Iterative Global Similarity Points : A robust coarse-to-fine integration solution for pairwise 3D point cloud registration

Yue Pan, Bisheng Yang, Fuxun Liang and Zhen Dong
Wuhan University
Wuhan, China, 430079
{panyue, bshyang, liangfuxun, dongzhenwhu}@whu.edu.cn

Abstract

In this paper, we propose a coarse-to-fine integration solution inspired by the classical ICP algorithm, to pairwise 3D point cloud registration with two improvements of hybrid metric spaces (e.g., BSC feature and Euclidean geometry spaces) and globally optimal correspondences matching. First, we detect the keypoints of point clouds and use the Binary Shape Context (BSC)[12] descriptor to encode their local features. Then, we formulate the correspondence matching task as an energy function, which models the global similarity of keypoints on the hybrid spaces of BSC feature and Euclidean geometry. Next, we estimate the globally optimal correspondences through optimizing the energy function by the Kuhn-Munkres[39] algorithm and then calculate the transformation based on the correspondences. Finally, we iteratively refine the transformation between two point clouds by conducting optimal correspondences matching and transformation calculation in a mutually reinforcing manner, to achieve the coarse-to-fine registration under an unified framework. The proposed method is evaluated and compared to several state-of-the-art methods on selected challenging datasets with repetitive, symmetric and incomplete structures. Comprehensive experiments demonstrate that the proposed IGSP algorithm obtains good performance and outperforms the state-of-the-art methods in terms of both rotation and translation errors.

1. Introduction

As the development of laser scanning and photogrammetry, point cloud, which depicts the world in 3D manner, has been widely collected by many platforms like airborne laser scanning (ALS) and terrestrial laser scanning (TLS). Point cloud registration is a fundamental problem in 3D computer vision and photogrammetry. Given several sets of points in different coordinate systems, the aim of registration is to find the transformation that best aligns all of them into a common coordinate system. Point cloud registration plays a significant role in many vision applications such as 3D model reconstruction [8][40], cultural heritage management[28][42], landslide monitoring[38] and solar energy analysis[19].In 3D object recognition, fitness degree between an existing model object and an extracted object in the scene can be evaluated with registration results [6][21].In the field of robotics, for simultaneous localization and mapping(SLAM), the registration can act as a visual odometry to realize structure from motion and locate the current view into the global scene[4][34].

Generally, there are four key challenges for unordered point clouds registration: (1) uneven point densities , (2) the huge amount of data , (3) repetitive, symmetric, and incomplete structures , and (4) limited overlaps between point clouds . All of these challenges can seriously affect the performance of point cloud registration methods. To address these challenges, extensive studies have been done to improve the accuracy, efficiency, and robustness of point cloud registration. Point cloud registration can be roughly categorized into pairwise and multi-view registration according to the number of input point clouds[20]. The pairwise registration is the prerequisite of multi-view registration, which is the focus of this paper. The registration process can be further divided into two major steps: coarse registration, in which an initial transformation between two point clouds is estimated, and fine registration, in which the initial transformation is then further refined[17].

The remainder of this paper is organized as follows. Following this introduction, Section 2 briefly reviews the representative work related to pairwise point cloud registration and introduces the paper's contribution. Section 3 gives a detailed description of the proposed pairwise point cloud registration method. The proposed method is validated in experimental studies in Section 4. Finally, the conclusions and future research directions are presented in Section 5.

2. Related work

2.1. Coarse Registration.

Normally, a pairwise coarse registration method has the following procedures. First, key primitive elements (*e.g.*, points, lines, and planes) are detected from each point cloud, where the point-based methods are more popular due to their feasibility to different scenes[5]. The keypoint detectors (*e.g.*, local surface patches[9], 2.5D SIFT[25], 3D SURF[24] and 3D Harris[33]) are exploited to extract keypoints from raw point clouds. Second, the feature descriptors (*e.g.*, Spin image[22], Fast Point Feature Histograms (FPFH)[31], Viewpoint Feature Histogram (VFH)[32] and Rotational Projection Statistics (RoPS)[18]) are calculated to encode the local shape information of each keypoint. Recently, [13] applies a deep neural network auto-encoder to realize the same effect. Third, various feature matching strategies are applied to determine the initial correspondences (*e.g.*, reciprocal correspondence[29], correlation coefficient[23], and chi-square test[44]). Finally, due to the fact that some of the obtained correspondences are incorrect, outliers need to be filtered from the initial correspondences, calling for a method to eliminate outliers and calculate the transformation between point clouds based on the remaining correspondences. This can be solved by some robust transformation estimation algorithms (*e.g.*, Random Sample Consensus (RANSAC)[14] , 3D Hough voting[37], geometric consistency constraints based on line distance or triangle property[36][41], and Game Theory based matching algorithms [2][43].)

Additionally, there are some point-based methods which do not follow the abovementioned workflow. For example, the 4-Points Congruent Sets (4PCS)[1] and its variants (*e.g.*, SUPER-4PCS[27], Keypoint-based 4PCS (K-4PCS)[35])determines the corresponding four-point base sets by taking intersection ratios of these four points instead of using feature descriptor for matching, thus improving the efficiency of RANSAC based global registration to a great extent.

Though most of the coarse registration algorithms can generally provide satisfactory registration results, they still have limitations. For symmetric and large-scale point clouds, a robust and efficient correspondence matching algorithm, which is capable of acquiring the globally optimal correspondences, is urgently needed.

2.2. Fine Registration.

As for fine registration, the Iterative Closest Point (ICP) algorithm [7] and its variants[3][16] are the most commonly used methods, which alternate between correspondence matching and transformation calculation until convergence. [3] proposed a Geometric Primitive ICP with Random sample consensus (GPICPR) in which the local surface normal vector and geometric curvature are used for matching and neighborhood searching.[16] used geometric features to improve the classical ICP algorithm. ICP and these variants are able to acquire registration results with high efficiency and accuracy. However, they require a good initialization to avoid converging to bad local minimum.

2.3. Contribution

To overcome the limitations and challenges, this paper proposes an Iterative Global Similarity Points (IGSP) algorithm to realize a coarse-to-fine integration solution to pairwise 3D point cloud registration. IGSP is inspired by the classical ICP algorithm with two improvements of hybrid metric spaces (*e.g.*, BSC feature and Euclidean geometry spaces) and globally optimal correspondences matching.

