FAST PAIRWISE COARSE REGISTRATION BETWEEN POINT CLOUDS OF CONSTRUCTION SITES USING 2D PROJECTION BASED PHASE CORRELATION

Rong Huang[1], Zhen Ye[1], Richard Boerner[1], Wei Yao[2], Yusheng Xu[1], Uwe Stilla[1]

[1]Photogrammetry and Remote Sensing, Technische Universität München, 80333 München, Germany
- (rong.huang, z.ye, richard.boerner, yusheng.xu, stilla)@tum.de
[2]Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hung Hom, Hong Kong
- wei.hn.yao@polyu.edu.hk

Commission II, WG III

KEY WORDS: Coarse registration, point cloud, automation construction, 2D matching, phase correlation

ABSTRACT:

For conducting change detection using 3D scans of a construction site, the registration between point clouds at different acquisition times is normally necessary. However, due to the complexity of constructing areas, the automatic registration of temporal scans is a challenging problem. In this work, we propose a fast and maker-free method for coarse registration between point clouds by converting the 3D matching problem into a 2D correlation problem, taking the special properties of building structures into consideration. Our proposed method consists of two major steps: the conversion from 3D points to 2D image data and the estimation of transformation parameters between 2D images in the frequency domain. In the conversion step, the point cloud of each scan is projected into a 2D grey image, by which the ground footprint of the point cloud is obtained. In the following step, we represent the 2D image in frequency-domain and estimate the horizontal transformation parameters by using Fourier-Mellin transformation. A real application is performed to validate the feasibility and effectiveness of our workflow using photogrammetric point clouds of a construction site in two different acquisition time. Regarding the real application of coarse registration of point clouds, our proposed method can achieve a registration error of less than 1 degree and more efficient than the classical baseline methods for the fast orientation between scans.

1. INTRODUCTION

In the recent decade, in the Architectural, Engineering and Construction (ACE) fields, there is an exponential increase demand for efficient and effective progress monitoring of construction sites (Bosché et al., 2015). Traditional methods of progress monitoring depend on visual inspections, and extensive manual records and data analysis (Seidell III., 2002). These traditional methods rely heavily on the personal skills, and experiences of the inspectors, which leads to a high possibility of incompetence and inaccuracy in recording the whole progress. Simultaneously, it is inevitable to be time-consuming with extensive manual work. Therefore, an increased automation is urgently required for construction progress monitoring.

For this need, a large number of researchers have reported related methods using 2D imaging, 3D photogrammetry and laser scanning for automatically acquiring measurements in construction sites. Among the aforementioned techniques, point clouds, acquired via laser scanning or stereo matching of 2D images, have been the best available data source for acquiring 3D information in construction monitoring tasks (Tuttas et al., 2014, Tuttas et al., 2017, Xu et al., 2018). In order to monitor the progress of a construction site and obtain a full coverage over time, several scans are needed to be acquired in different times during the construction progress. Hence, the registration of scans from different acquisition time is the base of further point cloud processing. However, for construction sites, it always has a very complex environment, so that acquired point clouds could be influenced by noises and outliers resulting from errors of stereo matching, uneven point cloud density caused by varying measuring distances of camera or scanners, occlusions caused

by the strained observation positions, and disturbances caused by dynamic objects (e.g., unrelated moving or temporary objects or workers) are quite common. Simultaneously, construction buildings and other structures often present symmetries and self-similarities, and the overlapping rate between different scans may be low when the acquisition time interval is large. Due to these problems, the automatic registration of point clouds on construction sites is a challenging task.

