle equals to a linear function with three unknown parameters representing shifts. By linear regression, a solution of these parameters can be achieved (Tong et al., 2015a). However, this is the ideal situation, and the real situation is that source and target point clouds are not fully overlapped. Moreover, there would be noise and outliers existing in the point clouds. All these will result in high-frequency noise in the cross-power spectrum tensor. To avoid these high-frequency disturbances, we need a selection of desired frequency components.

3.2.3 Extraction of low-frequency components Based on the computed Fourier spectra of the individual point clouds using FFT and their normalized cross-power spectrum matrix, it is necessary to conduct a selection of frequency components for extracting those low-frequency ones. Here, it is assumed that the 3D phase correlation between point clouds has similar characteristics like the 2D phase correlation between images. Thus, as spectral components at a high frequency are the most likely to be biased due to aliasing and noise, most of the energy is mainly concentrated in the low-frequency components for 2D natural images in the 2D case (Leprince et al., 2007, Tong et al., 2015b). For the cross-power spectrum from the 3D phase correlation, a similar strategy is used masking out around 80% of frequency components at the periphery of the tensor Q . Namely, only the core of the tensor Q will be preserved for estimating the parameters of a linear function.

3.3 Robust estimation of 3D shifts

Once the low-frequency components in the cross-power spectrum are extracted, one can still follow the linear function given in Eq.6. In order to estimate the parameters of this linear function, a RANSAC algorithm is adopted, in which a subset with a minimized required size of randomly sampled data is utilized to fit the model. In general, the idea of RANSAC is to estimate the model parameters with the sampled data with a minimized size (i.e., four points in this tensor) and then to check whether other data samples are within a predefined threshold. However, the directly obtained cross-power spectrum is a wrapped tensor. Therefore, prior to the estimation of parameters, a 3D unwrapping is required. Then, the parameters (δx , δy , δz) of the unwrapped phase angles of the identified components can be converted to the real estimated shift parameters (ΔX , ΔY , ΔZ):

$$\Delta X = \delta x M / (2\pi) \quad (7)$$

$$\Delta Y = \delta y N / (2\pi) \quad (8)$$

$$\Delta Z = \delta z L / (2\pi) \quad (9)$$

where M , N , L denote the dimensions of the input tensor. Afterward, considering the difference of the coordinates (X_0 , Y_0 , Z_0) calculated from the rough alignment, the estimated 3D shifts should be ($X_0 + \Delta X$, $Y_0 + \Delta Y$, $Z_0 + \Delta Z$), which are the final outputs.

4. EXPERIMENTS AND RESULTS

4.1 Testing datasets

The proposed method for performing automated registration was tested using TLS datasets of various scanners and scenes. Specifically, three datasets were tested. The first dataset is a pair of TLS point clouds from the ThermalMapper project acquired by

the Jacobs University Bremen (see Fig. 3a) (Borrmann et al., 2013). The second dataset is from a pair of TLS point clouds from the large-scale point cloud classification benchmark (Semantic3D) datasets published by ETH Zurich (Hackel et al., 2016), which is the area of Cathedral of St. Gallen (see Fig. 4a). The last one is a set of scans from the Real-world Scans with Small Overlap (RESSO) dataset (see Fig. 5a) (Chen et al., 2019). For this dataset, we used three pairs of TLS point clouds. The detailed information of the datasets is listed in Table 1. For the first two datasets, we manually aligned source and target point clouds as ground truth, while, as a registration benchmark, the last one provided the accurately aligned source and target scans as ground truth. Based on the ground truth, we manually translated the source point cloud with predefined 3D shifts. Then, the matching was performed between the source and target point clouds, and the root mean squared error (RMSE) of estimated 3D shifts served as the criterion for the evaluation. Our method was implemented in Matlab, and all the experiments were conducted on a computer with an Intel i7-4710MQ CPU and 16GB RAM.

Datasets	Number of points and overlaps		
	Source	Target	approx. overlap ratios
Bremen	15033238	15246878	0.85
StGallen	31640124	31149413	0.62
Resso Pair0-10	469284	461612	0.82
Resso Pair2-3	825915	618672	0.60
Resso Pair5-6	607874	501967	0.71

Table 1. Detailed information of experimental datasets.

