Grid-based scan-to-map matching for accurate 2D map building

Kunjin Ryu, Lakshitha Dantanarayana, Tomonari Furukawa & Gamini Dissanayake

RSJ

Taylor & Francis
Taylor & Francis Group

FULL PAPER

Grid-based scan-to-map matching for accurate 2D map building

Kunjin Ryu[a], Lakshitha Dantanarayana[b] [ORCID], Tomonari Furukawa[a,b] [ORCID] and Gamini Dissanayake[b] [ORCID]

[a]Department of Mechanical Engineering, Virginia Tech, Blacksburg, VA, USA; [b]Centre for Autonomous Systems, University of Technology, Sydney, Australia

ABSTRACT

This paper presents a grid-based scan-to-map matching technique for accurate 2D map building. At every acquisition of a new scan, the proposed technique matches the new scan to the previous scan similarly to the conventional techniques, but further corrects the error by matching the new scan to the globally defined map. In order to achieve best scan-to-map matching at each acquisition, the map is represented as a grid map with multiple normal distributions (NDs) in each cell, which is one contribution of this paper. Additionally, the new scan is also represented by NDs, developing a novel ND-to-ND matching technique. This ND-to-ND matching technique has significant potential in the enhancement of the global matching as well as the computational efficiency. Experimental results first show that the proposed technique accumulates very small errors after consecutive matchings and identifies that the scans are matched better to the map with the multi-ND representation than one ND representation. The proposed technique is then tested in a number of large indoor environments, including public domain datasets and the applicability to real world problems is demonstrated.

1. Introduction

Understanding its location and the surroundings is of paramount importance for a mobile robot placed in a totally unknown environment, before commencing any actions. The map of an environment is one of the most basic representations that can be used by a robot to understand the environment. This maps can be used for many aspects of robot navigation. The most common type of map that is popular amongst the community that are helpful for such navigation tasks are 2D grid maps.

The choicest way of addressing the problem of creating a map is known as Simultaneous Localization and Mapping (SLAM), which is building a map of the environment while at the same time localizing the robot in the map. One solution to this problem is given by iteratively observing the surrounding environment and associating a new observation containing some objects to the previous observation containing the same objects. The SLAM problem became even more important when the robot needs to autonomously explore the environment, and it is obvious that the ability of the robot is extremely limited without an accurate solution.[1,2] Since SLAM relies upon the successful association of the new observation to the past observations, existing SLAM techniques can be classified in terms of the type of data extracted from the observation into two categories, feature-based

and scan-based. In the feature-based SLAM techniques, a set of features including different types of distinct geometric models such as points, lines, curvatures, and any arbitrary shapes are extracted from the observation and used as landmarks to associate the new and previous observations.[3,4]

Although feature based SLAM has become extremely popular in the field and has many applications, in environments where it is not straight forward to extract features or when a grid based map is required from sensors such as laser range finders (LRF), scan based SLAM methods [5–8] are employed. These methods directly utilizes unprocessed scan readings as observations which are generally sets of points obtained from aforementioned range sensors such as a sonar [9] or LRF.[10] One of the fundamental issues of the scan-based SLAM has thus been the association or matching of one scan to another, namely the scan-to-scan matching.

The most common scan-to-scan matching approach is based on the Iterative Closest Point (ICP) technique [11] which allows the point-to-point matching between two scans by minimizing the total distance between them. Despite the popularity of the technique, the point-to-point matching may yield inappropriate data association since two corresponding points are not actually on the same position in the environment. Weiss and Puttkamer [12] proposed a technique that avoids the point-to-point

---

CONTACT Lakshitha Dantanarayana ✉ lakshitha.dantanarayana@uts.edu.au

matching problem by calculating an angle between two neighboring scan points for every scan point and using the angles to match the two scans. Biber and Straßer [13] represented a subdivided grid space and collectively describing a scan within each grid cell by a normal distribution (ND). This grid-based technique, the so-called Normal Distribution Transform (NDT), spatially associates every point of the new scan in a grid cell to the ND in the cell. The NDT does not suffer from the point-to-point matching, but the matching performance relies on the grid size and the initial guess.

Furthermore a number of approaches have been proposed in literature on integration of these scan matching techniques in SLAM and improving accuracy. Early efforts include the work of Lu and Milios [14], which performed the matching of the new scan to the previous scan and further matched all the scans by storing the past scans. Thrun et al. [15] used the expectation maximization (EM) algorithm that finds the best matching past scan to the new scan from all the past scans. The matching of the new scan to all the past scans is then achieved by the scan-to-scan matching of the new scan to this best past scan. Although they have demonstrated capabilities in accurate matching, the approaches could still see accuracy issues without a loop closure as they do not either implement a powerful scan-to-scan matching or utilize all the past scans. In order for accuracy improvement the NDT technique was used to match the new scan to multiple recent past scans sequentially, which also improve accuracy by the point-to-ND matching.[13] Gutmann et.al. [10] presents a similar method that matches the new scan with the past K scans which improved the accuracy of the map immensely. Due to the need for matching to all the past scans for the best accuracy, Bosse et al. [16] introduced a subspace-to-map matching technique where the new scan is matched to all the past scans of a subspace of concern with any scan-to-scan matching technique and the subspaces are subsequently associated to each other for global mapping. This technique achieves the matching of the new scan to all the past scans, but the accuracy could still drop since the new scan points not in the subspace are not matched to the past scans. KinectFusion by Nicombe et al. [17] and the method proposed by Tomono [18] constructs a global model by doing global transform of all previously acquired scans, and uses signatures from that model for tracking the next robot pose. But these methods can be computationally expensive as the metric that is used for matching needs to be extracted from that global model at every iteration by using expensive techniques such as ray-casting in the case of KinectFusion.

Another advantage of most of these techniques for scan matching is that they can directly be extended to 3D mapping with minimal effort.[19]

This paper presents a grid-based scan-to-map matching technique for making accurate maps. The proposed technique matches the new scan to the previous scan using the conventional ICP technique, but further corrects its scan-to-scan matching error by representing the new scan globally through the estimation of the robot pose and matching it to the global map, namely the scan-to-map matching. In order to best correct the scan-to-scan matching error in the scan-to-map matching process, the proposed technique represents the map to match as a grid map with cells each with multiple NDs. The matching of the new scan to a ND, inspired by the NDT technique, avoids the point-to-point matching error of the ICP technique. The representation of a cell with multiple NDs further avoids errors in the scan-to-map matching process since the scan in a cell may not be matched to a single map ND in the cell. In addition, the new scan is also represented in terms of NDs, thereby characterizing the proposed scan-to-map matching technique as a novel ND-to-ND matching technique. Matching the similarity of the normal distributions using the Kullback-Leibler (KL) divergence, which makes the proposed technique is theoretically sound. In addition, the ND-to-ND matching has the potential for additional accuracy improvement and further dramatically improving computational time and novel over all the point-to-X and NDT-to-X matching techniques.[19]

The proposed technique is able to exploit the high accuracy of sensors such as laser range finders to produce highly accurate maps in fairly large indoor environments even without the benefit of SLAM or other inputs such as odometry. Therefore the authors feel that the proposed technique has high potential of being applicable to full SLAM in even larger environments, which is proposed as future work.

This paper is organized as follows. The next section explains the fundamental formulations of two existing scan-to-scan matching techniques, the ICP and the NDT, which are the most associated with the proposed technique. The proposed grid-based scan-to-map matching technique is presented in the Section 3 including the globally defined map and the ND-to-ND matching. Section 4 presents a number of experiments which investigate the performance of the proposed technique based on the error accumulation, and the effectiveness of multiple ND representation together with the applicability to large indoor environments. Finally, Section 5 summarizes conclusions and proposes future works.

2. Scan-to-scan matching techniques

2.1. Overview

Figure 1 shows the schematic diagram of the general scan-to-scan matching technique. When scans are taken by a range sensor on a moving robot, they are sequentially obtained with respect to different robot coordinate systems. Let ${}^{\{R^-\}}Z_{k-1} = \left\{ {}^{\{R^-\}}\mathbf{z}_{k-1}^i | \forall i \in \{1, \ldots, m\} \right\}$ be the previous scan in the previous robot coordinate system, and ${}^{\{R\}}Z_k = \left\{ {}^{\{R\}}\mathbf{z}_k^i | \forall i \in \{1, \ldots, m\} \right\}$ be the new scan in the new robot coordinate system, where k is the time step, m is the number of points in the scan. $\{R^-\}$ and $\{R\}$ denote the previous robot coordinate system and the new robot coordinate system. Given the two scans, a scan-to-scan matching technique iteratively finds relative transformation parameters, ${}^{\{R^-\}}_{\{R\}}\mathbf{p}_k = [t_k^x, t_k^y, \phi_k]^\mathsf{T}$, composed of a translation, $[t_k^x, t_k^y]^\mathsf{T}$, and a rotation, ϕ_k, between the two coordinate systems by locally matching the two scans. The first step is to transform the new scan in the new robot coordinate system to that in the previous coordinate system using the currently guessed transformation parameters.

