A Single Correspondence Is Enough: Robust Global Registration to Avoid Degeneracy in Urban Environments

Hyungtae Lim\(^1,3\), Student Member, IEEE, Suyong Yeon\(^2\), Soohyun Ryu\(^2\), Yonghan Lee\(^2\), Youngji Kim\(^2\), Jaeseong Yun\(^2\), Euigon Jung\(^2\), Donghwan Lee\(^2\), and Hyun Myung\(^1\), Senior Member, IEEE

Abstract—Global registration using 3D point clouds is a crucial technology for mobile platforms to achieve localization or manage loop-closing situations. In recent years, numerous researchers have proposed global registration methods to address a large number of outlier correspondences. Unfortunately, the degeneracy problem, which represents the phenomenon in which the number of estimated inliers becomes lower than three, is still potentially inevitable. To tackle the problem, a degeneracy-robust decoupling-based global registration method is proposed, called Quatro. In particular, our method employs quasi-SO(3) estimation by leveraging the Atlanta world assumption in urban environments to avoid degeneracy in rotation estimation. Thus, the minimum degree of freedom (DoF) of our method is reduced from three to one. As verified in indoor and outdoor 3D LiDAR datasets, our proposed method yields robust global registration performance compared with other global registration methods, even for distant point cloud pairs. Furthermore, the experimental results confirm the applicability of our method as a coarse alignment. Our code is available: https://github.com/url-kaist/quatro

I. INTRODUCTION

3D point cloud registration is a method to align 3D point clouds by estimating relative pose or poses between two or more 3D point clouds [1]. 3D registration methods are widely utilized in various research fields such as object recognition [2]–[5], mapping [6]–[8], localization [9]–[12], and so forth. In particular, 3D registration is also a fundamental task for terrestrial mobile platforms to estimate egomotion [13], [14], i.e. odometry, and to manage loop closing situations [15], [16].

Accordingly, the registration algorithms are mainly divided into two categories. One is the local registration [17]–[22], and the other is the global registration [23]–[27]. Iterative Closest Point (ICP) [17] is a renowned local registration method and it has made an impact on subsequent studies. Unfortunately, the ICP-variants set point pairs by using a greedy, exhaustive nearest neighbor (NN) search for every iteration. Consequently, the local registration methods are only applicable when two point clouds, namely source and target, are close enough and nearly overlapped [22]. Otherwise, the correspondences via the NN search are likely to become invalid; thus, the result of the local registration may get caught in the local minima [20].

The global registration [23]–[27], on which this paper places emphasis, aims to estimate the relative pose between distant and partially overlapped point clouds. Because the global registration is relatively invariant to the initial pose difference, the global registration is usually used to provide an initial alignment close enough for the local registration to allow the estimate to converge on the global minima. The global registration can be further classified into correspondence-based [23], [26]–[28] and correspondence-free methods [29]–[31], but the latter ones are beyond our scope.

Some of the well-known correspondence-based methods are random sample consensus (RANSAC) [23] and its variants [32]–[35], yet these are known to become slow and brittle with high outlier rates (> 50%) [28]. On the other hand, branch-and-bound (BnB)-based methods [36]–[38] have been proposed for theoretical optimality guarantees. Unfortunately, BnB-based methods are not applicable to real-world applications because they are time-consuming (> 50 sec) [39].