To tackle aforementioned problems, we present a 3D-to-2D coarse registration method for fast aligning point clouds of construction sites in different acquisition time, which consists of the conversion from 3D points to 2D image data and the estimation of transformation parameters of 2D images in the frequency domain. In the conversion step, the grey value of each pixel represents the number of points projected into this pixel in the vertical direction. To be specific, we utilize the geometric constraints of the verticality of building facades and maintain the main structure of the footprint of the buildings in construction site by setting the threshold for the point number projected to each pixel. Besides, using 2D images instead of 3D points can also largely improve the efficiency of the registration. Thus, 3D registration problem is changed into a 2D image matching problem. Here, we utilize a method based on transform in the frequency domain. Compared with other methods based on the original grey values or features extracted from these images, Fourier-Mellin transform is able to deal with transition, rotation, scaling and osculation with high robustness. Simultaneously, after Fourier-Mellin transform, the rotation can be fast estimated by phase correlation. Taking these factors into consideration, the core concept of our method is to transform 3D problem into 2D problem and solve it in frequency domain, so that a fast alignment

between two point clouds can be obtained and a balance of effectiveness and efficiency is also achieved.

2. RELATED WORK

For 3D registration, in general, a common registration method should solve two core issues: the representation of the 3D scans and the identification of corresponding pairs based on the representation of the 3D scans. There is a variety of researches providing solutions for registering point clouds by the use of different abstraction of the 3D points. These registration methods can be categorized into two essential classes, namely registration in the original spatial domain and registration in the feature domain.

2.1 Registration in the original spatial domain

Registration methods based on the original spatial domain mainly refer to the registration between points, and the basic idea is to find corresponding points in the original scans directly.

The Iterative Closet Point (ICP) is one of the most used methods, which tends to minimize point-to-point distances in the overlapping areas between scans. Similarly, Random sample consensus (RANSAC) methods has also been widely used, in which an iteration step is also applied by randomly selecting a set of points and calculating the transformation parameters and selecting the optimal transformation parameters under the condition that the number of inlier points reaches to the maximum when applying these parameters to the entire point clouds. The 4-point congruent sets (4PCS) and its variants (Aiger et al., 2008, Mellado et al., 2014) are also good representatives which utilize the RANSAC-based strategy. In 4PCS, a set of congruent points are selected instead of randomly selecting among original point set, which can largely increase the efficiency of the registration by decreasing the number of point candidates.

However, for the aforementioned methods, although they can be used to align different point clouds in some cases, they are quite sensitive to noises and outliers of point clouds, and it is of high possibility that these methods do not work at all when the scans are acquired in different times and has large differences in space domain. Simultaneously, they are all computation-expensive and time-consuming due to the iterative process.

2.2 Registration in the feature domain

For the registration approaches based on feature domain, instead of using points in the original scans, feature points and primitives (e.g. lines or surfaces) are extracted as geometric features for the registration. The basic concept of this registration is to extract feature points or primitives and to find the correspondences in feature domain.

Scale invariant feature transform (SIFT) (Böhm and Becker, 2007), Difference of Gaussian (DoG) (Theiler et al., 2014), Fast Point Feature Histogram (FPFH) (Rusu et al., 2009) are methods which are commonly used for the selection of key points. However, for the keypoint-based methods, although they can largely decrease the number of candidates for the correspondence searching, they share the same problem as the methods using the original space points, namely sensitivity to noises and outliers. Compared with point-based methods, the line-based, and surface-based methods are more robust since

more geometric constraints are taken into consideration, which leads to higher reliability of the features. Line-based methods refer to utilizing lines, such as straight lines (Jaw and Chuang, 2008, Habib et al., 2005), and curves (Yang and Zang, 2014) as candidates for finding the correspondence. As for surface-based methods, planal surfaces (Theiler and Schindler, 2012, Xiao et al., 2012), plane sets (Xu et al., 2017) and curved surfaces (Flöry and Hofer, 2010, Gruen and Akca, 2005) are acquired for further identification of corresponding pairs. For the line, or surface based methods, the performance of the registration methods rely heavily on the accuracy and efficiency of the extraction of these geometric primitives. For instance, the extraction of planes based on the RANSAC strategy strategy is really time-consuming, which lead to the decrease of the efficiency of the registration.