Note that the initial transformation parameters can be estimated from readings of other sensors such as an odometer, or can be set as zeros assuming that the two scans are close enough. Mathematically, the transformation of a point of the new scan in the new robot coordinate system to that in the previous robot coordinate system is performed as,

$$
\begin{aligned}
{}^{\{R^-\}}\mathbf{z}_k^i \left({}^{\{R^-\}}_{\{R\}}\mathbf{p}_k \right) &= \mathbf{R}\left(\phi_k\right) {}^{\{R\}}\mathbf{z}_k^i + \mathbf{t}_k \\
&= \begin{bmatrix} \cos(\phi_k) & -\sin(\phi_k) \\ \sin(\phi_k) & \cos(\phi_k) \end{bmatrix} \begin{bmatrix} {}^{\{R\}}z_k^{x_i} \\ {}^{\{R\}}z_k^{y_i} \end{bmatrix} \\
&\quad + \begin{bmatrix} t_k^x \\ t_k^y \end{bmatrix}
\end{aligned}
\tag{1}
$$

where $\mathbf{t}_k = [t_k^x, t_k^y]^\mathsf{T}$, and ${}^{\{R\}}\mathbf{z}_k^i = [{}^{\{R\}}z_k^{x_i}, {}^{\{R\}}z_k^{y_i}]^\mathsf{T}$. Then, each point of the new scan ${}^{\{R^-\}}Z_k$ is associated with ${}^{\{R^-\}}Z_{k-1}$ and finds the correspondence set, ${}^{\{R^-\}}Y_k = \left\{ {}^{\{R^-\}}Y_k^i | \forall i \in \{1, \ldots, m\} \right\}$, to which the new scan is to be compared. Note that the number of corresponding elements may be less than m if any new scan point does not find a corresponding element. The new transformation parameters are finally computed by minimizing the error metric between the new scan and the correspondence, or equivalently maximizing the score function indicating how good the scan-to-scan matching is. The way of finding ${}^{\{R^-\}}Y_k$ and of computing the transformation parameters varies by scan-to-scan matching techniques, which will be detailed in the following subsections. The iterative identification of the transformation parameters

stops when the absolute value of the increment of computation is lower than the specified threshold value:

$$
\left| \Delta {}^{\{R^-\}}_{\{R\}}\mathbf{p}_k \right| < \delta
\tag{2}
$$

2.2. Iterative closest point technique

When the new scan is transformed to the coordinate system of the previous scan, the ICP scan-to-scan matching technique calculates the distance to all previous scan points from each new scan point and finds the corresponding point, i.e. ${}^{\{R^-\}}Y_k^i = {}^{\{R^-\}}\mathbf{y}_k^i$, that has the minimum distance. The corresponding point has the shortest distance to the new scan point:

$$
\begin{aligned}
&d({}^{\{R^-\}}\mathbf{z}_k^i, {}^{\{R^-\}}\mathbf{y}_k^i) \\
&= \min \left\{ d({}^{\{R^-\}}\mathbf{z}_k^i, {}^{\{R^-\}}\mathbf{z}_{k-1}^j) | \forall j \in \{1, \ldots, m\} \right\}
\end{aligned}
\tag{3}
$$

where $d(\cdot, \cdot)$ denotes a distance between two points. Given the correspondence the derivation of ${}^{\{R^-\}}_{\{R\}}\mathbf{p}_k$ is equivalent to solving the minimization problem of the error metric:

$$
\begin{aligned}
e({}^{\{R^-\}}_{\{R\}}\mathbf{p}_k) &= \sum_i \left\| {}^{\{R^-\}}\mathbf{y}_k^i - (\mathbf{R}(\phi_k){}^{\{R\}}\mathbf{z}_k^i + \mathbf{t}_k) \right\|^2 \\
&\rightarrow \min_{{}^{\{R^-\}}_{\{R\}}\mathbf{p}_k}
\end{aligned}
\tag{4}
$$

The ICP technique solves the minimization problem using the singular value decomposition (SVD).[20] The means of the new scan and its corresponding point set are first computed as:

$$
{}^{\{R\}}\bar{\mathbf{z}}_k = \frac{1}{m} \sum_i {}^{\{R\}}\mathbf{z}_k^i, \quad {}^{\{R^-\}}\bar{\mathbf{y}}_k = \frac{1}{m} \sum_i {}^{\{R^-\}}\mathbf{y}_k^i
\tag{5}
$$

Defining ${}^{\{R\}}\mathbf{a}_k^i = {}^{\{R\}}\mathbf{z}_k^i - {}^{\{R\}}\bar{\mathbf{z}}_k$ and ${}^{\{R^-\}}\mathbf{b}_k^i = {}^{\{R^-\}}\mathbf{y}_k^i - {}^{\{R^-\}}\bar{\mathbf{y}}_k$, the error metric, $e({}^{\{R^-\}}_{\{R\}}\mathbf{p}_k)$, in Equation (4) can be then rewritten as:

$$
\begin{aligned}
e({}^{\{R^-\}}_{\{R\}}\mathbf{p}_k) &= \sum_{i=1}^{n_c} \left\| {}^{\{R^-\}}\mathbf{b}_k^i - \mathbf{R}(\phi_k){}^{\{R\}}\mathbf{a}_k^i \right. \\
&\quad \left. + ({}^{\{R^-\}}\bar{\mathbf{y}}_k - \mathbf{R}(\phi_k){}^{\{R\}}\bar{\mathbf{z}}_k - \mathbf{t}_k) \right\|^2
\end{aligned}
\tag{6}
$$

Decoupling the rotation and the translation, the substitution of ${}^{\{R^-\}}\bar{\mathbf{y}}_k - \mathbf{R}(\phi_k){}^{\{R\}}\bar{\mathbf{z}}_k - \mathbf{t}_k = \mathbf{0}$ into Equation (6) yields:

$$
e({}^{\{R^-\}}_{\{R\}}\mathbf{p}_k) = \sum_{i=1}^{n_c} \left\| {}^{\{R^-\}}\mathbf{b}_k^i - \mathbf{R}(\phi_k){}^{\{R\}}\mathbf{a}_k^i \right\|^2
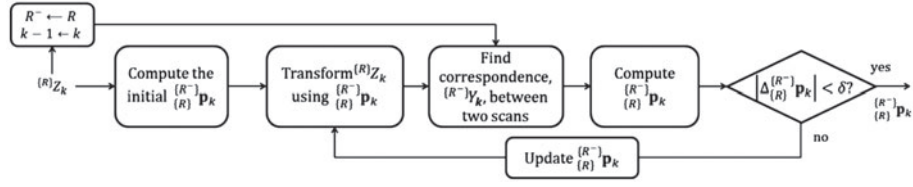$$

Figure 1. Scan-to-scan matching process.

$$= \sum_{i=1}^{n_c} \left\| {}^{\{R^-\}}\mathbf{b}_k^i \right\|^2 + \sum_{i=1}^{n_c} \left\| {}^{\{R\}}\mathbf{a}_k^i \right\|^2 - 2\mathrm{tr}(\mathbf{R}(\phi_k)\mathbf{N}_k) \tag{7}$$

where $\mathbf{N}_k = \sum_{i=1}^{n_c} {}^{\{R\}}\mathbf{a}_k^i {}^{\{R^-\}}\mathbf{b}_k^i{}^{\mathsf{T}}$. In the above equation, the error metric is minimized when $\mathrm{tr}(\mathbf{R}(\phi_k)\mathbf{N}_k)$ is maximized. Decomposing $\mathbf{N}_k$ by the SVD into $\mathbf{N}_k = \mathbf{U}_k\mathbf{D}_k\mathbf{V}_k^{\mathsf{T}}$, the transformation matrices, $\mathbf{R}(\phi_k)$, and $\mathbf{t}_k$ are finally given by

$$\mathbf{R}(\phi_k) = \mathbf{V}_k\mathbf{U}_k^{\mathsf{T}}, \quad \mathbf{t}_k = \bar{\mathbf{y}}_k - \mathbf{R}(\phi_k)\bar{\mathbf{z}}_k \tag{8}$$

where $\mathbf{U}_k$ and $\mathbf{V}_k$ are real or complex unitary matrices, and $\mathbf{D}_k$ is a rectangular diagonal matrix with nonnegative real number entries.[20] From $\mathbf{R}(\phi_k)$, the orientational transformation parameter, ϕ_k, can be derived as:

$$\phi_k = \mathrm{atan2}(R_{21}, R_{11}) \tag{9}$$

where R_{ij} is the entry of the matrix $\mathbf{R}$ in the ith row and the jth column.