4.2 Influence of different shifts

The first experiment is about the influence of different shifts. In this experiment, we manually translated the source point cloud of the Bremen dataset with predefined 3D shifts, ranging from 10m to 100m with an incremental rate of 10m. In these tests, the size of the voxel was set to 2m, and the threshold for RANSAC fitting was set to 0.1. The maximum iteration number of RANSAC process was limited to 20000 times. In Fig. 6, the 3D shifts estimation results are displayed.

As seen from Figure 6, we can find that the performance of the proposed method is insensitive to the shifts between source and target point clouds. The RMSE of the estimated 3D shift was around 0.65 m. This is because that the rough alignment using bounding boxes has offset any large shifts. After such a rough alignment, for the same pair of point clouds, the 3D shift estimation is always the same problem. It is also interesting that the execution time of tests reveals a slightly decreasing tendency, but it is still within a range between 25 s and 35 s. Considering the number of points, such an execution time indicates our method is an efficient solution. This is also a major advantage of global feature-based registration methods.

4.3 Influence of different resolutions in the discretization

Besides the shifts, for the proposed method, the resolution of voxels is also an essential factor that can significantly influence the quality of shift estimation. The resolution is actually the geometric size of each voxel used in the voxelization and binarization. In these experiments, we manually translated the source point cloud of the Bremen and St.Gallen datasets with predefined 3D shifts of 50 m. In these tests, the size of voxel ranged from 1.0 m to 5.0 m with an incremental rate of 1.0 m, and the threshold for

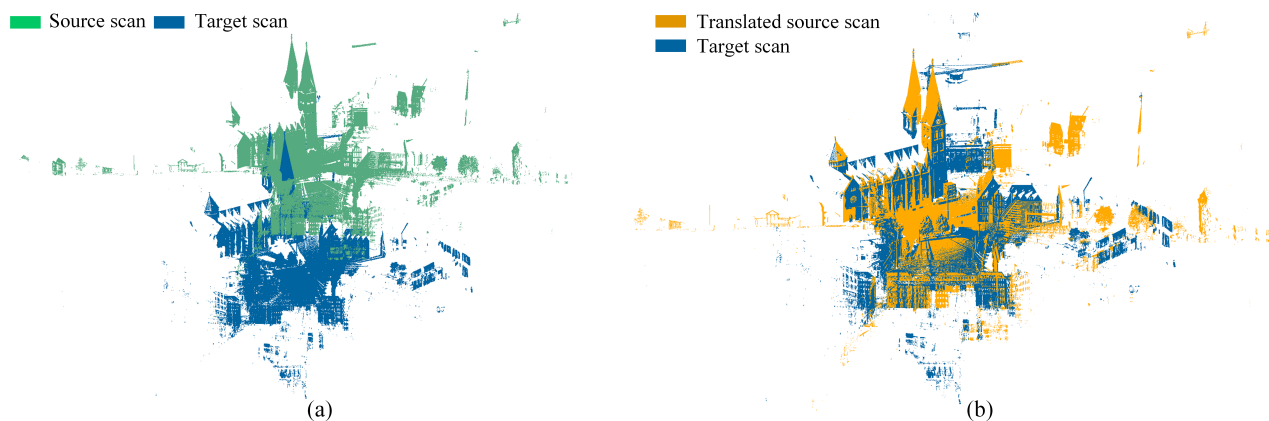


Figure 3. Data and results of Bremen dataset. (a) Source and target point clouds. (b) Matching result.