Specifically, the main contributions of the proposed method are as follows: i) we formulate the correspondence matching task as an energy function, which models the global similarity of keypoints on the hybrid metric spaces of BSC feature and Euclidean geometry, to get a more robust result . ii).we realize a coarse-to-fine registration by conducting optimal correspondences matching and transformation calculation in an iterative and mutually reinforcing manner, so that a good initialization is not essential.

3. Method

3.1. IGSP framework

An overview of the proposed IGSP algorithm is shown in Fig.1. Similar to ICP, an iteration process is also involved. IGSP iteratively conducts globally optimal correspondences matching and transformation calculation, until the transformation is negligible. The workflow of IGSP is shown in Fig.1a. First, we detect the keypoints of point clouds and generate the Binary Shape Context (BSC)[12] descriptors to encode their local features. Then, we formulate the correspondence matching task as an energy function, which models the global similarity of keypoints on the hybrid metric space of BSC feature and Euclidean geometry, as shown in Fig.1d-e. Next, we determine the globally optimal correspondences by optimizing the energy function by the Kuhn-Munkres(KM) algorithm[39] and then calculate the transformation based on them, as shown in Fig.1f-g. Finally, we refine the transformation between two point clouds by iteratively conducting optimal correspondences matching and transformation calculation to realize a coarse-to-fine registration under an unified framework, as shown in Fig.1h-i. The pseudo code of IGSP is shown in Algorithm 1.

3.2. Energy function construction

In this section, a global energy function for keypoints matching is constructed, as Eq.1, which consists of data cost and penalty cost. Data cost represents the global similar-
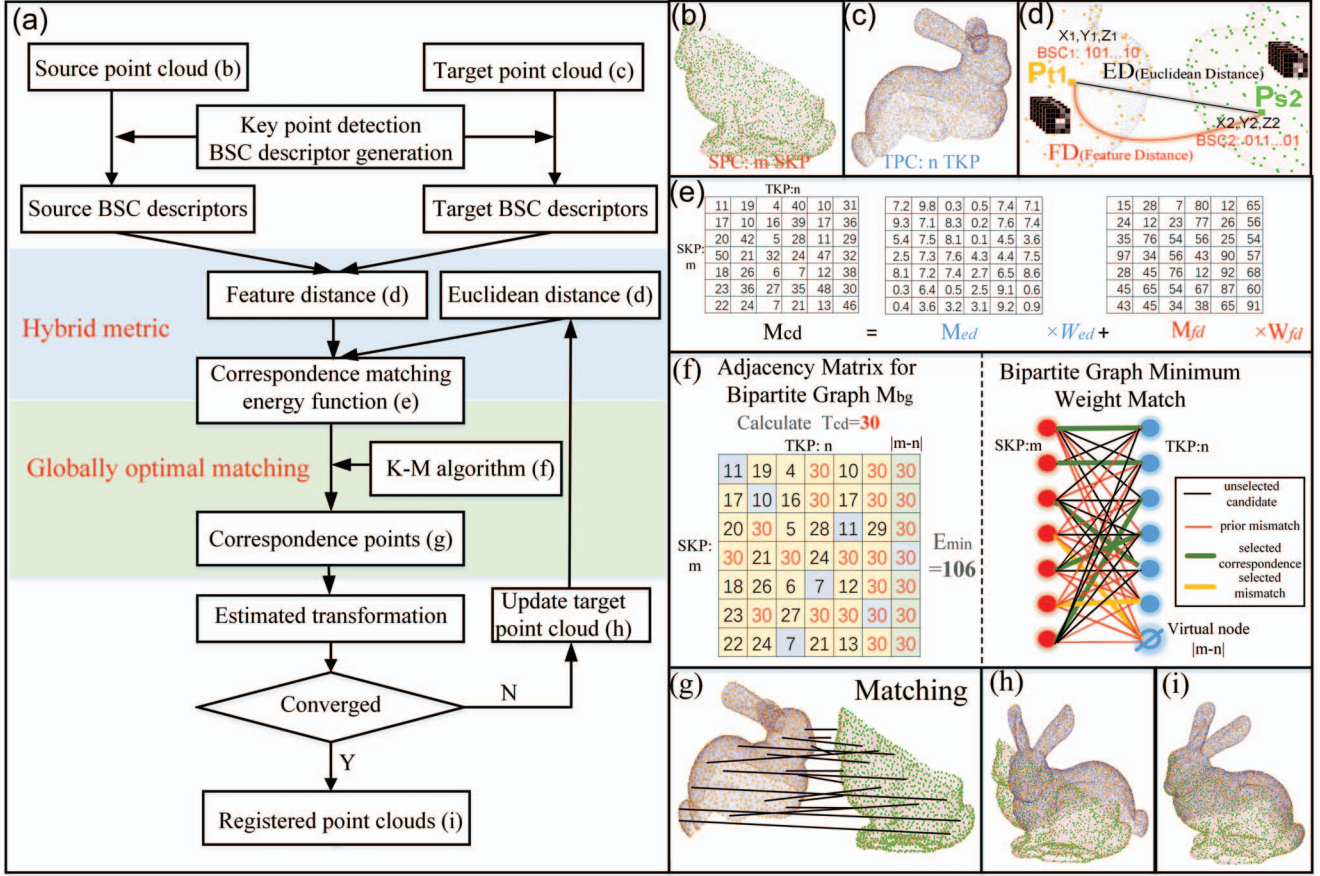
Figure 1. Overview of the proposed IGSP algorithm framework:(a)workflow, (b) the input Source Point Cloud (SPC, in red) and detected Source Keypoints (SKP, in green), (c) the input Target Point Cloud (TPC, in purple) and detected Target Keypoints (TKP, in orange), (d) a sketch map for the Euclidean Distance (ED) and Feature Distance (FD) (e) the pairwise Compound Distance (CD) and the corresponding ED matrix, FD matrix, (f) Left:an example of energy function optimization by KM algorithm. Numbers in grids represent the compound distance between keypoints. The mismatch threshold T_{cd} (in red) is firstly calculated (30), and a virtual column (the rightmost column) is inserted. Red numbers are the prior mismatches. The numbers in blue grids are used to calculate the result (106), which is considered as the minimum matching energy $E_{\min}$. Right:the corresponding sketch map of the example. SKP and TKP are shown in red and blue respectively. The virtual node is represented by Φ. Four different type of lines are illustrated (black: unselected candidate, red: prior mismatch, green: selected correspondence, orange: selected mismatch) (g) the correspondences solved by KM algorithm, (h) the updated result after several iterations, (i) final registered result.