![Fig. 1. (T-B, L-R): Before and after the application of our proposed method called Quatro in KITTI dataset [7] when two distant and partially overlapped point clouds, i.e. source (cyan) and target (yellow), are given. As the distance between the two viewpoints of source and target becomes farther, it gives rise to an increase in the ratio of outliers within the putative correspondences and simultaneously reduces the number of inliers, which results in the performance degradation of correspondence-based global registration methods in general [23], [26], [28]. Under the circumstances, our proposed method shows robust performance, overcoming the effect of outliers, as well as the degeneracy issue. The red and green lines denote outlier and inlier correspondences, respectively (best viewed in color).](https://github.com/url-kaist/quatro)
Aiming for both outlier-robust and fast registrations, gradient non-convexity (GNC) has been introduced based on Black-Rangarajan duality [26], [40]. For instance, Zhou et al. [26] showed that GNC outperforms previous approaches and is faster than existing methods by at least ten times, while tolerating 70-80% of outliers. As a result, many researchers have employed GNC to achieve robust pose estimation [27], [28], [41], [42]. In the meantime, semidefinite programming and sums of squares relaxation-based methods have been proposed to solve problems in polynomial-time with certifiable optimality guarantees [43]–[46]. Further, novel graph-theoretic methods also have been proposed to prune spurious correspondences effectively [47], [48].

Unfortunately, a degeneracy problem is still potentially inevitable. Degeneracy usually refers to various phenomena in the perceptually degraded environments, such as corridors, mines, and so forth [49]–[51]. However, in this paper, we specify degeneracy as the phenomenon in which the number of the estimated inliers becomes lower than the minimum number of inliers during registration, which is generally three in the 3D space. Note that this effect can occur not only in the perceptually degraded environments but also in some cases where two distant and partially overlapped point clouds are given, resulting in catastrophic failure of global registration.

To sum up, the reason for the degeneracy problem is mainly twofold. First, there is a sparsity issue with 3D LiDAR sensors [52]–[55], which means, as the distance from the sensor coordinate becomes farther away, a point cloud becomes too sparse; thus, it induces inaccurate feature matching. That is, even though the two point clouds represent the same space locally, the local areas have different densities because the source and target are captured in a different location. Consequently, the sparsity issue directly affects the quality of feature descriptors and then gives rise to more false correspondences. Thus, the ratio of the outliers within the putative correspondences increases, as shown in Fig. 1. Second, these outliers pair let the outlier-rejection algorithms, such as graph-theoretic pruning methods [47], [48] or weight update in GNC [26], [27], reject the putative correspondences dramatically to prune outliers. Accordingly, too many correspondences are occasionally filtered. Finally, the number of the estimated inliers occasionally becomes less than three in the process, resulting in degeneracy.

To tackle the degeneracy problem, we propose a novel global registration method, called Quatro, which is a combination of the words quasi- and cuatro meaning 4-DoF (degree of freedom). Some degeneracy-aware SLAM frameworks have been proposed in [49]–[51]. However, to the best of our knowledge, this is the first attempt to deal with the degeneracy in global registration by reducing the minimum number of required correspondences of SE(3) from three to one, named quasi-SE(3) by decoupling SE(3) estimation into quasi-SO(3) estimation followed by component-wise translation estimation (COTE) [27].

In summary, the contribution of this paper is threefold.

- To avoid degeneracy, we propose a novel rotation estimation method called quasi-SO(3) estimation by utilizing the characteristics of urban environments.
- As a result, our proposed method shows a more promising performance over the state-of-the-art methods on real-world indoor/outdoor 3D LiDAR datasets.
- In particular, it is remarkable that Quatro-c2ξ, a simple coarse-to-fine scheme, outperforms the state-of-the-art methods, including deep learning-based approaches.

Therefore, we finally confirm the suitability of our proposed method as a coarse alignment.

II. QUATRO: QUASI-SE(3) ESTIMATION

A. Problem Definition

First, we begin by denoting source and target point cloud which are captured by a 3D LiDAR sensor on different viewpoints as \( P \) and \( Q \), respectively. Then, let us define \( P = \{ p_1, p_2, \ldots, p_N \} \) and \( Q = \{ q_1, q_2, \ldots, q_M \} \), where each point of the clouds, \( p_i \) (\( 1 \leq i \leq N \)) and \( q_j \) (\( 1 \leq j \leq M \)) \( \in \mathbb{R}^3 \), consists of \( \{ x, y, z \} \) in the Cartesian coordinate. It is assumed that the \( xy \)-plane of \( P \) and \( Q \) are already aligned with that of mobile platforms to utilize the Atlanta world assumption [56] (see Section II.C).