Although the aforementioned literature is not comprehensive, few researches have addressed the characteristics in transform domain. However, transform-based methods, such as, phase correlation (Nagashima et al., 2006), and Walsh transform (Lazaridis and Petrou, 2006) are quietly common in 2D image processing and are methods of high robustness. In our work, we propose a method in which the 3D scans are projected into the ground plane and the transformation parameters are obtained by using phase correlation in transform plane. To this end, little work has been done for coarse registration using projection based approach. The scientific contributions of this paper are as follows:

- A method which provides the 2D representation of the 3D scene is proposed: Unlike commonly used registration method in a 3D scene, we conduct the registration in the projection plane.

- Based on the 2D projection, registration is conducted in the transform domain.

- A novel strategy using 2D projection based phase correlation between scans is developed, which overcomes the disadvantages of conventional 3D-based registration and improves the processing efficiency simultaneously.

3. METHODOLOGY

The implementation of our proposed method consists of two core steps: the conversion from 3D points to 2D image data and the estimation of transformation parameters of 2D images in the frequency domain. In the conversion step, the point cloud of each scan is projected into a 2D grey image, by which the ground footprint of the point cloud is obtained. In the following step, we represent the 2D image in frequency-domain and estimate the horizontal transformation parameters by using Fourier-Mellin transformation. The processing workflow is sketched in Fig.1. The detailed explanation on each step will be introduced in the following sections.

3.1 Projecting 3D scenes to 2D

The generation of building footprint is a vertical projection of 3D points, in which 3D points are converted into 2D grey image data by projecting points to the ground plane along the vertical direction. Thus, a projection image is obtained, in which grey values represent the numbers of points mapped to the grid of each pixel. For the vertical structure, we assume that the points belong to the vertical structure falls into the aligned pixels or
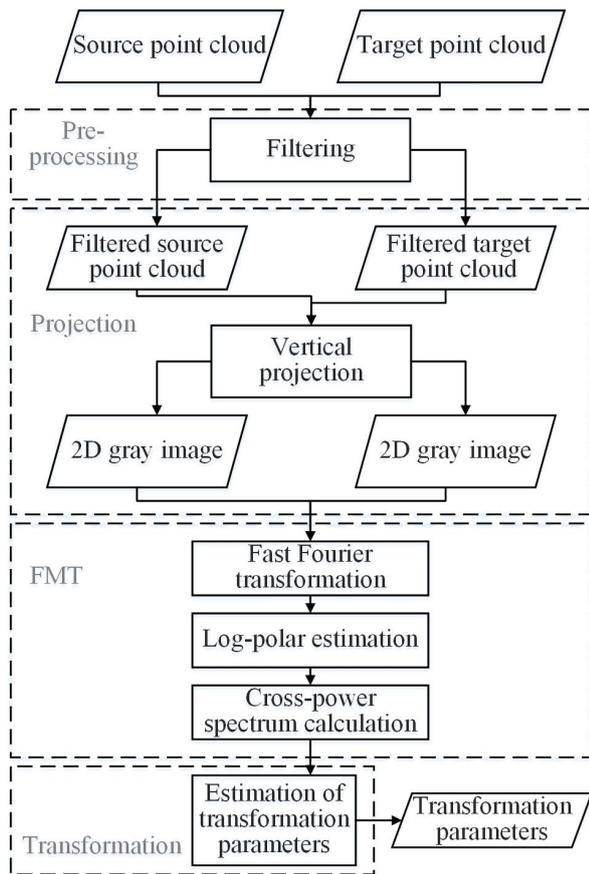
Figure 1. Workflow of the proposed fast coarse registration strategy.

neighboring pixels, which indicates that the bright lines in 2D projection refer to the vertical structure in 3D scans. Thus we apply a filter to eliminate the noises caused by noises, outliers or irrelevant objects to the main building structures in 3D scans in the projection image by setting a threshold to the grey values of the pixels.