It is to be noted that ICP has been used and researched on for many years and there are multiple variants of ICP available in the literature. But throughout this paper we use the more commonly used generic version of ICP, as it's not one of the strengths of this paper.

2.3. Normal distribution transform technique

Unlike the ICP technique the NDT scan-to-scan matching technique compares each new scan point to a ND. Since the NDT technique maps ${}^{\{R^-\}}\mathbf{z}_k^i$ onto a grid space having cells each represented with a ND. The NDT technique first defines a grid space with respect to the previous robot coordinate system and derives a ND for each grid cell after identifying ${}^{\{R^-\}}Z_{k-1}$ on the space as shown in Figure 2. For the jth cell, the mean and covariance matrix are computed by

$$^{\{R^-\}}\bar{\mathbf{z}}_{k-1}^j = \frac{1}{m_{k-1}^j} \sum_{i=1}^{m_{k-1}^j} {}^{\{R^-\}}\mathbf{z}_{k-1}^{j_i} \tag{10}$$

$$^{\{R^-\}}\bar{\Sigma}_{k-1}^j = \frac{1}{m_{k-1}^j} \sum_{i=1}^{m_{k-1}^j} \left({}^{\{R^-\}}\mathbf{z}_{k-1}^{j_i} - {}^{\{R^-\}}\bar{\mathbf{z}}_{k-1}^j \right) \left({}^{\{R^-\}}\mathbf{z}_{k-1}^{j_i} - {}^{\{R^-\}}\bar{\mathbf{z}}_{k-1}^j \right)^{\mathsf{T}} \tag{11}$$

where ${}^{\{R^-\}}\mathbf{z}_{k-1}^{j_i}$ is the ith point of the previous scan in the jth cell, and m_{k-1}^j is the number of scan points in the cell.

After transforming every new scan point using the currently guessed transformation parameters, each point is located in some grid cell. If ${}^{\{R^-\}}\mathbf{z}_k^{j_i}$ sees a ND created by the previous scan in the cell, the correspondence or the properties of the ND, i.e. ${}^{\{R^-\}}Y_k^i = \left\{ {}^{\{R^-\}}\hat{\mathbf{z}}_k^{j_i}, {}^{\{R^-\}}\hat{\Sigma}_k^{j_i} \right\}$, are those of the ND of the previous scan:

$$^{\{R^-\}}\hat{\mathbf{z}}_k^{j_i} \leftarrow {}^{\{R^-\}}\bar{\mathbf{z}}_{k-1}^j, \quad {}^{\{R^-\}}\hat{\Sigma}_k^{j_i} \leftarrow {}^{\{R^-\}}\bar{\Sigma}_{k-1}^j \tag{12}$$

Upon completion of the identification of the correspondence, the derivation of ${}^{\{R^-\}}_{\{R\}}\mathbf{p}_k$ is equivalent to solving the maximization problem of the score function, $s({}^{\{R^-\}}_{\{R\}}\mathbf{p}_k)$:

$$s\left({}^{\{R^-\}}_{\{R\}}\mathbf{p}_k\right)$$
$$= \sum_{j}\sum_{i=1}^{m_k^j} \exp \frac{-({}^{\{R^-\}}\mathbf{z}_k^{j_i} - {}^{\{R^-\}}\hat{\mathbf{z}}_k^{j_i})^{\mathsf{T}} ({}^{\{R^-\}}\hat{\Sigma}_k^{j_i})^{-1} ({}^{\{R^-\}}\mathbf{z}_k^{j_i} - {}^{\{R^-\}}\hat{\mathbf{z}}_k^{j_i})}{2}$$
$$\to \max_{{}^{\{R^-\}}_{\{R\}}\mathbf{p}_k} \tag{13}$$

Since the score function is the sum of piecewise smooth functions, a standard quadratic optimization method can be used. Applying Newton's method, ${}^{\{R^-\}}_{\{R\}}\mathbf{p}_k$ is iteratively computed by the increment $\Delta{}^{\{R^-\}}_{\{R\}}\mathbf{p}_k$:

$$\Delta{}^{\{R^-\}}_{\{R\}}\mathbf{p}_k = -\mathbf{H}_k^{-1}\,\mathbf{g}_k \tag{14}$$

where $\mathbf{H}_k$ and $\mathbf{g}_k$ are the sums of the Hessian, $\tilde{\mathbf{H}}_k$, and the gradient, $\tilde{\mathbf{g}}_k$, of the score function:

$$\mathbf{H}_k = \sum_{j}\sum_{i=1}^{m_k^j} \tilde{\mathbf{H}}_k^{j_i}, \quad \mathbf{g}_k = \sum_{j}\sum_{i=1}^{m_k^j} \tilde{\mathbf{g}}_k^{j_i} \tag{15}$$
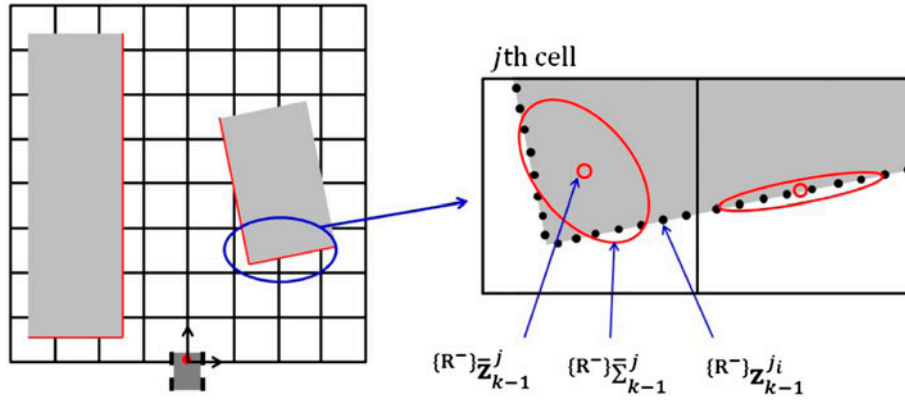
Figure 2. Two-dimensional grid space overlaped with the environment (left) and the ND based on scan points in the jth cell (right).
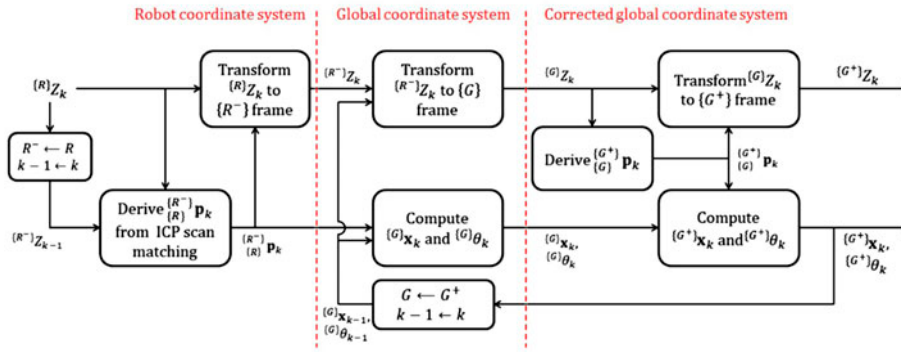

Figure 3. Proposed grid-based scan-to-map matching technique.