Next, let \( (i, j) \in A \) be candidate correspondences where \( i \) and \( j \) denote the indices of points in \( P \) and \( Q \), respectively. In practice, \( A \) inevitably includes inherent outlier set \( O \), such that \( O \cup O^c = A \). This happens in most of the 3D point feature matching algorithms [28], [57], [58]. Accordingly, the relationship of each pair can be expressed as follows:

\[
q_j = Rp_i + t + \epsilon_{ij}
\]  

(1)

where \( R \in SO(3) \) and \( t \in \mathbb{R}^3 \) are the relative rotation and translation, respectively, and \( \epsilon_{ij} \in \mathbb{R}^3 \) denotes the unknown measurement noise. That is, \( \epsilon_{ij} \) is Gaussian noise if \( (i, j) \in O^c \) or is irregular error if \( (i, j) \in O \). Finally, our objective function could be defined as follows:

\[
\hat{R}, \hat{t} = \arg \min_{R \in SO(3), t \in \mathbb{R}^3} \sum_{(i, j) \in A} \rho \left( r(q_j - Rp_i - t) \right)
\]  

(2)

where \( \rho(\cdot) \) denotes surrogate function [28] to suppress undesired large errors produced by \( O \) and \( r(\cdot) \) denotes the squared residual function, i.e. \( |\cdot|^2 \). In summary, our goal is to estimate \( \hat{R} \) and \( \hat{t} \) while suppressing the effect of outliers as much as possible by employing \( \rho(\cdot) \).

B. Decoupling Rotation and Translation Estimation

According to [59] and [60], \( \hat{R} \) and \( \hat{t} \) can be easily obtained in closed form by decoupling rotation and translation estimation. To this end, first of all, two pairs \( (i, j) \) and \( (i', j') \) are subtracted to cancel out the effect of \( t \) by using (1). For simplicity, let \( \alpha_k = p_{i'} - p_i; \beta_k = q_{j'} - q_j; \epsilon_k = \epsilon_{ij'} - \epsilon_{ij} \), and let \( K \) be the total number of these translation invariant measurements (TIMs) [27]. Note that \( K = |A| \) by subtracting consecutive pairs to build TIMs in a chain form to minimize computational cost and the first TIM is made by the subtraction between the first and \( |A| \)-th pair.

Then, the relationship between \( \alpha_k \) and \( \beta_k \) can be expressed as \( \beta_k = R\alpha_k + \epsilon_k \). Accordingly, \( \hat{R} \) is estimated as follows:
of pitch and roll were observed to be much smaller than $z$ to be coplanar based on the Atlanta world assumption [56].

One might argue that this assumption does not hold against degeneracy. For simplicity, the approximated $R$ where $\bar{I}$ quasi-SO(3) directly. Finally, this assumption reduces DoF of rotation from three to one, making our method robust against degeneracy. For simplicity, the approximated $R$ is denoted by $R_+$. One might argue that this assumption does not hold in non-flat regions or if two viewpoints are no longer located in coplanar regions as the distance between two viewpoints increases [63]–[65]. However, most current mapping/navigation systems employ an inertial navigation system (INS). Accordingly, roll and pitch angles are fully observable by estimating the horizontal plane from the gravity vector [65], [66]. Consequently, the roll and pitch can be expressed as absolute states in the world coordinate, which means these are drift-free [67]. Thus, this still allows our goal to be valid, while restricting the relative rotation to relative yaw rotation (see Section IV.D).