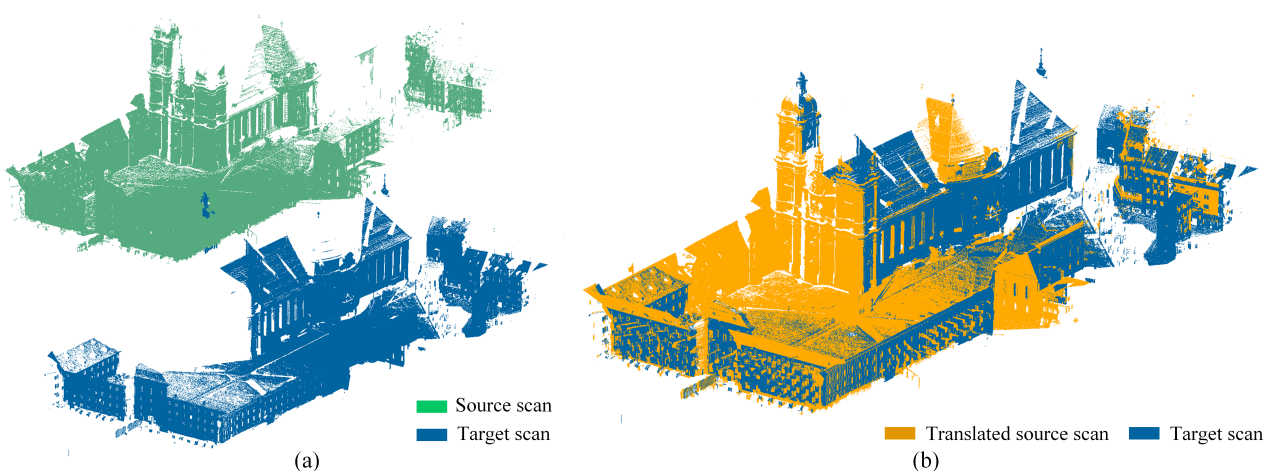


Figure 4. Data and results of StGallen dataset. (a) Source and target point clouds. (b) Matching result.

RANSAC fitting was set to 0.1. The maximum iteration number of RANSAC process was also limited to 20000 times. In Fig. 7b, the 3D shifts estimation results are given.

As one can expect, the larger the voxel is, the faster the execution of our method will be. The experimental result proves this assumption. For both two datasets, the execution time experiences a significant decrease along with the increase of voxel resolution. On the other hand, it is also evident that the RMSE of the estimated 3D shift for both two datasets shows a drastic improvement from less than 0.5 m to more than 1.5 m. This is mainly because the voxel resolution in the discretization directly determines the sampling rate of 3D points. A large voxel size, namely a low voxel resolution, indicates a sparse sampling of the point cloud, which would be easier influenced by the aliasing effect. To give a better illustration of the dataset and shift estimation results, in Figs. 3b and 4b, source and target point clouds and matching results using a voxel resolution of 2.0 m are displayed.

4.4 Influence of different overlapping ratio

The overlapping ratio is another factor that matters to the proposed method. This is because scans in the real-world typically have unpredictable varying overlap ratios, which challenges any registration algorithms (Chen et al., 2019). In this experiment, six scans from the RESSO dataset were arranged into three pairs,

and the source point clouds of these pairs were manually translated by 50 m. The size of voxels was set 1.0 m, 2.0 m, and 3.0 m, respectively, and the threshold for RANSAC fitting was set to 0.1. The maximum iteration number of RANSAC process was also limited to 20000 times. In Fig. 8, the 3D shifts estimation results are shown.

As seen from Figure 8, we can observe the tendency that the execution time reveals a significant decrease along with the increase of voxel resolution, while the RMSE will increase when the voxel size is getting larger, except the result of the third pair. According to these observed results, we can conclude that our proposed method is robust to the datasets with various overlapping ratios. In Fig. 5b, the source and target point clouds and matching results using a voxel resolution of 2.0 m are displayed. From these illustrated results, we can definitely confirm that our proposed method shows a satisfying performance of efficient estimation of 3D shifts between point clouds of different overlapping ratios.

5. CONCLUSION

This work proposed an automatic marker-free method using global features for a fast and robust estimation of 3D shifts between point clouds, which converted the 3D translation calculation between two point clouds from the space domain to the estimation of phase difference in the frequency domain. By utilizing