For the ith new scan point $^{\{R^-\}}z_k^{j_i}$ in the jth cell, the gradient vector $\tilde{g}_k^{j_i}$ is given by

$$
\begin{aligned}
&\tilde{g}_k^{j_i}\left(^{\{R^-\}}z_k^{j_i}, ^{\{R^-\}}\hat{z}_k^{j_i}, ^{\{R^-\}}\hat{\Sigma}_k^{j_i}\right) \\
&= \tilde{z}_k^{\mathsf{T}}\left(^{\{R^-\}}\hat{\Sigma}_k^{j_i}\right)^{-1}\frac{\partial \tilde{z}_k}{\partial(t_k^x, t_k^y, \phi_k)} \\
&\exp\frac{-\tilde{z}_k^{\mathsf{T}}(^{\{R^-\}}\hat{\Sigma}_k^{j_i})^{-1}\tilde{z}_k}{2}
\end{aligned}
\tag{16}
$$

and the m, n entry of $\tilde{\mathbf{H}}_k^{j_i}$ is computed by

$$
\begin{aligned}
&\tilde{H}_k^{j_i}\left(^{\{R^-\}}z_k^{j_i}, ^{\{R^-\}}\hat{z}_k^{j_i}, ^{\{R^-\}}\hat{\Sigma}_k^{j_i}\right)[m, n] \\
&= -\exp\frac{-\tilde{z}_k^{\mathsf{T}}\left(^{\{R^-\}}\hat{\Sigma}_k^{j_i}\right)^{-1}\tilde{z}_k}{2} \\
&\times\left\{-\tilde{z}_k^{\mathsf{T}}\left(^{\{R^-\}}\hat{\Sigma}_k^{j_i}\right)^{-1}\frac{\partial \tilde{z}_k}{\partial p_k[m]}\right\} \\
&\times\left\{-\tilde{z}_k^{\mathsf{T}}\left(^{\{R^-\}}\hat{\Sigma}_k^{j_i}\right)^{-1}\frac{\partial \tilde{z}_k}{\partial p_k[n]}\right\} \\
&+\left\{-\tilde{z}_k^{\mathsf{T}}\left(^{\{R^-\}}\hat{\Sigma}_k^{j_i}\right)^{-1}\frac{\partial^2 \tilde{z}_k}{\partial p_k[m]\partial p_k[n]}\right\} \\
&+\left\{-\frac{\partial \tilde{z}^{\mathsf{T}}}{\partial p_k[n]}(^{\{R^-\}}\hat{\Sigma}_k^{j_i})^{-1}\frac{\partial \tilde{z}_k}{\partial p_k[m]}\right\}
\end{aligned}
\tag{17}
$$

where $\tilde{z}_k = ^{\{R^-\}}z_k^{j_i} - ^{\{R^-\}}\hat{z}_k^{j_i}$ and the first and second partial derivative of $\tilde{z}_k$ can be derived as

$$
\begin{aligned}
\frac{\partial \tilde{z}_k}{\partial(t_k^x, t_k^y, \phi_k)} &= \left[\frac{\partial \tilde{z}_k}{\partial p_k[1]}, \frac{\partial \tilde{z}_k}{\partial p_k[2]}, \frac{\partial \tilde{z}_k}{\partial p_k[3]}\right] \\
&= \begin{bmatrix} 1 & 0 & -^{\{R\}}z_k^{x_i}\sin(\phi_k) - ^{\{R\}}z_k^{y_i}\cos(\phi_k) \\ 0 & 1 & ^{\{R\}}z_k^{x_i}\cos(\phi_k) - ^{\{R\}}z_k^{y_i}\sin(\phi_k) \end{bmatrix}
\end{aligned}
\tag{18}
$$

$$
\frac{\partial^2 \tilde{z}_k}{\partial p_k[m]\partial p_k[n]}
= \begin{cases} \begin{bmatrix} -^{\{R\}}z_k^{x_i}\cos(\phi_k) + ^{\{R\}}z_k^{y_i}\sin(\phi_k) \\ -^{\{R\}}z_k^{x_i}\sin(\phi_k) - ^{\{R\}}z_k^{y_i}\cos(\phi_k) \end{bmatrix} & \text{for } m=n=3 \\ \begin{bmatrix} 0 \\ 0 \end{bmatrix} & \text{otherwise} \end{cases}
\tag{19}
$$

When $\Delta_{\{R\}}^{\{R^-\}}\mathbf{p}_k$ is computed, $_{\{R\}}^{\{R^-\}}\mathbf{p}_k$ is then updated by:

$$
_{\{R\}}^{\{R^-\}}\mathbf{p}_k \leftarrow {}_{\{R\}}^{\{R^-\}}\mathbf{p}_k + \Delta_{\{R\}}^{\{R^-\}}\mathbf{p}_k
\tag{20}
$$

As can be seen in this section, the concept of the ICP is very simple and easy to implement, but it requires the exact point-to-point matching. Therefore, finding $_{\{R\}}^{\{R^-\}}\mathbf{p}_k$ can be significantly affected by wrong correspondence,

and the point-to-point error metric tends to converge to local minima. The NDT avoids this point-to-point matching problem by using a collectively represented correspondence of ND as in Equation (13). However, due to its representation of the scan on the grid space, the performance of the NDT relies on the size of the grid and the initial guess.

3. The proposed grid-based scan-to-map matching

3.1. Overall process

Figure 3 shows the overall process of the proposed grid-based scan-to-map matching technique which is based on the NDT's grid-based matching. Instead of the previous scan, the proposed technique matches the new scan to the globally defined map which is an accumulation of new scans after the scan-to-map matching. When the new scan ${}^{\{R\}}Z_k$ is obtained, the proposed technique first performs the ICP scan-to-scan matching to derive the transformation parameters, ${}^{\{R^-\}}_{\{R\}}\mathbf{p}_k^{\mathrm{ICP}}$, and transforms each new scan point in the $\{R\}$ coordinate system to that in the $\{R^-\}$ coordinate system:

$$ {}^{\{R^-\}}\mathbf{z}_k^i({}^{\{R^-\}}_{\{R\}}\mathbf{p}_k^{\mathrm{ICP}}) = \mathbf{R}(\phi_k^{\mathrm{ICP}}) \, {}^{\{R\}}\mathbf{z}_k^i + \mathbf{t}_k^{\mathrm{ICP}} \qquad (21) $$

where ${}^{\{R^-\}}_{\{R\}}\mathbf{p}_k^{\mathrm{ICP}} = [\mathbf{t}_k^{\mathrm{ICP}\mathsf{T}}, \phi_k^{\mathrm{ICP}}]^\mathsf{T}$.

Having the new scan matched to the previous scan, each new scan point ${}^{\{R^-\}}\mathbf{z}_k^i$ is further transformed to that in the global coordinate system, $\{G\}$, using the robot pose estimated at the previous time step in the $\{G\}$ coordinate system:

$$ {}^{\{G\}}\mathbf{z}_k^i = \mathbf{R}({}^{\{G\}}\theta_{k-1}){}^{\{R^-\}}\mathbf{z}_k^i + {}^{\{G\}}\mathbf{x}_{k-1} \qquad (22) $$

where ${}^{\{G\}}\mathbf{x}_{k-1} = [{}^{\{G\}}x_{k-1}, {}^{\{G\}}y_{k-1}]^\mathsf{T}$ and ${}^{\{G\}}\theta_{k-1}$ are the robot pose in the global coordinate system estimated at time step $k-1$. The iterative estimation of the robot pose in the global coordinate system is performed by considering the robot movement, ${}^{\{R^-\}}\mathbf{x}_k$ and ${}^{\{R^-\}}\theta_k$, which is equivalent to $\mathbf{t}_k^{\mathrm{ICP}}$ and ϕ_k^{ICP}, respectively:

$$ \begin{aligned} {}^{\{G\}}\mathbf{x}_k &= \mathbf{R}\left({}^{\{G\}}\theta_{k-1}\right){}^{\{R^-\}}\mathbf{x}_k + {}^{\{G\}}\mathbf{x}_{k-1} \\ &= \mathbf{R}\left({}^{\{G\}}\theta_{k-1}\right)\mathbf{t}_k^{\mathrm{ICP}} + {}^{\{G\}}\mathbf{x}_{k-1} \\ {}^{\{G\}}\theta_k &= {}^{\{R^-\}}\theta_k + {}^{\{G\}}\theta_{k-1} = \phi_k^{\mathrm{ICP}} + {}^{\{G\}}\theta_{k-1} \end{aligned} \qquad (23) $$

This global coordinate system is, however, incorrectly located due to the misalignment of the previous robot coordinate system by the ICP scan-to-scan matching as well as the error of estimation of the robot pose. Once the

new scan is transformed to the $\{G\}$ coordinate system, the proposed technique iteratively matches the new scan to the map in the $\{G^+\}$ coordinate system, which is the global coordinate system corrected by the proposed technique from the original guess of the global coordinate system, and derives the new scan in the $\{G^+\}$ coordinate system:

$$ {}^{\{G^+\}}\mathbf{z}_k^i \left({}^{\{G^+\}}_{\{G\}}\mathbf{p}_k\right) = \mathbf{R}(\phi_k) \, {}^{\{G\}}\mathbf{z}_k^i + \mathbf{t}_k \qquad (24) $$

where ${}^{\{G^+\}}_{\{G\}}\mathbf{p}_k = [\mathbf{t}_k^\mathsf{T}, \phi_k]^\mathsf{T}$ is the error correction parameters, or the scan-to-map matching transformation parameters, and transforms the new scan to the corrected global coordinate system. The derivation of the error correction parameters is detailed in the next subsections. Simultaneously, the robot pose in the $\{G\}$ coordinate system is also corrected by ${}^{\{G^+\}}_{\{G\}}\mathbf{p}_k$:

$$ \begin{aligned} {}^{\{G^+\}}\mathbf{x}_k &= \mathbf{R}(\phi_k){}^{\{G\}}\mathbf{x}_k + \mathbf{t}_k \\ {}^{\{G^+\}}\theta_k &= {}^{\{G\}}\theta_k + \phi_k \end{aligned} \qquad (25) $$

Because the misalignment of the previous robot coordinate system and the error of the robot pose are corrected by matching the new scan to the map, the proposed technique does not accumulate the scan-to-scan matching error as well as the pose estimation error.