D. Quasi-SO(3) Estimation using Graduated Non-Convexity

To estimate $R_+\, GNC$ with a truncated least square [28] is introduced. For simplicity, specifying the $k$-th measurement pair, i.e. $\alpha_k$ and $\beta_k$, as $\zeta_k$, [26] can be redefined as $arg\min_{R_+} \sum_{k=1}^K \rho_\mu (r(\zeta_k, R_+))$, where $\rho_\mu (\cdot)$ denotes surrogate function governed by parameter $\mu$, as shown in Fig. 2(b). For this, the equation is first rewritten by leveraging Black-Rangarajan duality as follows [26], [27]:

$$
R_+ = \arg\min_{R_+ \in \text{SO}(3), w_k \in [0, 1]} \sum_{k=1}^K \left[ w_k r(\zeta_k, R_+) + \Phi_{\rho_\mu} (w_k) \right]
$$

where $\Phi_{\rho_\mu} (w_k) = \mu (1 - w_k) + \epsilon^2$ is a penalty term [42]. Unfortunately, the objective function cannot be directly solved [26]. Thus, (5) is solved by using alternating optimization as follows:

$$
\tilde{R}_+^{(t)} = \arg\min_{R_+} \sum_{i=1}^K \tilde{w}_k^{(t-1)} r(\zeta_k, R_+), \quad (6)
$$

$$
\tilde{W}^{(t)} = \arg\min_{w_k \in [0, 1]} \sum_{k=1}^K \left[ w_k r(\zeta_k, \tilde{R}_+^{(t)}) + \Phi_{\rho_\mu} (w_k) \right] \quad (7)
$$

where the superscript $(t)$ denotes the $t$-th iteration and $\tilde{W} = \text{diag} (\tilde{w}_1, \tilde{w}_2, \ldots, \tilde{w}_K)$. Each $\tilde{w}_k$ can be solved in a truncated closed form as follows:

$$
\tilde{w}_k^{(t)} = \begin{cases} 
0 & \text{if } \tilde{r}_k \in \left[ \frac{\mu + 1}{\mu + \epsilon^2}, +\infty \right] \\
\tilde{c}_k \sqrt{\frac{\mu (\mu + 1)}{\epsilon^2}} & \text{if } \tilde{r}_k \in \left[ \frac{\mu + 1}{\mu + \epsilon^2}, \frac{\mu + 1}{\mu + \epsilon^2} \right] \\
1 & \text{if } \tilde{r}_k \in \left[ 0, \frac{\mu + 1}{\mu + \epsilon^2} \right]
\end{cases} \quad (8)
$$

where $\tilde{r}_k$ denotes $r(\zeta_k, \tilde{R}_+^{(t)})$ for simplicity. The overall procedure is illustrated in Fig. 3. For each iteration, $\mu$ is updated as $\mu(t) \leftarrow \kappa \cdot \mu(t-1)$, where $\mu(0) = \epsilon^2 / (\max(r(\zeta_k, I_3)) - \epsilon^2)$ and $\kappa > 1$ is a factor that increases the magnitude of non-convexity gradually, as presented in Fig. 2(b). The iteration ends if the differential of $\sum_{k=1}^K \tilde{w}_k^{(t)} r(\zeta_k, \tilde{R}_+^{(t)})$ becomes small enough.

Unfortunately, the outliers have a direct negative impact on the optimization at the first iteration because all the weights are initialized to 1, that is, $\tilde{W}^{(0)} = I_3$ (Fig. 3(c)). Besides, in weight update, as in (8), most weights are occasionally assigned to zero (Fig. 3(d)), resulting in degeneracy in the 3D space (Figs. 3(d) and 3(g)). However, $R_+$ is designed to suppress the effect of outliers, as well as to be robust against degeneracy; thus, catastrophic failure can be prevented, as shown in Fig. 3(e).
quasi-inliers rarely affect the estimation of $\epsilon$ of the ground. Accordingly, the so forth, are likely to be similar along the normal direction geometrical characteristics, e.g. surface normal, density, and because the ground can be considered locally flat enough and plane, $\epsilon$ two terms: a) a parallel term to the ground plane, i.e. $xy$. However, note that $\epsilon$ has advantages over $\epsilon$ and $\epsilon$. In short, $\epsilon$ has advantages over $\epsilon$ for the following three reasons. First, estimation of $\epsilon$ resolves the degeneracy issue in itself. As previously mentioned, the number of estimated inliers is not always guaranteed to be more than the minimum DoF of rotation in GNC. However, note that the minimum DoF of $\epsilon$ is three, whereas that of $\epsilon$ is one. As a result, $\epsilon$ can be more robust against degeneracy. Empirically, when the distance between the source and target viewpoints becomes more than 9 meters away, for example, in outdoor environments, the degeneracy occasionally occurs. However, our method can conduct robust registration, even if the two viewpoints are farther apart (see Section IV.A).