ity of matched keypoints in source and target point clouds, while penalty cost indicates the number of keypoints without correspondence. For data cost, p and q represent a matching keypoint pair. M is the set of correspondences between two point clouds and $CD(p,q)$ means the compound distance between keypoint p and q ,which is defined as a weighted sum of feature distance $FD(p,q)$ and Euclidean distance $ED(p,q)$, as shown in Eq.2. W_{ed} and W_{fd} are the weights of ED and FD. As shown in Eq.3, $FD(p,q)$ is defined as the Hamming distance between the BSC descriptors of p and q . $ED(p,q)$ is calculated as a scale factor related to point density s_{ed} times the Euclidean distance between p and q , as shown in Eq.4. W_{ed} and W_{fd} are as-

signed as Eq.5, in which k is the iteration number counted from zero and m controls the weight changing rate. For penalty cost in Eq.1, φ is the set of unmatched keypoints in two point clouds and $|\varphi|$ is the number . W_p is the weight of penalty cost, which will be further used as the criterion for mismatch judgement in 3.3.

$$E = \underbrace{\sum_{p \in S, q \in T, (p,q) \in M} CD(p,q)}_{Data\_cost} + \underbrace{W_p |\varphi|}_{Penalty\_cost} \quad (1)$$

$$CD(p,q) = W_{fd}FD(p,q) + W_{ed}ED(p,q) \quad (2)$$

182

Algorithm 1 Iterative Global Similarity Points (IGSP)

Notation:

SPC, TPC :Source and Target Point Clouds

$SKP(sk_1, sk_2, ...sk_m), TKP(tk_1, tk_2, ...tk_n)$:Source and Target Keypoint sets

$SBSC(sb_1, sb_2, ...sb_m), TBSC(tb_1, tb_2, ...tb_n)$:BSC features extracted from SKP and TKP

$M_{ed}, M_{fd}, M_{cd}, M_{bg}$:Matrix of ED, FD, CD and the bipartite graph (ED, FD and CD represents Euclidean, feature and compound distance respectively)

σ_r, σ_t :convergence threshold for translation and rotation variation

M :the set of correspondences in SKP and TKP

Rt_{temp} : transformation estimated at the current iteration

Rt : transformation accumulated from the first to the current iteration

k :iteration number

Input: SPC and TPC

Output: Rt

1: Detect m keypoints $SKP(sk_1, sk_2, ...sk_m)$ from SPC and n keypoints $TKP(tk_1, tk_2, ...tk_n)$ from TPC
2: Extract BSC feature $SBSC(sb_1, sb_2, ...sb_m)$ of $SKP(sk_1, sk_2, ...sk_m)$ and $TBSC(tb_1, tb_2, ...tb_n)$ of $TKP(tk_1, tk_2, ...tk_n)$
3: Calculate the $m \times n$ matrix M_{fd} using $SBSC(sb_1, sb_2, ...sb_m)$ and $TBSC(tb_1, tb_2, ...tb_n)$, as Eq.3
4: Initialization:$k \leftarrow 0, Rt \leftarrow I, \Delta r \leftarrow \infty, \Delta t \leftarrow \infty$
5: **while** $\Delta r > \sigma_r \wedge \Delta t > \sigma_t$ **do**
6:   Calculate the $m \times n$ matrix M_{ed} using $SKP(sk_1, sk_2, ...sk_m)$ and $TKP(tk_1, tk_2, ...tk_n)$, as Eq.4
7:   Calculate W_{ed} and W_{fd}, as Eq.5
8:   $M_{cd} \leftarrow W_{ed}M_{ed} + W_{fd}M_{fd}$
9:   Calculate T_{cd}, as Eq.10
10:   Calculate $\max(m,n) \times \max(m,n)$ matrix M_{bg} using M_{cd} and T_{cd}, as Eq.7
11:   $M = KM(M_{bg}, T_{cd})$, KM represents Kuhn-Munkres algorithm, which takes the bipartite graph adjacency matrix and the mismatch threshold as input and outputs the matching set.
12:   Calculate Rt_{temp} from M using SVD, as Eq.9
13:   $TKP \leftarrow Rt_{temp}TKP$
14:   Calculate Δr and Δt from Rt_{temp}
15:   $Rt \leftarrow Rt_{temp}Rt, k \leftarrow k+1$
16: **end while**

$$FD(p,q) = HD(f_p, f_q) \tag{3}$$

$$ED(p,q) = s_{ed}\|p-q\| \tag{4}$$

$$\begin{cases} W_{fd} = e^{-\frac{k}{m}} \\ W_{ed} = 1 - e^{-\frac{k}{m}} \end{cases} \quad k = 0,1,2... \tag{5}$$

By minimizing the energy function, we can obtain the globally optimal correspondences $\{M, \varphi\}^*$, as shown in Eq.6.

$$\{M, \varphi\}^* = \underset{\{M,\varphi\}}{arg\min}E = \underbrace{W_p|\varphi|}_{penalty\_cost} + \quad \text{why?}$$
$$\underbrace{\sum_{p\in S, q\in T, \{p,q\}\in M} \left(\left(1 - e^{-\frac{k}{m}}\right)ED(p,q) + e^{-\frac{k}{m}}FD(p,q)\right)}_{Data\_cost} \tag{6}$$

3.3. Energy function optimization by KM algorithm 二分图

In this section, the global energy function Eq.6 is generalized into a bipartite graph minimum weight match problem and then solved by using slacked KM algorithm.