Having the overall process of the scan-to-map matching identified, the representation of the grid map having multiple NDs in each cell is first defined in Section 3.2. In addition to the map NDs, the scan NDs are then derived from the new scan and paired with map NDs for scan-to-map matching. Section 3.3 presents the derivation of ${}^{\{G^+\}}_{\{G\}}\mathbf{p}_k$ via the ND-to-ND matching, whereas the update of the grid map using the derived ${}^{\{G^+\}}_{\{G\}}\mathbf{p}_k$ is detailed in Section 3.4. In order to simplify the notation the corrected global coordinate system, $\{G^+\}$, will be dropped from now on, and all notations without the coordinate system are considered as being in the corrected global coordinate system.

3.2. Grid map representation and selection of matching map normal distribution

Figure 4 illustratively shows the grid map with multiple map NDs in each cell together with the matching of new scan to the map NDs. As shown in the figure, the new scan of an object can be significantly different depending on where the scan is taken. The grid map with multiple NDs allows the matching of the new scan to a map ND irrespective of the robot pose. Mathematically, such a grid map updated up to time step $k-1$ for deriving ${}^{\{G^+\}}_{\{G\}}\mathbf{p}_k$ is

represented as

$$M_{1:k-1} = \left\{ M_{1:k-1}^j | \forall j \in \{1, \ldots, n^g\} \right\}$$

where $M_{1:k-1}^j$ is the property of the jth grid cell, and n^g is the number of grid cells. $M_{1:k-1}^j$ is given by

$$M_{1:k-1}^j = \left\{ M_{1:k-1}^{jl} = \left\{ \bar{z}_{1:k-1}^{jl}, \bar{\Sigma}_{1:k-1}^{jl}, m_{1:k-1}^{jl} \right\} | \forall l \in \left\{1, \ldots, n_{k-1}^j\right\} \right\}$$

where $M_{1:k-1}^{jl}$ is the property of the lth map ND in the jth cell with the mean, $\bar{z}_{1:k-1}^{jl}$, covariance matrix, $\bar{\Sigma}_{1:k-1}^{jl}$, and the total number of scan points, $m_{1:k-1}^{jl}$. n_k^j denotes the total number of map NDs in the jth cell.

With the new scan transformed to the $\{G\}$ coordinate system, the scan ND in the jth cell to match to a map ND in the same cell is derived simply as

$$^{\{G\}}\bar{z}_k^j = \frac{1}{m_k^j} \sum_{i=1}^{m_k^j} {}^{\{G\}}z_k^{j_i} \tag{26}$$

$$^{\{G\}}\bar{\Sigma}_k^j = \frac{1}{m_k^j} \sum_{i=1}^{m_k^j} \left({}^{\{G\}}z_k^{j_i} - {}^{\{G\}}\bar{z}_k^j \right) \left({}^{\{G\}}z_k^{j_i} - {}^{\{G\}}\bar{z}_k^j \right)^\top \tag{27}$$

where $^{\{G\}}z_k^{j_i}$ is the ith scan point in the jth cell and m_k^j is the total number of points in the jth cell. The selection of a matching map ND for the scan ND in the proposed technique starts with quantifying the similarity of the scan ND to each map ND in the same cell. The similarity can be computed by the KL divergence, D_{KL}, which is a mathematically solid method for measuring the distance between two probability distributions:

$$S\left(N({}^{\{G\}}\bar{z}_k^j, {}^{\{G\}}\bar{\Sigma}_k^j), N(\bar{z}_{1:k-1}^{jl}, \bar{\Sigma}_{1:k-1}^{jl})\right)$$
$$= -D_{KL}\left(N({}^{\{G\}}\bar{z}_k^j, {}^{\{G\}}\bar{\Sigma}_k^j) || N(\bar{z}_{1:k-1}^{jl}, \bar{\Sigma}_{1:k-1}^{jl})\right)$$
$$= -\frac{1}{2} \left\{ \text{tr}\left((\bar{\Sigma}_{1:k-1}^{jl})^{-1} {}^{\{G\}}\bar{\Sigma}_k^j \right) \right.$$
$$+ \left(\bar{z}_{1:k-1}^{jl} - {}^{\{G\}}\bar{z}_k^j \right)^\top \left(\bar{\Sigma}_{1:k-1}^{jl} \right)^{-1} \left(\bar{z}_{1:k-1}^{jl} - {}^{\{G\}}\bar{z}_k^j \right)$$
$$\left. - \ln \frac{\det ({}^{\{G\}}\bar{\Sigma}_k^j)}{\det (\bar{\Sigma}_{1:k-1}^{jl})} - \lambda \right\} \tag{28}$$

where $l \in \left\{1, \ldots, n_{k-1}^j, \right\}$, λ is the dimension of the NDs, and $N({}^{\{G\}}\bar{z}_k^j, {}^{\{G\}}\bar{\Sigma}_k^j)$ and $N(\bar{z}_{1:k-1}^{jl}, \bar{\Sigma}_{1:k-1}^{jl})$ are the scan ND and the lth map ND, respectively. Out of the map NDs the most similar one to the scan ND is that with the

highest similarity value:

$$S\left(N({}^{\{G\}}\bar{z}_k^j, {}^{\{G\}}\bar{\Sigma}_k^j), N(\bar{z}_{1:k-1}^{jl*}, \bar{\Sigma}_{1:k-1}^{jl*})\right)$$
$$= \min \left\{ S\left(N({}^{\{G\}}\bar{z}_k^j, {}^{\{G\}}\bar{\Sigma}_k^j), N(\bar{z}_{1:k-1}^{jl}, \bar{\Sigma}_{1:k-1}^{jl})\right) | \right.$$
$$\left. \forall l \in \left\{1, \ldots, n_{k-1}^j\right\} \right\} \tag{29}$$

The l^*th map ND is regarded as the matching map ND for the scan ND if the similarity is greater than the specified threshold value:

$$S\left(N({}^{\{G\}}\bar{z}_k^j, {}^{\{G\}}\bar{\Sigma}_k^j), N(\bar{z}_{1:k-1}^{jl*}, \bar{\Sigma}_{1:k-1}^{jl*})\right) > \gamma \tag{30}$$

Having the matching map ND identified for each scan ND, the derivation of $^{\{G^+\}}_{\{G\}}\mathbf{p}_k$ is possible by matching all the scan NDs to the corresponding matching map NDs.