Second, $\epsilon$ helps to suppress the effect of the outliers. Because the ground can be considered locally flat enough and most structures tend to be orthogonal to the ground, the local geometrical characteristics, e.g. surface normal, density, and so forth, are likely to be similar along the normal direction of the ground. Accordingly, the $\epsilon_k$ can be decomposed into two terms: a) a parallel term to the ground plane, i.e. $xy$-plane, $\epsilon_k$ and b) a perpendicular one $\epsilon_k^\perp$, which satisfies $\epsilon_k = \epsilon_k^\parallel + \epsilon_k^\perp$. Consequently, the outliers are classified into two groups. One is the quasi-inliers, $Q$, which satisfies $|\epsilon_k^\parallel| \approx 0$ and $|\epsilon_k^\perp| \gg 0$, and the other is the definite outliers, $D$, which satisfies $|\epsilon_k^\perp| \gg 0$. Note that $\hat{Q} \cup \hat{D} = \hat{O}$, where $\hat{O}$ denotes the estimated output of GNC. In that context, $\epsilon_k^\perp$ of quasi-inliers rarely affect the estimation of $\epsilon$ because the estimation of $\epsilon$ is invariant to $z$-values. Thus, $Q$ can be considered as additional inliers when optimizing $\epsilon$.

Third, as an extension of the second reason, $Q$ has a positive impact on divergence prevention by increasing the ratio of inliers (i.e. $|\hat{O}^|\{\epsilon\}|$ to $|\hat{O}^|\{\epsilon\}|$). This is because GNC-based optimization sometimes diverge when the ratio of the outlier correspondences is too large.

E. Component-wise Translation Estimation

Finally, the relative translation is estimated in a component-wise way, as shown in Fig. 3(f). Let the boundary interval set $\mathcal{E}$ be the $2|A|$-tuples that comprises the lower bound, $i_{ij} = \sigma_{ij}c$, and the upper bound, $i_{ij} + \sigma_{ij}c$, and assume that all the elements of $\mathcal{E}$ is sorted in ascending order. Then, let the $g$-th consensus set be $\mathcal{I}_g = \{(i,j) | (\frac{\sigma_{ij}}{|\mathcal{I}_g|})^2 \leq \epsilon^2\}$, where $\phi_g \in \mathbb{R}$ is any value that satisfies $\mathcal{E}(g) < \phi_g < \mathcal{E}(g+1)$ for $g = 1, 2, \ldots, 2|A|-1$. Then, $\hat{t}_g$ is estimated by the weighted average for non-empty $\mathcal{I}_g$ as follows:

$$\hat{t}_g = \left( \sum_{(i,j) \in \mathcal{I}_g} \frac{1}{\sigma^2_{ij}} \right)^{-1} \sum_{(i,j) \in \mathcal{I}_g} \frac{i_{ij}}{\sigma^2_{ij}}.$$  (9)

Finally, among up to $2|A|-1$ candidates $\hat{t}_g$, $\hat{t}$ is selected which minimizes the truncated objective function $\phi$. Note that COTE is based on the premise that the estimated rotation is precise enough.

Therefore, our method finally enables degeneracy-robust global registration even though a single pair is left due to poor feature matching and drastic correspondence pruning.