A bipartite graph is a graph whose vertices can be divided into two disjoint sets S and T[15]. Each edge connects a vertex in S to a vertex in T in this graph. Given a bipartite graph and its corresponding edges, the optimal weight matching guarantees that each node in one subgraph can be matched to only one node in the other subgraph, and this matching can achieve global matches with the maximal or minimum summation of edge weight.

The task to find the best matching for two keypoint sets can be modeled as an optimal matching task of bipartite graph. In the weighted bipartite graph $G = (S, T, E)$, each keypoint in source and target cloud is represented by one node respectively in set S and T. Suppose there are m and n keypoints detected from source and target cloud respectively, and when $m \neq n$, we add $|m-n|$ virtual unmatched nodes N_v to the set with less nodes to make $|S| = |T| = \max(m,n)$. Each edge $e(p,q) \in E$ corresponds to a distance between node p in S and node q in T. With T_{cd} as the mismatch threshold for CD, $e(p,q)$ is computed as follows:

$$e(p,q) = \begin{cases} CD(p,q), & CD(p,q) < T_{cd} \wedge p \notin N_v \wedge q \notin N_v \\ T_{cd} & , else \end{cases} \tag{7}$$

Next, the sum of all edge weights can be minimized as $E_{\min}^{bgm}$ in Eq.6. When $T_{cd} = 2W_p$, there is only a constant difference between $E_{\min}^{bgm}$ and the minimum energy E_{min} for the energy function, so that the minimum weight match is equivalent to the optimization of energy function. The selected edges whose weight $e^*(p,q) < T_{cd}$ make up the optimal matching set M^*, and the unmatched keypoints set

183

φ^* is also determined.

$$
\begin{aligned}
E_{\min}^{bgm} &= \sum e^* (p,q) + T_{cd} (\max(m,n) - |M^*|) \\
&= \sum e^* (p,q) + W_p (m+n+|m-n|-2|M^*|) \\
&= \sum_{p \in S, q \in T, \{p,q\} \in M^*} CD(p,q) + W_p |\varphi| + W_p |m-n| \\
&= E_{\min} + W_p |m-n|
\end{aligned}
\tag{8}
$$

Given such weighted bipartite graph $G = (S,T,E)$, the Kuhn-Munkres(KM) algorithm is employed. It outputs a complete bipartite matching with minimum matching weight by transforming the problem from an optimization problem of <u>finding a minimum weight matching into a combinatorial one of finding a perfect matching</u>. For efficiency concern, we apply the KM algorithm with <u>slacked</u> terms whose time complexity is $O\left(|V|^3\right)$, in which $|V|$ is the number of vertexes of the graph. Fig.1 illustrates a simple example of this process.

It is proved that given a point cloud pair with m and n keypoints respectively and a fixed threshold W_p or T_{cd}, E_{min} can be calculated by solving $E_{\min}^{bgm}$ via KM algorithm.

Once the correspondence is determined, the optimal transformation can be estimated via Singular Value Decomposition(SVD), as shown in Eq.9 .

$$
\{R|t\}^* = \arg\min_{\{R,t\}} J = \sum_{p \in S, q \in T, \{p,q\} \in M^*} \|p - Rq + t\|^2
\tag{9}
$$

3.4. Iteration Process

Although a correspondence set $\{M, \varphi\}$ with global similarity can be solved efficiently, it is not robust enough. Inspired by ICP, we propose an <u>iteration</u> process by alternating between <u>solving for the correspondence</u> using the method presented above and <u>estimating the transformation using SVD</u> until convergence. For each iteration, W_{ed} and W_{fd} are updated according to Eq.5. In the iteration process, W_{fd} decreases from 1 to 0, while W_{ed} increases from 0 to 1. At first, $W_{ed} = 0$ and the correspondence is estimated only based on keypoints feature descriptor, thus producing a coarse registration result. As the process continues, the weight of Euclidean distance increases, introducing the geometric restriction and refining the registration over and over.

As mentioned before, W_p (or T_{cd}) is the weight of penalty cost used as the criterion for mismatch judgement. In the context of bipartite graph minimum weight match, candidate pairs whose $CD > T_{cd}$ would be regarded as mismatches. A self-adaptive scheme for the determination of this subtle parameter is devised, as shown in Eq.10.
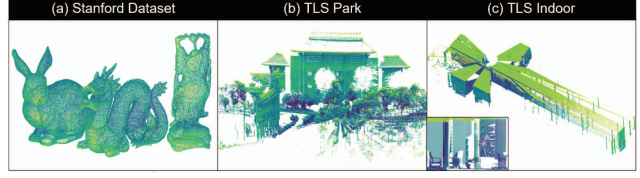

Figure 2. The sampled point clouds from three datasets, rendered by ambient occlusion. (for indoor datasets, some details are also shown)

$$
T_{cd} = \begin{cases} \mu_{cd} + p_1^t \sigma_{cd} & , k = 0 \\ p_2^t \left(1 - e^{-\frac{k}{m}}\right) \overline{ED_{k-1}^c} + p_3^t e^{-\frac{k}{m}} \overline{FD_{k-1}^c} & , k > 0 \end{cases}
\tag{10}
$$

When $k = 0$, given a pairwise match matrix of CD, the mean value μ_{cd} and standard deviation σ_{cd} can be calculated. With a threshold parameter p_1^t, we get the initial T_{cd}. When $k > 0$, $\overline{ED_{k-1}^c}$ and $\overline{FD_{k-1}^c}$ are the average Euclidean and feature distance of matched keypoints pairs of last iteration. Besides, p_2^t and p_3^t are two threshold parameter to be determined. The search space of candidate match keeps narrowing under the restriction of the average geometric and feature similarity of last iterations correspondence set. Generally, the larger parameters $p_i^t \, (i = 1,2,3)$ are, the larger the search space of candidate match is.

Finally, as for the condition of convergence, we set thresholds for both translation and rotation in practice. Once the transformation difference between two iterations meets the condition, the iterative process stops.

4. Experiment

4.1. Experimental setup

Experiment Platform. The experiments are implemented with a 16 GB RAM and an Intel Core i7-6700HQ @ 2.60GHz CPU. The proposed IGSP algorithm and the compared baselines are all implemented in C++ with the help of point cloud library (PCL) [30].