3.3. Derivation of error correction parameters

Since a scan ND in each cell is compared to a matching map ND, the correspondence is derived not for every point but for every cell, i.e. $Y_k^j = \left\{\hat{z}_k^j, \hat{\Sigma}_k^j\right\}$. The correspondence for each cell is equivalent to the property of the matching map ND in the same cell:

$$\begin{aligned} \hat{z}_k^j &\leftarrow \bar{z}_{1:k-1}^{jl*} \\ \hat{\Sigma}_k^j &\leftarrow \bar{\Sigma}_{1:k-1}^{jl*} \end{aligned} \quad \text{if } S\left(N({}^{\{G\}}\bar{z}_k^j, {}^{\{G\}}\bar{\Sigma}_k^j), N(\bar{z}_{1:k-1}^{jl*}, \bar{\Sigma}_{1:k-1}^{jl*})\right) > \gamma \tag{31}$$

Note that a scan ND that does not have a matching map ND is not thus considered in the derivation of $^{\{G^+\}}_{\{G\}}\mathbf{p}_k$. Given the correspondence of the scan NDs, the derivation of $^{\{G^+\}}_{\{G\}}\mathbf{p}_k$ begins with the initial values set to $\mathbf{0}$ as it is valid to assume that the ICP scan-to-scan matching and the previous robot pose estimation is reasonably correct. The proposed technique first transforms the mean and covariance matrix of each scan ND to those in the $\{G^+\}$ coordinate system using the currently guessed $^{\{G^+\}}_{\{G\}}\mathbf{p}_k$:
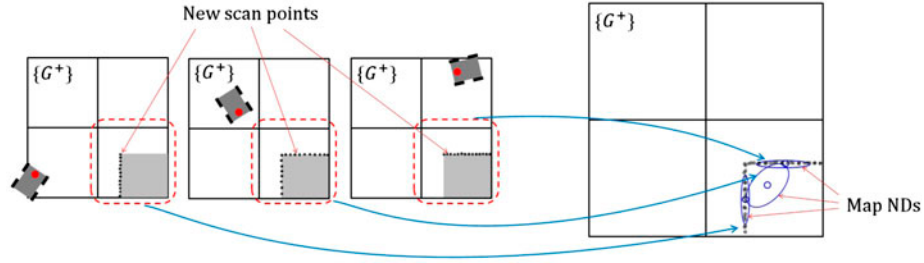
$$\bar{z}_k^j = \mathbf{R}(\phi_k) \, {}^{\{G\}}\bar{z}_k^j + \mathbf{t}_k \tag{32}$$
$$\bar{\Sigma}_k^j = \mathbf{R}(\phi_k) \, {}^{\{G\}}\bar{\Sigma}_k^j \, \mathbf{R}(\phi_k)^{-1} \tag{33}$$

With all the scan NDs and the matching map NDs described in the $\{G^+\}$ coordinate system, the transformation parameters $^{\{G^+\}}_{\{G\}}\mathbf{p}_k$ can be then computed by maximizing the objective function given by the sum of similarities between the scan NDs and the matching map NDs:

$$f(^{\{G^+\}}_{\{G\}}\mathbf{p}_k) = \sum_j S\left(N(\bar{z}_k^j, \bar{\Sigma}_k^j), N(\hat{z}_k^j, \hat{\Sigma}_k^j)\right) \tag{34}$$

Figure 4. The grid map is represented by multiple NDs (right) and new scans to be matched to the grid map (left).

The objective function of the proposed technique equally sums the similarities. In other words, similarities with a small number of scan points can be treated as equally as those with a large number of scan points. This could allow the proposed technique to match the new scan to the map more globally than the conventional point-to-X techniques. The ND-to-ND matching could also dramatically improve the computation time.

Although the analytical expressions of the gradient and the Hessian may be obtained for the objective function, the small-size optimization problem, with only three parameters for the two-dimensional scan, could be easily solved with the Newton method numerically computing the gradient and the Hessian.

3.4. The update of the grid map

The grid map is initially that with the first scan NDs, and this is regarded as the first grid map updated up to the previous time step. Given the mean and the covariance matrix of the scan ND of each cell in the {G} coordinate system shown in Equations (26) and (27), the proposed technique then updates map NDs in the same cell differently depending on whether there is a matching map ND. If there is a matching map ND, only this matching map ND is updated with the scan ND. The mean and covariance matrix of the matching map ND in the jth cell are updated according to the weighted mean formulation:

$$\bar{z}^{j_{l*}}_{1:k} = \frac{m^j_{1:k-1} \bar{z}^{j_{l*}}_{1:k-1} + m^j_k \bar{z}^j_k}{m^j_{1:k-1} + m^j_k} \tag{35}$$

$$\bar{\Sigma}^{j_{l*}}_{1:k} = \frac{m^j_{1:k-1} \bar{\Sigma}^{j_{l*}}_{1:k-1} + m^j_k \bar{\Sigma}^j_k}{m^j_{1:k-1} + m^j_k} \tag{36}$$

After the update, the number of scan points for the map ND is also updated:

$$m^{j_{l*}}_{1:k} = m^{j_{l*}}_{1:k-1} + m^j_k \tag{37}$$

On the other hand, if the scan ND has found no matching map ND, the scan ND is simply added as a new map ND without any update to the current map NDs. Let the index of the new map ND be $l^+ = n^j_k + 1$. The mean and the covariance matrix of the map ND in the jth cell are given by

$$\bar{z}^{j_{l+}}_{1:k} = \bar{z}^j_k \tag{38}$$

$$\bar{\Sigma}^{j_{l+}}_{1:k} = \bar{\Sigma}^j_k \tag{39}$$

The number of scan points of the map ND is, similarly, the number of scan points of the scan ND:

$$m^{j_{l+}}_{1:k} = m^j_k \tag{40}$$

After the scan ND is added, the number of the map NDs becomes $n^j_k \leftarrow n^j_k + 1$. The update of the grid map completes by applying the cell-wise update to all the grid cells.

4. Experimental results

This section is aimed at investigating the performance of the proposed scan-to-map matching technique and demonstrating the applicability of the proposed technique in real indoor environments. All experiments were conducted using a ground mobile robot with a forward-facing LRF, Hokuyo UTM-30lx, mounted on the robot (Figure 5). No other sensors such as an odometer and an IMU were used to estimate the pose of the robot and to build a map. In the first experiment, the performance of the proposed technique is investigated based on the position and orientation error seen from landmarks at every matching of the new scan to the map. The second experiment focuses on showing the effectiveness of multi-ND representation within a grid cell instead of having a single ND. Finally, the proposed technique is tested within a number of real indoor environments each of which is relatively large and unstructured. This includes the popular public domain dataset from Freiburg University Building 079.[21] Table 1 shows the parameters used in the experiments.
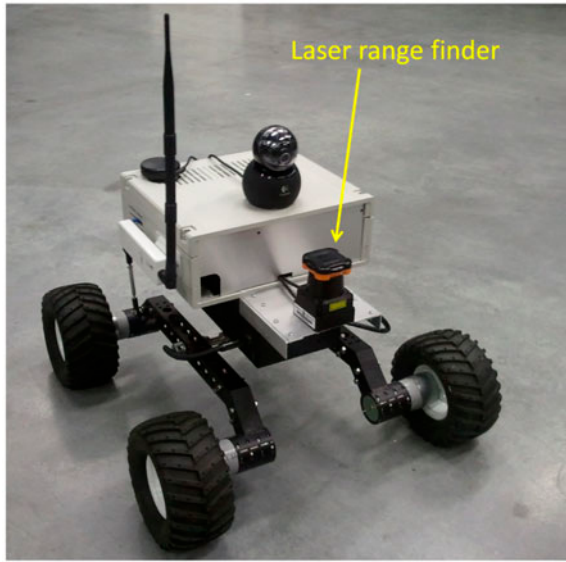
Figure 5. Ground mobile robot with laser range finder.

Table 1. Default parameters for the scan-to-map matching technique.

Parameters	Value
LRF scanning angle	0° to 180°
Scanning interval	0.5°
LRF scanning frequency	10 Hz
Grid cell size	$1m$ by $1m$
Threshold (δ) for the scan-to-scan matching technique	0.001
Threshold similarity (γ)	-2

4.1. Effect of the scan-to-map matching

Figure 6 shows the first experiment where there is a L-shaped object at the end of a corridor. The robot was initially located at the starting point which was known in the global coordinate system. In order to exclude environmental parameters that might have influence on the experiment, the environment was selected to be simple. In the experiment the robot observed the entire object at all time and was manually driven along two different paths, one of which was a straight line and the other was a curvature. The robot took 190 scans and 348 scans for the linear and the nonlinear motion, respectively. At every acquisition of the new scan, the robot performed the scan-to-map matching and every scan points were mapped onto the global coordinate system.

The left and right edge, and the center point of the object are considered as detectable features, and the position error $\epsilon_k^{f_i}$ of the ith feature at time step k is given by

$$\epsilon_k^{f_i} = \sqrt{\left(x_0^{f_i} - x_k^{f_i}\right)^2 + \left(y_0^{f_i} - y_k^{f_i}\right)^2} \qquad (41)$$

where f_i is either left or right edge (i.e. f_l or f_r), $\left[x_0^{f_i}, y_0^{f_i}\right]^\mathsf{T}$ is the initial position of f_i, and $\left[x_k^{f_i}, y_k^{f_i}\right]^\mathsf{T}$ represents the position mapped onto the corrected global coordinate system at time step k. The slope of the line connecting the center point and the left edge of the object is also calculated in order to to see the orientation error.