3.2 2D image matching using FMT

3.2.1 FMT Fourier space image registration methods provide a way to recover all rigid transformation parameters, i.e. rotation, translation and scale. They differ from other registration methods in that they search for the optimal match in the frequency domain. These methods make use of the Fourier Shift Theorem and the Fourier Rotation Theorem to provide invariance to rotation, translation and scale. Image registration is then performed by phase correlation of the cross-power spectra. It is possible to compute the Fourier Transform of an image efficiently by using the Fast Fourier Transform (FFT).

Let $F_1(\epsilon, \lambda)$ and $F_2(\epsilon, \lambda)$ be the Fourier transforms corresponding to images $I_1(u, v)$ and $I_2(u, v)$. If I_1 and I_2 are related by a rotation θ_0 and translation $(\Delta u, \Delta v)$, both in the image plane, their Fourier transforms are related by

$$F_2(\epsilon, \lambda) = -e^{j2\pi(\epsilon\Delta u + \lambda\Delta v)} \times$$
$$F_1(\epsilon cos(\theta_0) + \lambda sin(\theta_0), -\epsilon sin(\theta_0) + \lambda cos(\theta_0)) \quad (1)$$

Let M_1 and M_2 be the magnitudes of F_1 and F_2, respectively,

which are related by

$$M_2(\epsilon, \lambda) = M_1(\epsilon cos(\theta) + \lambda sin(\theta), -\epsilon sin(\theta) + \lambda cos(\theta)) \quad (2)$$

To recover both rotation and scale simultaneously, the Fourier magnitude spectra are transformed to a log-polar representation (ρ, θ), which enables the representation of rotation as a shift. M_1 and M_2 are then related by

$$MP_2(\rho, \theta) = MP_1(\rho/s, \theta - \theta_0) \quad (3)$$

where s is the scaling factor, $\epsilon = \log(\rho)$ and $\lambda = \log(s)$.

The normalized cross-power spectrum is then defined as

$$\widehat{Corr}(\epsilon, \lambda) = \frac{F_1(\epsilon, \lambda)}{F_2(\epsilon, \lambda)} = \frac{F_1(\epsilon cos(\theta) F_2'^*(\epsilon, \lambda)}{|F_1(\epsilon cos(\theta) F_2'^*(\epsilon, \lambda)|}$$
$$= e^{j2\pi(\epsilon\Delta u + \lambda\Delta v)} \quad (4)$$

where F^* is the complex conjugate of F. Taking the inverse Fourier transform

$$Corr(u, v) = F^{-1}(\widehat{Corr}(\epsilon, \lambda)) = \sigma(u - \Delta u, v - \Delta v) \quad (5)$$

which means that $Corr(u, v)$ is nonzero only at

$$(\Delta u, \Delta v) = \arg\max_{(u,v)} \{Corr(u, v)\} \quad (6)$$

In FMT, the estimation of rotation, scale and transition are all turned into a phase correlation problem.

3.2.2 2D image registration In order to perform an image registration algorithm, the FMT has been chosen. The FMT is a global method that takes the contributions from all points in the images into account in order to provide a way to recover all rigid transformation parameters, i.e. rotation, translation. It is an efficient and accurate method to process a couple of images that are fairly similar. The steps of the grey image registration algorithm are described in Alg.

Algorithm 1: Steps of the FMT algorithm for the estimation of transformation parameters