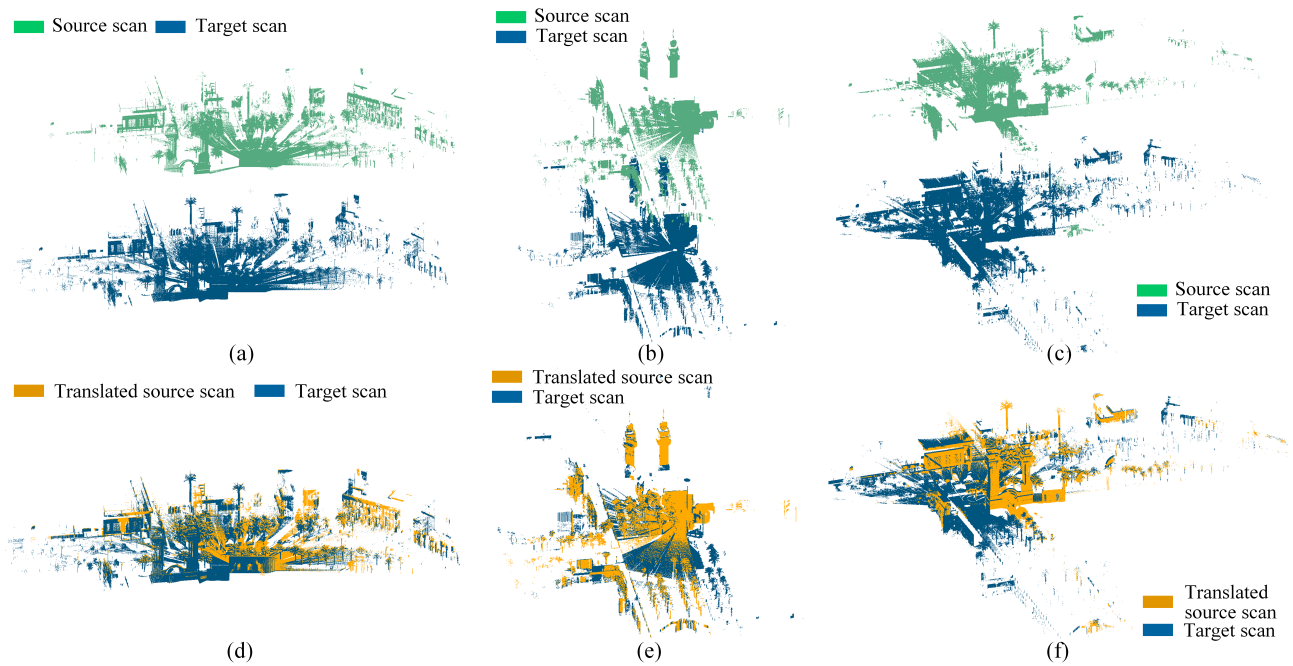


Figure 5. Data and results of Resso dataset. (a), (b) and (c) source and target point clouds. (d), (e), and (f) matching results.