Figure 7(a) and (d) show accumulated scan points after the scan-to-map matching for the linear and nonlinear motion, respectively. To address the effect of the global correction by the proposed technique, the figure also shows scan point mapped onto the global coordinate system after sequential scan-to-scan matchings by the ICP and NDT technique, but without the global correction. As shown in the figure, the scans transformed by the proposed technique at different time steps are well aligned to one another. This indicates that the estimations of the transformation parameters by the local scan-to-scan matchings are well corrected by the global scan-to-map matching parameters. On the other hand, the scans transformed by the sequential scan-to-scan matchings without the global correction do not lie on the same position. From these results the effect of the global correction can be qualitatively verified. In both cases of the linear and nonlinear motion the ICP technique without the global correction generates relatively larger position error than the others. This is because the L-shaped corner is initially away and thus only captured by a small number of scan points. ICP match of the close-by walls represented by high-density scan points.

The error produced by NDT is due to the absence of global-correction

Usage of different ICP techniques available in literature, would exploit different features in the environment and yield slightly different results. For example using a point-to-line ICP for the particular example shown in Figure 7 would produce much better results as it exploints the lines in the environment. But if the environment was to be composed of curved lines, that technique would not produce good results. Therefore as explained in Section 2.2, the proposed algorithm uses generic ICP throughout.

Figures 8 and 9 quantitatively show the position error and the slope of the line connecting the center point and the left edge for the linear and nonlinear motion. As expected from the previous figure showing the accumulated scan points, the proposed technique has the smallest position errors for both motions. There is nearly no difference between the linear and nonlinear motion case and the error is consistent in its value regardless of time step. The slope does not change a lot with respect to time step, indicating that the orientation error is small and not accumulated with time. Note that these errors are caused not only by the matching process, but also

Table 2. Initial and final positions of left (LE) and right edges (RE) estimated by three methods and position errors.
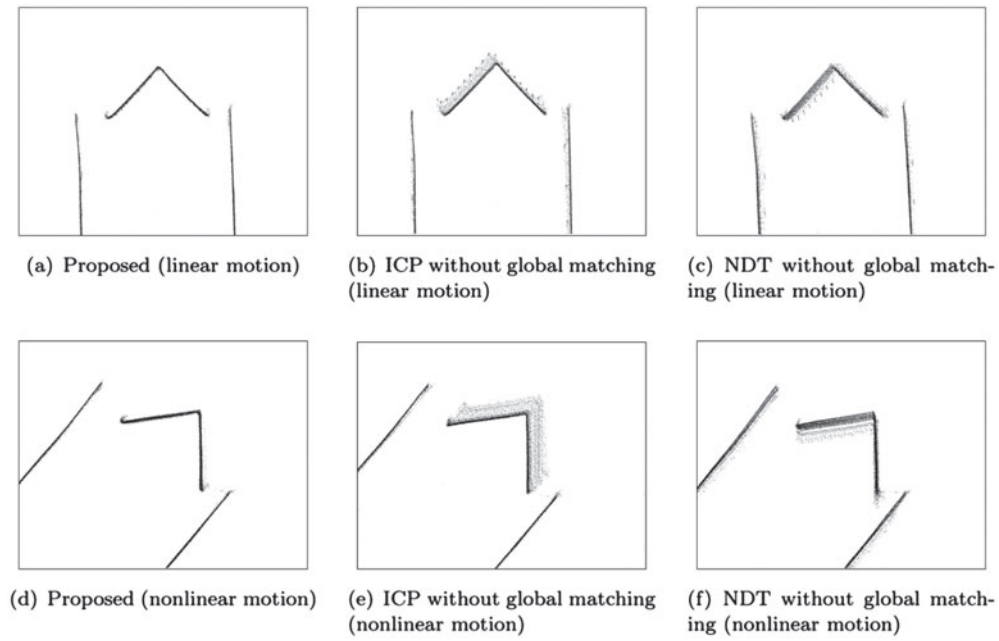
			Proposed	ICP	NDT
Linear motion					
	LE	$[x_0^{f_l}, y_0^{f_l}]$		$[-0.85,\ 9.64]$	
		$[x_{190}^{f_l}, y_{190}^{f_l}]$	$[-0.89, 9.64]$	$[-0.81, 9.45]$	$[-0.99, 9.73]$
		$\epsilon_{190}^{f_l}$ (m)	0.044	0.194	0.174
	RE	$[x_0^{f_r}, y_0^{f_r}]$		$[0.5,\ 9.69]$	
		$[x_{190}^{f_r}, y_{190}^{f_r}]$	$[0.52, 9.67]$	$[0.61, 9.45]$	$[0.42, 9.81]$
		$\epsilon_{190}^{f_r}$ (m)	0.025	0.26	0.141
Nonlinear motion					
	LE	$[x_0^{f_l}, y_0^{f_l}]$		$[5.92,\ 7.85]$	
		$[x_{348}^{f_l}, y_{348}^{f_l}]$	$[5.89, 7.82]$	$[5.71, 7.57]$	$[5.9, 8.04]$
		$\epsilon_{348}^{f_l}$ (m)	0.041	0.353	0.185
	RE	$[x_0^{f_r}, y_0^{f_r}]$		$[6.98, 6.97]$	
		$[x_{348}^{f_r}, y_{348}^{f_r}]$	$[6.94, 6.97]$	$[6.74, 6.71]$	$[6.95, 7.21]$
		$\epsilon_{348}^{f_r}$ (m)	0.034	0.358	0.233
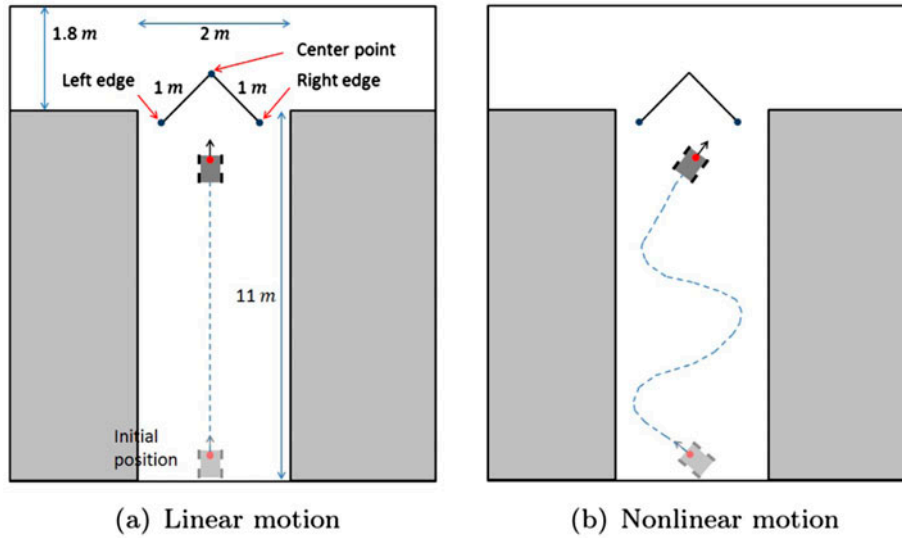
Figure 6. Robot takes scan images and matches them to previous scans and global map following (a) a straight line and (b) a curvature.
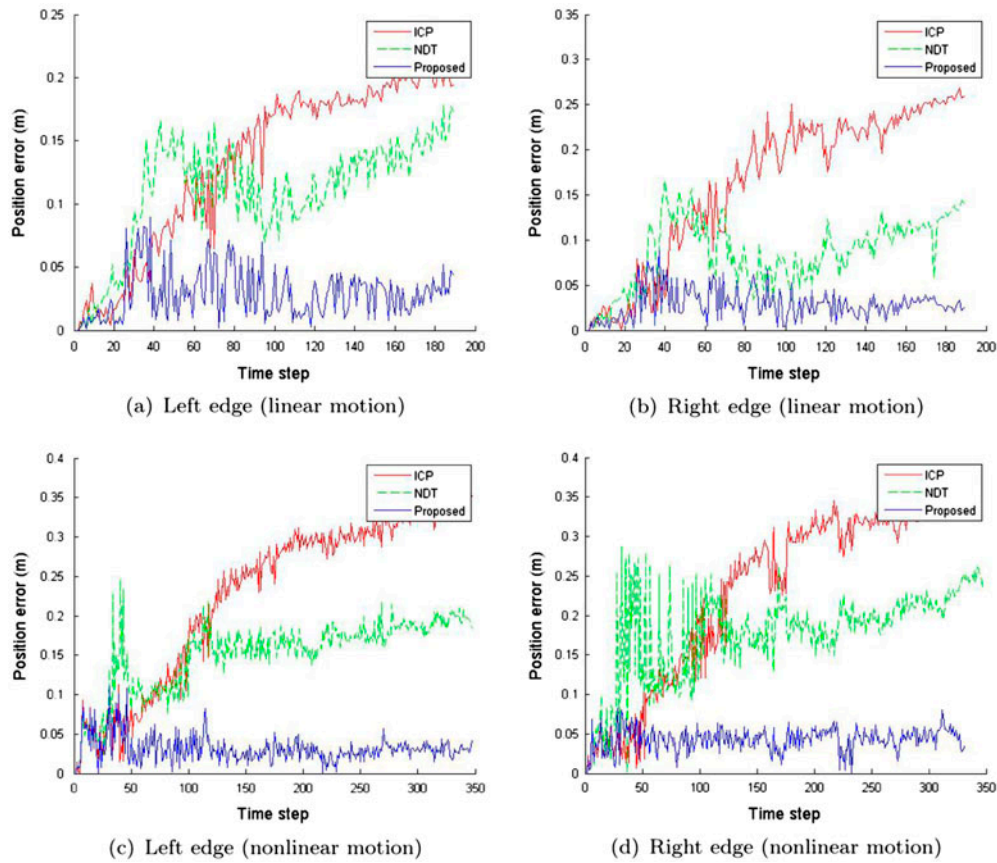
by the LRF with the scanning interval of 0.5° which observes the features at different positions for each scan. However, when new scans are matched only to their previous scans using the ICP and NDT technique, position errors become large. For both the linear and nonlinear motion the ICP scan-to-scan matching technique shows the largest position errors which increase with time. Position errors by the NDT scan-to-scan matching are less than the ICP, whereas, slopes at the initial and the last time step are slightly distinct, which implies there exist orientation errors. According to Table 2 the proposed technique has successfully removed the position errors by the scan-to-scan matchings with 6.67 and 4.62 times lesser errors than the ICP and NDT technique without the global correctness for the linear motion, and 9.62 and 5.64 times lesser errors for the nonlinear motion, respectively. It is important that the position errors do not accumulate with time step, but it stays within 4cm.