F. Preprocessing of Correspondences

In this paper, fast point feature histogram (FPFH) is utilized, which is widely used as a conventional descriptor for the registration. However, the original FPFH for a 3D point cloud captured by a 64-channel LiDAR sensor takes tens of seconds, which is too slow. For this reason, we employ voxel-sampled FPFH, which is preceded by voxel-sampling with voxel size $v$. This is followed by the correspondence test that outputs $A_{raw}$, as represented in Fig. 3(a).

Thereafter, the max clique inlier selection (MCIS) is applied, which takes $A_{raw}$ as input and outputs $A$, as shown in Fig. 3(b). TEASER++ is the representative method to adopt the MCIS first, and the researchers showed that MCIS successfully discards gross outliers. We found that finding the exact maximal clique is the bottleneck of the algorithm in terms of speed. Thus, we revise it to find a heuristic maximal clique whose cardinality is large enough, which is named as MCIS-heuristic.

III. EXPERIMENTS

A. Dataset

We experimented in the indoor/outdoor environments, i.e. KITTI dataset and NAVER LABS localization dataset. In particular, we tested the algorithms in a) loop-closing situations, b) augmented rotation situations, and c) odometry situations.
Fig. 4. Degeneracy cases (T-B): Registration results between 1,398 to 3,554 frames in Seq. 00, which are 9.88 m apart, and between 177 and 1,726 frames in Seq. 08, which are 9.82 m apart. The red, green, and blue points denote source, target, and estimate, respectively. Note that RANSAC \cite{23} returns $I_4$ if it fails to find correct putative correspondences. The closer the green and blue points are, the better (best viewed in color).

**B. Error Metrics**

As quantitative metrics, the average relative pose errors, $t_{\text{avg}}$ for translation and $r_{\text{avg}}$ for rotation, are used as follows:

- $t_{\text{avg}} = \frac{\sum_{n=1}^{N} (t_{n,\text{GT}} - \hat{t}_{n})^2}{N}$,
- $r_{\text{avg}} = \frac{180 \pi \sum_{n=1}^{N} \cos^{-1}\left(\frac{\text{Tr}(\hat{\mathbf{R}}_n^T \mathbf{R}_{n,\text{GT}}) - 1}{2}\right)}{N}$

where the subscript GT denotes that the value is from the ground truth. For the odometry test, the relative odometry errors, $t_{\text{rel}}$ and $r_{\text{rel}}$, are used, which are calculated by \cite{71}.

**C. Parameters of Quatro**

Empirically, we set $\bar{c} = 0.15$ and $\kappa = 1.4$. Because of the sparsity differences depend on the number of channels of LiDAR sensors, the parameters for voxel-sampled FPFH should be changed depending on the sensor configuration. Thus, we set $\nu = 0.3$ m, radius for normal estimation $r_{\text{norm}} = 0.5$ m, and radius for FPFH $r_{\text{FPFH}} = 0.65$ m in the KITTI dataset \cite{7}, and $\nu = 0.1$ m, $r_{\text{norm}} = 0.3$ m, and $r_{\text{FPFH}} = 0.45$ m in the NAVER LABS localization dataset \cite{70}.

**IV. EXPERIMENTAL RESULTS**

To check the effectiveness of our proposed method, Quatro was quantitatively compared with state-of-the-art methods, namely, RANSAC \cite{23}, (specifically, RANSAC1K and RANSAC10K, where each number denotes the number of iteration), FGR \cite{26}, and TEASER++ \cite{27}. We leveraged the open-source implementations for the experiment. Quatro-c2f comprises the proposed Quatro as a coarse alignment, followed by G-ICP \cite{20} as a fine alignment.

**A. Impact of Quasi-SE(3) in Distant Global Registration**

In general, the state-of-the-art methods estimate precise relative pose by overcoming the effect of gross outliers. However, our Quatro exhibits noticeable robustness in the case where the relative pose between two viewpoints of source and target is distant, as shown in Fig. 4 and Table I. In particular, it was observed that FGR was more likely to fail to conduct registration compared with TEASER++ and our Quatro. This is because FGR inherently uses linearization of $\text{SE}(3)$ in optimization. Thus, if the linearization assumption does not hold, its performance becomes degraded, as shown in Fig. 5. Furthermore, FGR has no additional outlier rejection procedure, i.e. MCIS. Consequently, as the ratio of outliers increases, the performance obviously decreases.