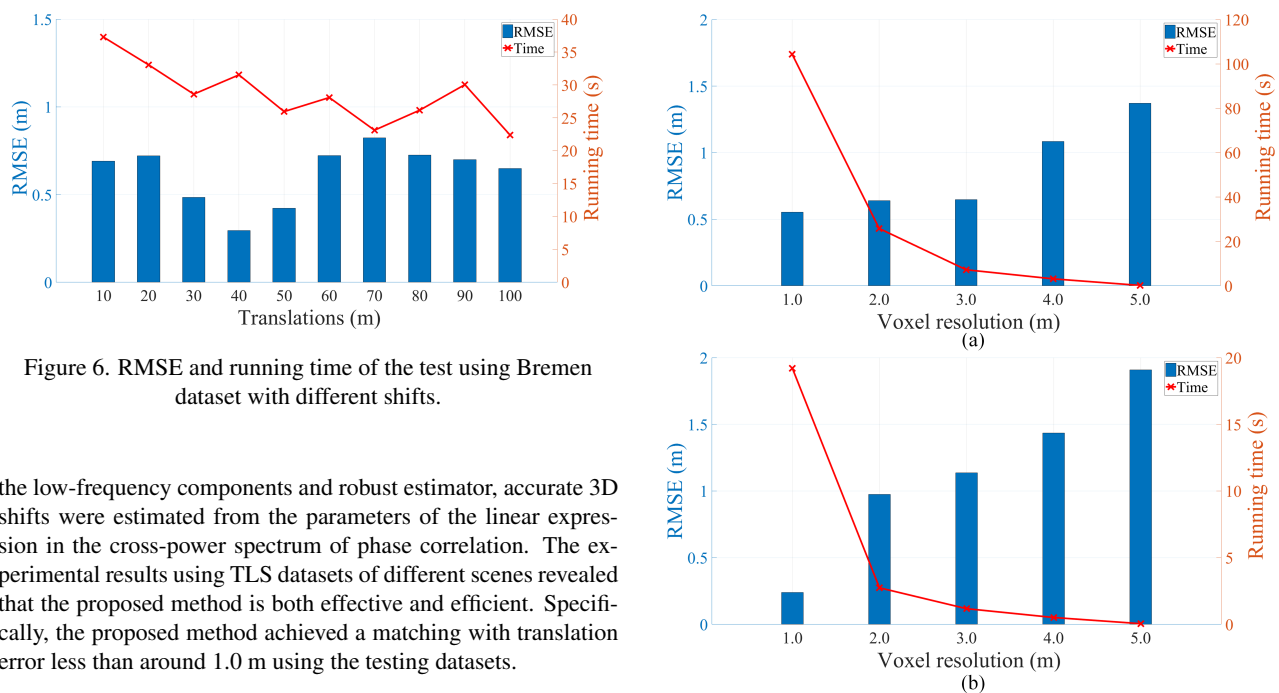


Figure 6. RMSE and running time of the test using Bremen dataset with different shifts.

Figure 7. RMSE and running time of the tests using (a) Bremen dataset and (b) StGallen dataset with different voxel resolutions.

the low-frequency components and robust estimator, accurate 3D shifts were estimated from the parameters of the linear expression in the cross-power spectrum of phase correlation. The experimental results using TLS datasets of different scenes revealed that the proposed method is both effective and efficient. Specifically, the proposed method achieved a matching with translation error less than around 1.0 m using the testing datasets.

However, the current method and conducted experiments still have some drawbacks that could be improved in future work. For example, the rotation is not considered in the current solution, which is actually a more challenging task than just estimating the translation. Fortunately, the angle difference between two point clouds could also be converted into a phase difference in the frequency domain using log-polar transformation. Successful cases have been achieved for the 2D image via Fourier-Mellin transformation, which can be drawn on for the 3D case as well. Moreover, in the binarization of the 3D point cloud, the current method considers only the occupancy of the 3D space, but actually the point density also plays a role in the binarization, which can increase the quantization accuracy of the 3D signal, provid-

ing richer information in the phase correlation. Therefore, future work will focus on the following points, on the basis of the current method:

- The estimation of rotation between point clouds using global features in the frequency domain.
- The registration of multi-source point clouds with varying

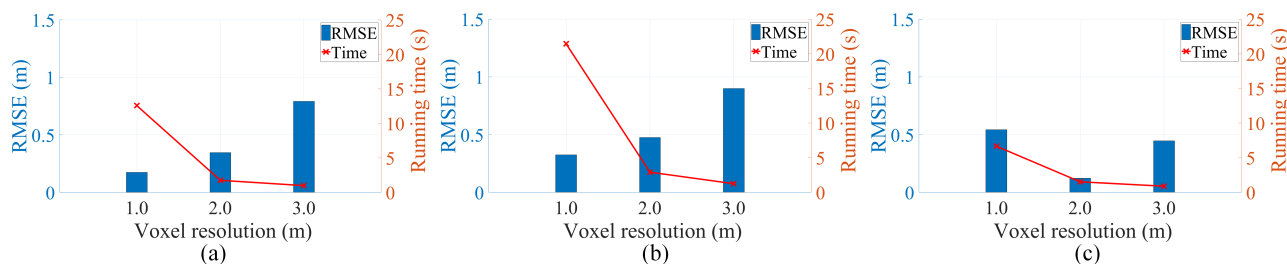


Figure 8. RMSE and running time of the test using RESSO dataset with different overlapping ratios.

density, overlaps, and noise levels.

- A refinement based on the coarse registration results for a fine registration.

All these points would be further tasks that needed to be addressed so that the potential of using global features for point clouds registration can be fully explored.