---

Input: Images I_1, and I_2
Output: Rotation θ, scale s, transition $(\Delta u, \Delta v)$
1: Apply thresholding filter to eliminate noises of images.
2: Apply FFT to images $I_1 \to F_1$ and $I_2 \to F_2$
3: Compute the magnitudes $M_1 = |F_1|$, $M_2 = |F_2|$.
4: Transform the resulting values from rectangular to polar coordinates. $M() \to MP()$.
5: Apply the FFT to polar images, a bilinear interpolation is used. $MP() \to \widehat{MP}()$.
6: Compute $\widehat{Corr}(\epsilon_\rho, \lambda_\theta)$ between $\widehat{MP}_1(\epsilon_\rho, \lambda_\theta)$ and $\widehat{MP}_2(\epsilon_\rho, \lambda_\theta)$ using Eq. 4.
7: Compute the inverse FFT $Corr(\rho, \theta)$ of $\widehat{Corr}(\epsilon_\rho, \lambda_\theta)$
8: Find the location of the maximum of $Corr(\rho, \theta)$ and obtain θ and s.
9: Construct a new image I_r by applying reverse rotation and scaling to I_2.
10: Apply FFT to image I_r.
11: Compute the correlation $\widehat{Corr}(\epsilon_{\Delta u}, \lambda_{\Delta v})$ using Eq. 4.
12: Take inverse FFT $Corr(u, v)$ of $\widehat{Corr}(\epsilon_{\Delta u}, \lambda_{\Delta v})$.
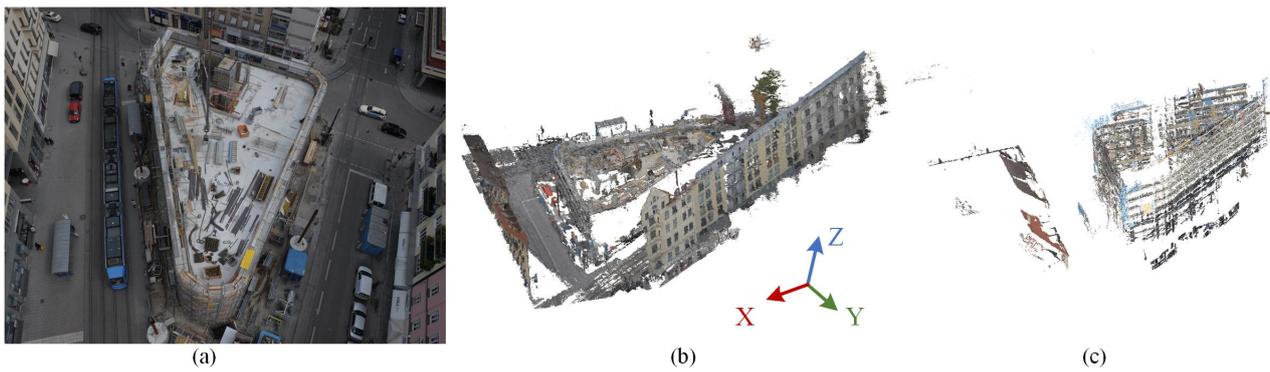13: Obtain the values $(\Delta u, \Delta v)$ of the shift.

(a)   (b)   (c)

Figure 2. Experimental data. (a) Image scene of the construction site. (b) Source and (c) target scans of a construction site.