Datasets description. As shown in Fig.2, the performance of the proposed IGSP algorithm is evaluated both on small scale scans of 3D models (Stanford Repository[10]) and on challenging real-world TLS datasets (*e.g.* Park, Indoor). These TLS datasets are challenging as (i) each of the datasets contains several scans and billions of points (ii) repetitive, symmetric and incomplete structures are common.

Evaluation criteria. We evaluate the performance of the proposed IGSP method in terms of rotation error and translation error which are commonly used for the evaluation of point cloud registration[18]. Given the estimated transform T and ground truth transform T^G, rotation error e^r and
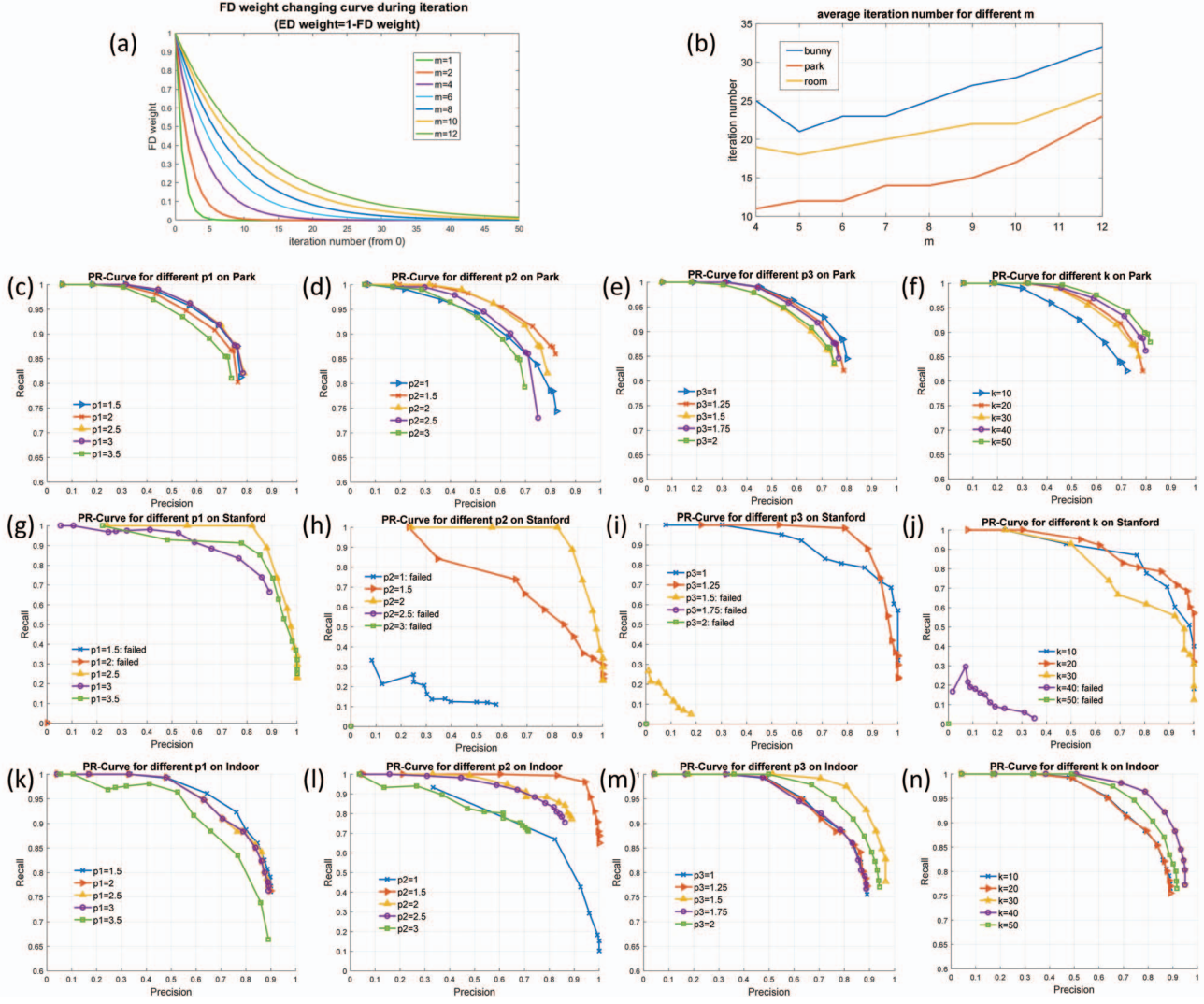
184

Figure 3. Parameter settings experiment result: (a-b) m's effect on ED, FD's weight and total iteration number for IGSP. (c-n) precision and recall performance of different parameter settings for IGSP's correspondence grouping on Park, Standford and Indoor datasets respectively.

translation error e^t can be calculated as Eq.11,12.

$$\Delta T = T \left(T^G\right)^{-1} = \begin{bmatrix} \Delta R & \Delta t \\ 0 & 1 \end{bmatrix} \quad (11)$$

$$\begin{cases} e^r = \arccos\left(\frac{tr(\Delta R)-1}{2}\right) \\ e^t = \|\Delta t\| \end{cases} \quad (12)$$

Besides, the keypoint correspondences quality can be evaluated with precision and recall as Eq.13, where TP is the number of true positive correspondences, FP is the number of false positive ones, and FN is the number of false negative ones.

$$\begin{cases} precision = \frac{TP}{TP+FP} \\ recall = \frac{TP}{TP+FN} \end{cases} \quad (13)$$

Parameter settings. To get the result of correspondence grouping step with relatively high recall and precision, the parameters should be set reasonably. As efficiency is another important concern, parameters that result in less total iteration number would be preferred. Experiment results on three datasets are shown in Fig.3, from which IGSP's main parameters are determined in consideration of the aforementioned criteria. Table 1 shows the parameter settings of the proposed IGSP method. As for the denotation, r_k is the Non-maximum suppression radius of BSC. The parameter settings are used for all the experiments in this paper.