On the other hand, TEASER++, which employs MCIS similar to our proposed method, shows robust performance compared with FGR in the distant cases. Unfortunately, MCIS also occasionally rejected numerous matching pairs, so it was no longer guaranteed that there are more than three estimated inliers left in the estimating rotation. Note that TEASER++ also utilized decoupling of rotation and translation, so if TEASER++ fails to estimate the rotation,
then it inevitably fails to estimate translation as a domino effect, as mentioned in Section II.E. Thus, TEASER++ shows relatively large variance as the range becomes larger, compared with our Quatro, as shown in Fig. 5(a).

In particular, the variance difference between TEASER++ and Quatro in estimating rotation becomes significant, as presented in Fig. 5(a). This is because our quasi-SE(3) reduces its minimum number of the required pairs to one via quasi-SO(3) estimation and COTE. Accordingly, it successfully avoids degeneracy, even if it is given less than three estimated inliers. As a result, catastrophic failure in the rotation estimation can be somewhat prevented. Furthermore, due to the effect of the quasi-inliers, some irregular errors (i.e. $c_k$) no longer affect the estimation of $R_y$. Thus, our method generally has a success rate of more than 99%.

Therefore, these experimental results corroborate that our quasi-SE(3) estimation enables robust global registration, while overcoming the degeneracy problem.

**B. Comparison in Loop Closing/Odometry Test**

As explained earlier, some registration methods show performance degradation as the distance or difference of rotation between two viewpoints of source and target becomes larger, as shown in Fig. 5. In contrast, our proposed method shows promising performance, especially with little variance in performance relative to other methods. Furthermore, it was shown that our quasi-SE(3) is additionally advantageous in indoor situations where the ground is mostly flat and the urban structures are orthogonal to the ground; thus, our assumption $\angle R_y \cdot R_x \approx I_3$ is obviously met. As a result, our method maintains the smallest rotation error, even though source and target clouds are in diametrically opposite directions.

Furthermore, our method shows better odometry results, as shown in Table 1. It is natural that local registration methods show better performance when the frame interval is small, but as the interval becomes larger, their performance drastically degraded. In contrast, the global registration methods show robust performance even though the frame interval becomes larger. In particular, performance of our method decreases little, showing smaller odometry errors than the other global registration methods.

**C. Quatro as a Coarse Alignment**

Finally, Quatro-c2f shows successful coarse-to-fine registration, as shown in Fig. 5(b) and Table 1. In particular, Quatro-c2f even shows better performance compared with the state-of-the-art methods, including conventional and deep learning-based approaches in a sequence that is used as a training dataset. Therefore, we finally confirm that Quatro can provide a sufficiently accurate coarse alignment, thus helping local registration algorithms conduct fine alignment successfully.

**D. Application: Leveraging INS in Non-flat Regions**

In addition, we conducted a feasibility study on the utilization of an INS system. That is, Quasi-SO(3) estimation is followed by the estimation of roll and pitch angles via INS measurements, i.e. $R_y \cdot R_x$. As presented in Table III, even though the raw measurements were used, leveraging pitch and roll measurements reduced the rotation error effectively. Therefore, the results show the possibility of generalization of our proposed method in non-flat regions.

**E. Registration Speed**

Furthermore, our proposed method shows the fastest optimization time by virtue of MCIS-heuristic, as represented in Fig. 6. On average, our method only takes 5 msec per optimization, which is sufficient for using our proposed method as a coarse alignment in real-time.

**V. Conclusion**

In this study, a robust global registration method, Quatro, has been proposed. Our proposed method proved to be more robust against degeneracy compared with the state-of-the-art methods. In future works, we plan to apply the concept of Quatro for generalization on various platforms, including UAVs, backpack-type mapping systems, and so forth.