(a)   (b)   (c)
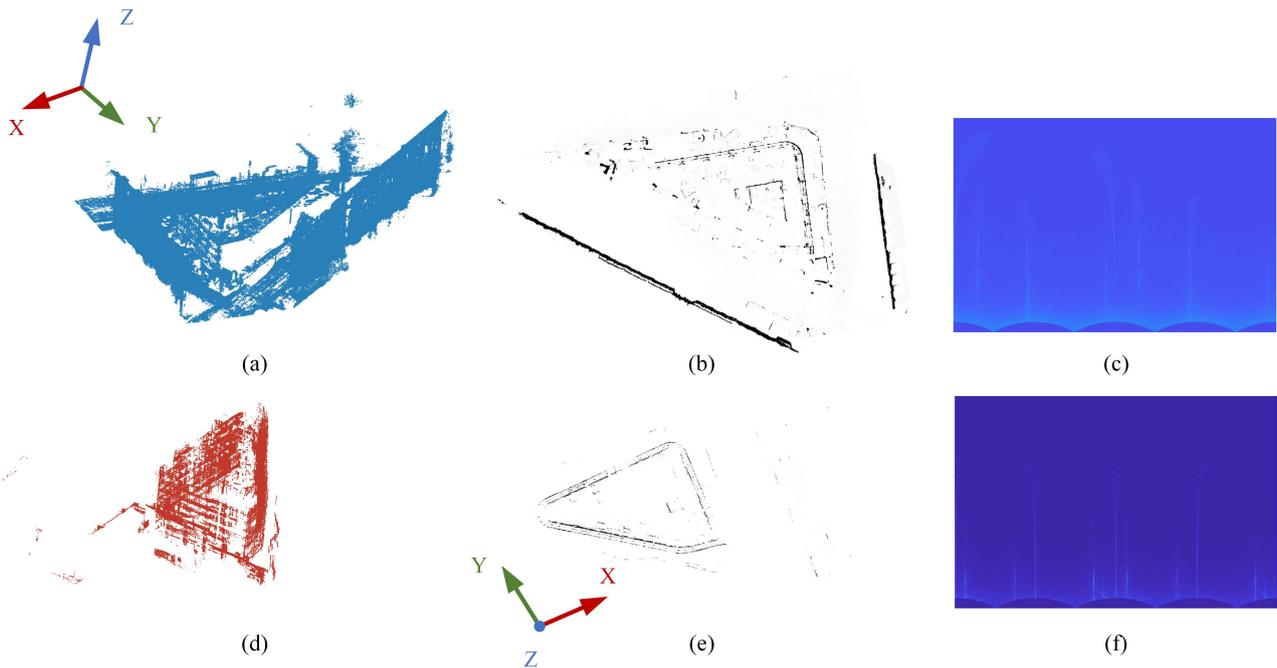
(d)   (e)   (f)

Figure 3. Registration results of two scans of the construction site. (a) and (d) Original scans. (b) and (e) 2D projection images. (c) and (f) Magnitude images of FFT in log-polar space. The above figures are results from the target scan, and the below figures are from the source scan.

4. EXPERIMENTS

4.1 Experimental datasets

The proposed method for performing registration is tested using photogrammetric point cloud datasets, generated from a structure from motion (SfM) system and multi-view stereo matching method. The images are acquired by a terrestrial, and hand-held SLR camera with the focal length 24mm and the sensor size 36×24mm(4256×2832 pixels). Test Site is an inner-city construction site with limited surrounding space in Munich. Streets and buildings enclose all sides of the site. The Z axis in the coordinate system of the point cloud is perpendicular to the earth ground plane. The scans used were acquired on June 27, 2013, and September 4, 2013, respectively.

4.2 Registration results

The coarse registration results of our dataset are illustrated in Fig.3 and Fig.4. In order to better display the performance of

Point Cloud	Size (number of points)	
	Source	Target
Original	6413022	889686
Filtered	6175127	874490

Table 1. Information of experimental datasets

our proposed method, we manually rotate the source scan, which is well aligned with the target scan with manual registration, along Z-axis for 30 degrees counterclockwise. In the step of 2D projection, the threshold for grey value is set to 228. As illustrated in Fig.4, the source scan is well aligned to the target scan. When it comes to the quantitative results of our experiments, the registration error is less than 1 degree, and the computation time is less than 20s, but the time rely heavily on the number of points in the scans in the projection step. For the proposed method, it should be noted that the performance of this method is highly influenced by the threshold which is set for
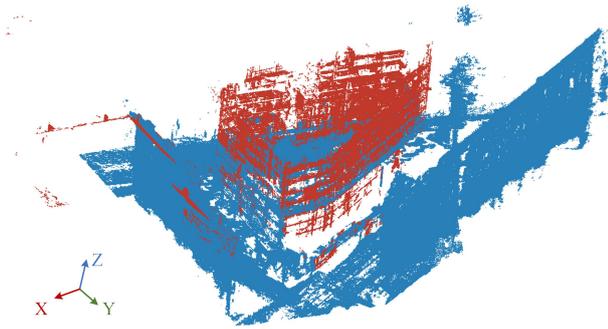
Figure 4. Registration results of two scans of the construction site.

eliminating the effect of noises and outliers.