4.2. Results, evaluation and analysis

Registration results. Fig.4 show different phases of registration results using the proposed IGSP method on three
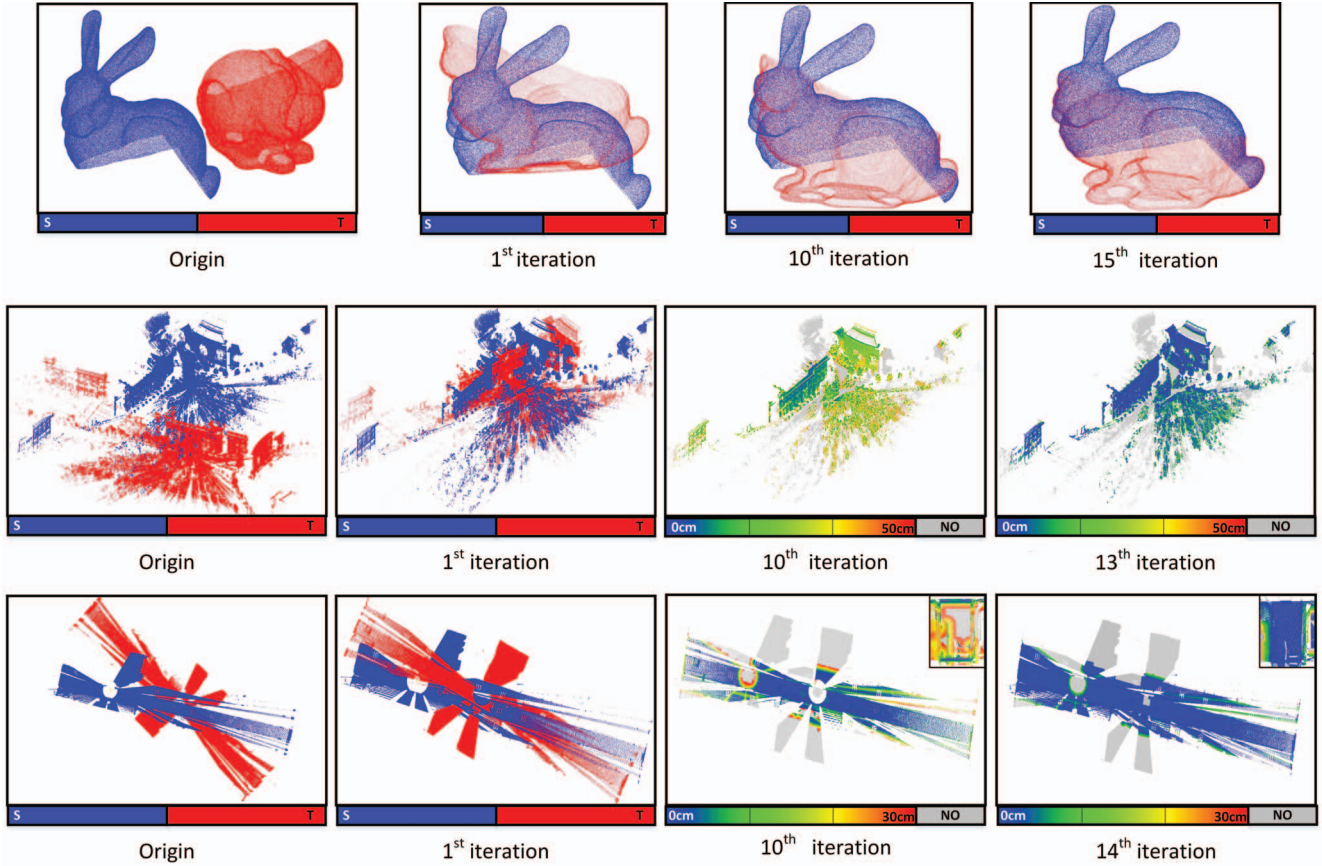
185

Figure 4. Registration result of (a)Stanford Bunny, (b)TLS Park and (c)TLS Indoor datasets during the iterative process. Source and target point clouds are shown in blue and red respectively. The Gradient color indicate the cloud to cloud distance of overlapping area and the non-overlapping points are shown in gray.

Table 1. Parameter Settings of the proposed IGSP method.

Parameter	Description	Value
s_{ed}	Scale factor of ED	$\frac{k}{r_k}, k = 30$
m	Iterative weight changing rate	8
p_1^t	Initial threshold parameter	2.5
p_2^t	Threshold parameter for ED	1.5
p_3^t	Threshold parameter for FD	1.25

Figure 5. Registration result of other challenging real-world datasets

testing datasets respectively. As seen in these figures,by the IGSP method, the point cloud pairs are iteratively converged from coarse to fine and get registered successfully. To further test the robustness, the algorithm is applied on other challenging real-world datasets, as shown in Fig.5. These results show that the proposed IGSP method performs well for various scenes, including those with repetitive, symmetric, noisy and incomplete structures, which are quite challenging for previous methods.
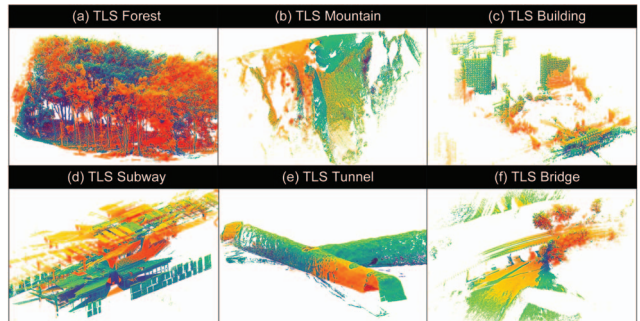
Registration accuracy evaluation. Average rotation error and translation error on three datasets are reported in Table 3. The errors show that the proposed IGSP method performs well in aligning both the tiny model and the real-world datasets, with average rotation error less than 0.1 degree and translation error less than 0.1 meter, which provides a good foundation for further applications like 3D re-
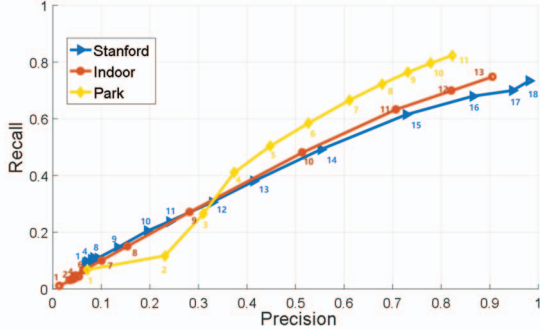
186

Figure 6. PR-Curve during iterative process on different Datasets

Table 2. Time performance of the proposed IGSP method.