ACKNOWLEDGEMENTS

The authors would like to appreciate the help from Dr. Zhen Ye from Remote Sensing Technology in Technische Universität München. This work was carried out within the frame of Leonhard Obermeyer Center (LOC) at Technische Universität München (TUM) [www.loc.tum.de].

REFERENCES

- Aiger, D., Mitra, N. J. and Cohen-Or, D., 2008. 4-points congruent sets for robust pairwise surface registration. In: *ACM SIGGRAPH 2008 papers on - SIGGRAPH 08*, Vol. 27number 3, ACM Press, p. 85.
- Al-Durgham, M. and Habib, A., 2013. A framework for the registration and segmentation of heterogeneous lidar data. *Photogrammetric Engineering & Remote Sensing* 79(2), pp. 135–145.
- Besl, P. J. and McKay, N. D., 1992. Method for registration of 3-d shapes. In: *Robotics-DL tentative*, International Society for Optics and Photonics, pp. 586–606.
- Böhm, J. and Becker, S., 2007. Automatic marker-free registration of terrestrial laser scans using reflectance. In: *Proceedings of 8th Conference on Optical 3D Measurement Techniques, Zurich, Switzerland*, pp. 338–344.
- Borrmann, D., Elseberg, J. and Nüchter, A., 2013. Thermal 3d mapping of building façades. In: *Intelligent Autonomous Systems 12*, Springer, pp. 173–182.
- Chen, S., Nan, L., Xia, R., Zhao, J. and Wonka, P., 2019. Plade: A plane-based descriptor for point cloud registration with small overlap. xxx(xxx), pp. xxx.
- Dold, C. and Brenner, C., 2006. Registration of terrestrial laser scanning data using planar patches and image data. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 36(5), pp. 78–83.
- Dong, Z., Liang, F., Yang, B., Xu, Y., Zang, Y., Li, J., Wang, Y., Dai, W., Fan, H., Liang, X. and Stilla, U., 2020. Registration of large-scale tls point clouds: A review and benchmark. *ISPRS Journal of Photogrammetry and Remote Sensing*.
- Dong, Z., Yang, B., Liang, F., Huang, R. and Scherer, S., 2018. Hierarchical registration of unordered tls point clouds based on binary shape context descriptor. *ISPRS Journal of Photogrammetry and Remote Sensing* 144, pp. 61 – 79.
- Ge, X., 2017. Automatic markerless registration of point clouds with semantic-keypoint-based 4-points congruent sets. *ISPRS Journal of Photogrammetry and Remote Sensing* 130, pp. 344–357.
- Ge, X. and Wunderlich, T., 2016. Surface-based matching of 3d point clouds with variable coordinates in source and target system. *ISPRS Journal of Photogrammetry and Remote Sensing* 111, pp. 1–12.
- Gehring, J., Hebel, M., Arens, M. and Stilla, U., 2016. A framework for voxel-based global scale modeling of urban environments. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 42, pp. 45–51.
- Habib, A., Datchev, I. and Bang, K., 2010. A comparative analysis of two approaches for multiple-surface registration of irregular point clouds. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*.
- Habib, A., Ghanma, M., Morgan, M. and Al-Ruzouq, R., 2005. Photogrammetric and lidar data registration using linear features. *Photogrammetric Engineering & Remote Sensing* 71(6), pp. 699–707.
- Hackel, T., Wegner, J. D. and Schindler, K., 2016. Fast semantic segmentation of 3d point clouds with strongly varying density. *ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences*.
- Huang, R., Xu, Y., Hong, D., Yao, W., Ghamisi, P. and Stilla, U., 2020. Deep point embedding for urban classification using als point clouds: A new perspective from local to global. *ISPRS Journal of Photogrammetry and Remote Sensing* 163, pp. 62–81.
- Huang, R., Ye, Z., Boerner, R., Yao, W., Xu, Y. and Stilla, U., 2019. Fast pairwise coarse registration between point clouds of construction sites using 2d projection based phase correlation. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XLII-2/W13*, pp. 1015–1020.
- Leprince, S., Barbot, S., Ayoub, F. and Avouac, J.-P., 2007. Automatic and precise orthorectification, coregistration, and subpixel correlation of satellite images, application to ground deformation measurements. *IEEE Transactions on Geoscience and Remote Sensing* 45(6), pp. 1529–1558.
- 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.