5. CONCLUSION

In this work, we present an automatic marker-free method for fast and coarse registration between point clouds by converting the 3D matching problem into a 2D correlation problem, taking the special properties of building structures into consideration. Our proposed method consists of the conversion from 3D points to 2D image data and the estimation of transformation parameters between 2D images in the frequency domain. The projection plays like the reduction of dimensionality of the data. Disturbance and small structural elements are removed during the projection process. Thus, in the generated 2D images, only essential structures of the as-built building are preserved, representing the ground footprint. Then, the correlation between 2D images in frequency-domain estimate the horizontal transformation parameters by using Fourier-Mellin transformation. In the future, we will try to apply the idea of this dimensionality reduction to the registration of point cloud, and conduct thorough experiments to fully analyze the potential of phase correlation for fast alignment between point clouds.

ACKNOWLEDGEMENTS

REFERENCES

Aiger, D., Mitra, N. J. and Cohen-Or, D., 2008. 4-points congruent sets for robust pairwise surface registration. In: *ACM Transactions on Graphics (TOG)*, Vol. 27number 3, ACM, p. 85.

Böhm, J. and Becker, S., 2007. Automatic marker-free registration of terrestrial laser scans using reflectance. In: *Proceedings of the Proceedings of 8th Conference on Optical 3D Measurement Techniques, Zurich, Switzerland*, pp. 9–12.

Bosché, F., Ahmed, M., Turkan, Y., Haas, C. and Haas, R., 2015. The value of integrating scan-to-bim and scan-vs-bim techniques for construction monitoring using laser scanning and bim: The case of cylindrical mep components. *Automation in Construction* 49, pp. 201–213.

Flöry, S. and Hofer, M., 2010. Surface fitting and registration of point clouds using approximations of the unsigned distance function. *Computer Aided Geometric Design* 27(1), pp. 60–77.

Gruen, A. and Akca, D., 2005. Least squares 3d surface and curve matching. *ISPRS Journal of Photogrammetry and Remote Sensing* 59(3), pp. 151–174.

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.

Jaw, J. J. and Chuang, T. Y., 2008. Registration of ground-based lidar point clouds by means of 3d line features. *Journal of the Chinese Institute of Engineers* 31(6), pp. 1031–1045.

Lazaridis, G. and Petrou, M., 2006. Image registration using the walsh transform. *IEEE Transactions on Image Processing* 15(8), pp. 2343–2357.

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.

Nagashima, S., Aoki, T., Higuchi, T. and Kobayashi, K., 2006. A subpixel image matching technique using phase-only correlation. In: *Intelligent Signal Processing and Communications, 2006. ISPACS'06. International Symposium on*, IEEE, pp. 701–704.

Rusu, R. B., Blodow, N. and Beetz, M., 2009. Fast point feature histograms (fpfh) for 3d registration. In: *Robotics and Automation, 2009. ICRA'09. IEEE International Conference on*, Citeseer, pp. 3212–3217.

Seidell III., F. M., 2002. Management of construction projects: A constructor's perspective. *Cost Engineering* 44(12), pp. 36.

Theiler, P. and Schindler, K., 2012. Automatic registration of terrestrial laser scanner point clouds using natural planar surfaces. *ISPRS Annals of 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.

Tuttas, S., Braun, A., Borrmann, A. and Stilla, U., 2014. Comparison of photogrammetric point clouds with bim building elements for construction progress monitoring. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*.

Tuttas, S., Braun, A., Borrmann, A. and Stilla, U., 2017. Acquisition and consecutive registration of photogrammetric point clouds for construction progress monitoring using a 4d bim. *PFG–Journal of Photogrammetry, Remote Sensing and Geoinformation Science* 85(1), pp. 3–15.

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

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 Photogrammetry, Remote Sensing & Spatial Information Sciences*.

Xu, Y., Tuttas, S., Hoegner, L. and Stilla, U., 2018. Reconstruction of scaffolds from a photogrammetric point cloud of construction sites using a novel 3d local feature descriptor. *Automation in Construction* 85, pp. 76–95.

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.