Dataset	$T_1(s)$	$T_2(s)$	#Iteration	$T(s)$
Stanford	3.31	0.53	15	11.26
Park	8.12	5.65	13	81.57
Indoor	6.25	3.92	14	61.13

Table 3. Registration accuracy and time performance comparison

Dataset	Method	e^r(mdeg)	e^t(mm)	T(s)
Stanford 30% overlapped	ICP[7]	/	/	5.6
	3DNDT[26]	/	/	**5.3**
	Super4PCS[27]	74.21	0.17	10.5
	FM+GC[11]	758.23	2.59	8.9
	IGSP	**9.89**	**0.02**	11.3
Park 65% overlapped	ICP	/	/	9.4
	3DNDT	/	/	**8.1**
	Super4PCS	205.14	187.65	38.0
	FM+GC	184.56	414.61	35.2
	IGSP	**93.74**	**85.12**	81.6
Indoor 70% overlapped	ICP	/	/	6.2
	3DNDT	/	/	**6.0**
	Super4PCS	209.85	/	47.1
	FM+GC	486.15	431.90	29.8
	IGSP	**100.42**	**25.53**	61.1

construction and object extraction.

Accuracy analysis. Fig.6 shows the precision-recall curve during IGSP's iterative process for three datasets, in which each point represents a temporal correspondence result. Since the global optimal correspondences are applied in each iteration, both the precision and recall of the correspondence increase through the process and finally exceed 0.75 on all these datasets, thus resulting in excellent registration performance.

Time performance analysis. A time efficiency evaluation of the proposed IGSP method is conducted with regard to the runtime in each step after proper keypoint detection parameters and Iterative convergence condition are set. Table 2 lists the the average iteration number k, time for registration preparation T_1, time for each iteration T_2 and the total runtime ($T = T_1 + kT_2$). As analyzed before, the time complexity of KM algorithm is $O(n^3)$, in which n is the key point number.

Performance comparison and analysis. To further analyze the performance of the proposed IGSP method, several pairwise point cloud registration methods (ICP[7],3D-NDT[26],Super4PCS[27] and feature matching with geometric consistency (FM+GC)[11]) are selected for performance comparison using the Park dataset and the Indoor dataset. Key parameters of all the compared methods are set according to the parameter settings recommended in the original articles.

Table 3 lists the average registration errors and runtime of the compared methods on three datasets, in which '/' means the registration failed (the value is great than 1000). It is found that ICP and 3D-NDT fail when good initial alignment or prior knowledge is not provided. Super4PCS

has poor performance on point clouds with limited overlapping and too many similar structures, especially for the indoor corridor. IGSP outperforms the FM+GC since many correct matching keypoint pairs are also rejected by geometric consistency filter, which leads to a relatively low recall of keypoints matching. However, the time efficiency of IGSP is inferior to all the compared methods due to the iteration process and the $O(n^3)$ time complexity of KM algorithm, which will be the main concern of our future work.

5. Conclusion and future work

Nowadays, point cloud registration is the basis of many applications. This paper presented the Iterative Global Similarity Points (IGSP) algorithm, which iteratively find the corresponding keypoints considering global similarity and estimate the rigid body transformation to achieve coarse-to-fine pairwise point clouds registration. We validated its performance on different scenarios. Comprehensive experiments indicated that the proposed IGSP algorithm obtained good performance in correspondence precision, recall and registration accuracy. Although the proposed method provides satisfactory registration results, it is time consuming due to the iteration process and the $O(n^3)$ time complexity of KM algorithm. In future work, we will try other efficient and reasonable methods (*e.g.* Minimum Cost Max Flow (MCMF) and Graph Cut) to solve the proposed energy function. A smooth term for geometric consistency in energy function will also be considered.

References

[1] D. Aiger, N. J. Mitra, and D. Cohen-Or. 4-points congruent sets for robust surface registration. *ACM Transactions on Graphics*, 27(3):#85, 1–10, 2008.

[2] A. Albarelli, S. R. Bul, A. Torsello, and M. Pelillo. Matching as a non-cooperative game. In *2009 IEEE 12th International Conference on Computer Vision*, pages 1319–1326, Sept 2009.

[3] K.-H. Bae and D. D. Lichti. A method for automated registration of unorganised point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing*, 63(1):36 – 54, 2008. Theme Issue: Terrestrial Laser Scanning.

[4] M. Barczyk and S. Bonnabel. Towards realistic covariance estimation of icp-based kinect v1 scan matching: The 1d case. In *2017 American Control Conference (ACC)*, pages 4833–4838, May 2017.

[5] S. Barnea and S. Filin. Keypoint based autonomous registration of terrestrial laser point-clouds. *ISPRS Journal of Photogrammetry and Remote Sensing*, 63(1):19 – 35, 2008. Theme Issue: Terrestrial Laser Scanning.

[6] S. Belongie, J. Malik, and J. Puzicha. Shape matching and object recognition using shape contexts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(4):509–522, Apr 2002.

[7] P. Besl and N. D. McKay. A method for registration of 3-d shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):239–256.

[8] G. Blais and M. D. Levine. Registering multiview range data to create 3d computer objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(8):820–824, Aug 1995.

[9] H. Chen and B. Bhanu. 3d free-form object recognition in range images using local surface patches. *Pattern Recognition Letters*, 28(10):1252 – 1262, 2007.

[10] B. Curless and M. Levoy. A volumetric method for building complex models from range images. In *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '96, pages 303–312, New York, NY, USA, 1996. ACM.

[11] Z. Dong, B. Yang, F. Liang, R. Huang, and S. Scherer. Hierarchical registration of unordered tls point clouds based on binary shape context descriptor. *ISPRS Journal of Photogrammetry and Remote Sensing*, 144:61 – 79, 2018.

[12] Z. Dong, B. Yang, Y. Liu, F. Liang, B. Li, and Y. Zang. A novel binary shape context for 3d local surface description. *ISPRS Journal of Photogrammetry and Remote Sensing*, 130:431 – 452, 2017.