- Mellado, N., Aiger, D. and Mitra, N. J., 2014. Super 4pcs fast global pointcloud registration via smart indexing. In: *Computer Graphics Forum*, Vol. 33number 5, Wiley Online Library, pp. 205–215.
- Myronenko, A. and Song, X., 2010. Point set registration: Coherent point drift. *IEEE transactions on pattern analysis and machine intelligence* 32(12), pp. 2262–2275.
- Stamos, I. and Leordeanu, M., 2003. Automated feature-based range registration of urban scenes of large scale. In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol. 2, IEEE, pp. II–II.
- Theiler, P. and Schindler, K., 2012. Automatic registration of terrestrial laser scanner point clouds using natural planar surfaces. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 3, pp. 173–178.
- Theiler, P. W., Wegner, J. D. and Schindler, K., 2014. Keypoint-based 4-points congruent sets—automated marker-less registration of laser scans. *ISPRS Journal of Photogrammetry and Remote Sensing* 96, pp. 149–163.
- Tong, X., Xu, Y., Ye, Z., Liu, S., Li, L., Xie, H., Wang, F., Gao, S. and Stilla, U., 2015a. An improved phase correlation method based on 2-d plane fitting and the maximum kernel density estimator. *IEEE Geoscience and Remote Sensing Letters* 12(9), pp. 1953–1957.
- Tong, X., Ye, Z., Xu, Y., Liu, S., Li, L., Xie, H. and Li, T., 2015b. A novel subpixel phase correlation method using singular value decomposition and unified random sample consensus. *IEEE Transactions on Geoscience and Remote Sensing* 53(8), pp. 4143–4156.
- Tsin, Y. and Kanade, T., 2004. A correlation-based approach to robust point set registration. In: *European conference on computer vision*, Springer, pp. 558–569.
- Von Hansen, W., 2006. Robust automatic marker-free registration of terrestrial scan data. *Proc. Photogramm. Comput. Vis* 36, pp. 105–110.
- Vosselman, G. and Maas, H.-G., 2010. *Airborne and terrestrial laser scanning*. CRC Press.
- Wang, J., Lindenbergh, R. and Menenti, M., n.d. Voxel based scalable registration of laser scanned point cloud data by neighbourhood voting. In: *PROCEEDINGS, IQMULUS IST WORKSHOP ON PROCESSING LARGE GEOSPATIAL DATA*, p. 39.
- Weber, T., Hänsch, R. and Hellwich, O., 2015. Automatic registration of unordered point clouds acquired by kinect sensors using an overlap heuristic. *ISPRS Journal of Photogrammetry and Remote Sensing* 102, pp. 96–109.
- Weinmann, M., Weinmann, M., Hinz, S. and Jutzi, B., 2011. Fast and automatic image-based registration of tls data. *ISPRS Journal of Photogrammetry and Remote Sensing* 66(6), pp. S62–S70.
- Xu, Y., Boerner, R., Yao, W., Hoegner, L. and Stilla, U., 2017. Automated coarse registration of point clouds in 3d urban scenes using voxel based plane constraint. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* IV-2/W4, pp. 185–191.
- Xu, Y., Boerner, R., Yao, W., Hoegner, L. and Stilla, U., 2019. Pairwise coarse registration of point clouds in urban scenes using voxel-based 4-planes congruent sets. *ISPRS journal of photogrammetry and remote sensing* 151, pp. 106–123.
- Xu, Y., Heogner, L., Tuttas, S. and Stilla, U., 2018. A voxel- and graph-based strategy for segmenting man-made infrastructures using perceptual grouping laws: comparison and evaluation. *Photogrammetric Engineering & Remote Sensing*.
- Yang, B. and Zang, Y., 2014. Automated registration of dense terrestrial laser-scanning point clouds using curves. *ISPRS Journal of Photogrammetry and Remote Sensing* 95, pp. 109–121.
- Yang, B., Dong, Z., Liang, F. and Liu, Y., 2016. Automatic registration of large-scale urban scene point clouds based on semantic feature points. *ISPRS Journal of Photogrammetry and Remote Sensing* 113, pp. 43–58.