[13] G. Elbaz, T. Avraham, and A. Fischer. 3d point cloud registration for localization using a deep neural network autoencoder. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 00, pages 2472–2481, July 2017.

[14] M. A. Fischler and R. C. Bolles. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. In M. A. Fischler, , and O. Firschein, editors, *Readings in Computer Vision*, pages 726 – 740. Morgan Kaufmann, San Francisco (CA), 1987.

[15] Y. Gao, Q. Dai, and N.-Y. Zhang. 3d model comparison using spatial structure circular descriptor. *Pattern Recognition*, 43(3):1142 – 1151, 2010.

[16] A. Gressin, C. Mallet, J. Demantk, and N. David. Towards 3d lidar point cloud registration improvement using optimal neighborhood knowledge. *ISPRS Journal of Photogrammetry and Remote Sensing*, 79:240 – 251, 2013.

[17] Y. Guo, F. Sohel, M. Bennamoun, M. Lu, and J. Wan. Rotational projection statistics for 3d local surface description and object recognition. *International Journal of Computer Vision*, 105(1):63–86, Oct 2013.

[18] Y. Guo, F. Sohel, M. Bennamoun, J. Wan, and M. Lu. An accurate and robust range image registration algorithm for 3d object modeling. *IEEE Transactions on Multimedia*, 16(5):1377–1390, Aug 2014.

[19] P. Huang, M. Cheng, Y. Chen, D. Zai, C. Wang, and J. Li. Solar potential analysis method using terrestrial laser scanning point clouds. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10(3):1221–1233, March 2017.

[20] D. F. Huber and M. Hebert. Fully automatic registration of multiple 3d data sets. *Image and Vision Computing*, 21(7):637 – 650, 2003. Computer Vision beyond the visible spectrum.

[21] A. E. Johnson and M. Hebert. Using spin images for efficient object recognition in cluttered 3d scenes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(5):433–449, May 1999.

[22] A. E. Johnson and M. Hebert. Using spin images for efficient object recognition in cluttered 3d scenes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(5):433–449, May 1999.

[23] A. E. Johnson and M. Hebert. Using spin images for efficient object recognition in cluttered 3d scenes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(5):433–449, May 1999.

[24] J. Knopp, M. Prasad, G. Willems, R. Timofte, and L. Van Gool. Hough transform and 3d surf for robust three dimensional classification. In K. Daniilidis, P. Maragos, and N. Paragios, editors, *Computer Vision – ECCV 2010*, pages 589–602, Berlin, Heidelberg, 2010. Springer Berlin Heidelberg.

[25] T.-W. R. Lo and J. P. Siebert. Local feature extraction and matching on range images: 2.5d sift. *Computer Vision and Image Understanding*, 113(12):1235 – 1250, 2009. Special issue on 3D Representation for Object and Scene Recognition.

[26] M. Magnusson. *The three-dimensional normal-distributions transform : an efficient representation for registration, surface analysis, and loop detection*. PhD thesis, ?rebro University, School of Science and Technology, 2009.

[27] N. Mellado, D. Aiger, and N. J. Mitra. Super 4pcs fast global pointcloud registration via smart indexing. *Computer Graphics Forum*, 33(5):205–215.

[28] A. Montuori, G. Luzi, S. Stramondo, G. Casula, C. Bignami, E. Bonali, M. G. Bianchi, and M. Crosetto. Combined use of ground-based systems for cultural heritage conservation

monitoring. In *2014 IEEE Geoscience and Remote Sensing Symposium*, pages 4086–4089, July 2014.

[29] T. Pajdla and L. V. Gool. Matching of 3-d curves using semi-differential invariants. In *Proceedings of IEEE International Conference on Computer Vision*, pages 390–395, Jun 1995.

[30] R. Rusu and S. Cousins. 3d is here: Point cloud library (pcl). In *IEEE International Conference on Robotics and Automation 2011 (ICRA 2011)*, 05 2011.

[31] R. B. Rusu, N. Blodow, and M. Beetz. Fast point feature histograms (fpfh) for 3d registration. In *2009 IEEE International Conference on Robotics and Automation*, pages 3212–3217, May 2009.

[32] R. B. Rusu, G. Bradski, R. Thibaux, and J. Hsu. Fast 3d recognition and pose using the viewpoint feature histogram. In *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 2155–2162, Oct 2010.

[33] I. Sipiran and B. Bustos. Harris 3d: a robust extension of the harris operator for interest point detection on 3d meshes. *The Visual Computer*, 27(11):963, Jul 2011.

[34] A. N. K. L. J. H. H. Surmann. 6d slamł3d mapping outdoor environments. *Journal of Field Robotics*, 24(8):699–722.

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

[36] F. Tombari, S. Salti, and L. Di Stefano. Unique signatures of histograms for local surface description. In *Proceedings of the 11th European Conference on Computer Vision Conference on Computer Vision: Part III*, ECCV'10, pages 356–369, Berlin, Heidelberg, 2010. Springer-Verlag.

[37] F. Tombari and L. D. Stefano. Object recognition in 3d scenes with occlusions and clutter by hough voting. In *2010 Fourth Pacific-Rim Symposium on Image and Video Technology*, pages 349–355, Nov 2010.

[38] G. Vosselman and H.-G. Maas. *Airborne and terrestrial laser scanning*. CRC Press, 2010.

[39] K. H. W. The hungarian method for the assignment problem. *Naval Research Logistics Quarterly*, 2(1):83–97, 1955.

[40] Y. Xu, R. Boerner, W. Yao, L. Hoegner, and U. Stilla. 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/W4:185–191, 2017.

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

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

[43] D. Zai, J. Li, Y. Guo, M. Cheng, P. Huang, X. Cao, and C. Wang. Pairwise registration of TLS point clouds using covariance descriptors and a non-cooperative game. *ISPRS Journal of Photogrammetry and Remote Sensing*, 134:15–29, Dec. 2017.

[44] Y. Zhong. Intrinsic shape signatures: A shape descriptor for 3d object recognition. In *2009 IEEE 12th International Conference on Computer Vision Workshops, ICCV Workshops*, pages 689–696, Sept 2009.