4.2. Effect of multiple normal distributions in a single cell

Having identified the effect of global correction capability by the proposed scan-to-map matching technique, this experiment investigates the effectiveness of maintaining multiple NDs instead of a ND in a single grid cell. In this experiment the robot possibly sees different parts of an object while it operates in a simulated environment, so that NDs from scan points in the jth cell can be largely different depending on the pose of the robot. The robot took 43 scans and two different grid maps were independently updated where the first grid map maintained multiple NDs within each grid cell whereas the other grid map contained a single ND in each cell. As shown in Figure 10(a), the robot could observe only one side of the corner at the initial time step, however, at time step k_1 it started seeing both sides of the corner. From time step k_2 the robot again observed only

(a) Proposed (linear motion) (b) ICP without global matching (linear motion) (c) NDT without global matching (linear motion)

(d) Proposed (nonlinear motion) (e) ICP without global matching (nonlinear motion) (f) NDT without global matching (nonlinear motion)

Figure 7. Accumulated scan points of the object by the proposed, ICP, NDT.

(a) Left edge (linear motion) (b) Right edge (linear motion)

(c) Left edge (nonlinear motion) (d) Right edge (nonlinear motion)

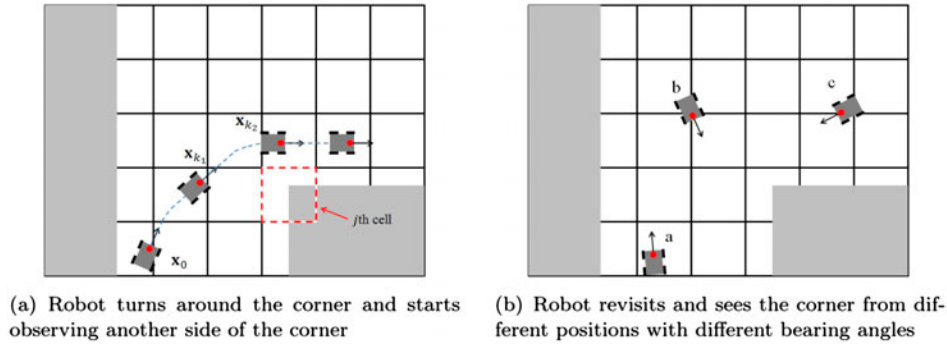**Figure 8.** Position errors seen at the left and right edge.

one side of the corner. With the two grid maps updated using the 43 scans, the robot was located at three different positions and the similarities between map NDs and the scan NDs are computed to find the most similar map NDT for each case (Figure 10(b)).

Figure 11(a) and (b) show the similarities between the scan NDs and the most similar map NDs in the jth cell when the grid map maintains multiple NDs and a single ND (includes methods such as [22,23]), respectively. For the multi-ND representation, when the robot first

(a) Linear motion

(b) Nonlinear motion
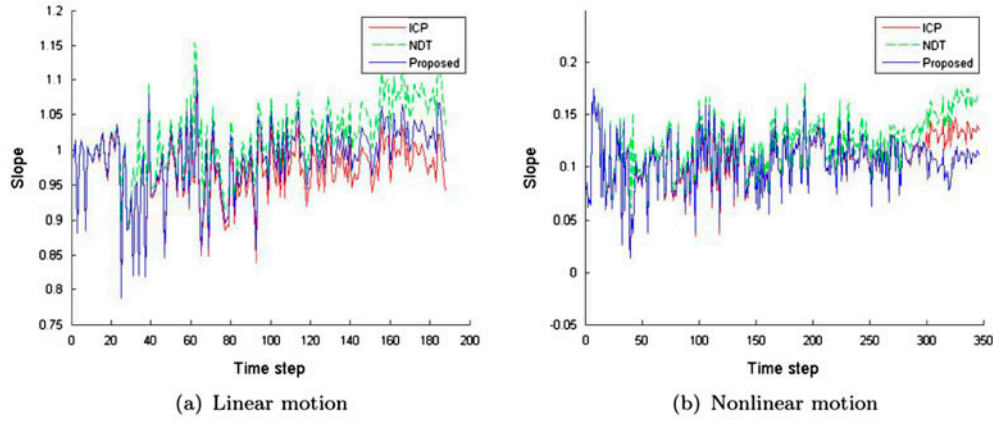
Figure 9. The slope of the line connecting the center point and the left edge of the object.

(a) Robot turns around the corner and starts observing another side of the corner

(b) Robot revisits and sees the corner from different positions with different bearing angles

Figure 10. Experiment for investigating the effect of multiple normal distributions in a single cell.

(a) Multi-ND representation

(b) Single-ND representation
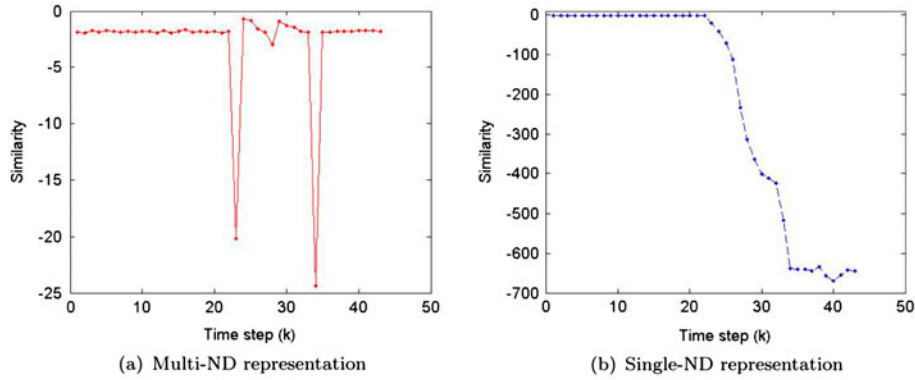
Figure 11. Similarity between the scan ND and the most similar map ND.

Table 3. Similarities between the scan ND and the map NDs.

| Position	Map ND 1	Map ND 2	Map ND 3	Map ND 4
a	−1.72	−2.89	−3.63	−98.37
b	−47.12	−1.12	−2.89	−61.38
c	−173.28	−4.27	−2.793	−1.732

observes different side of the corner at time step k_1 (i.e. $k_1 = 23$), the similarity drops down to −20.17. However, the similarity immediately goes up at the next time step after the scan ND at time step k_1 is added as a new map

ND. Similarly, at time step $k = 28$ and k_2 (i.e. $k_2 = 34$), when the similarity is lower than the threshold value, a new map NDs is added and the similarity increases again at the next time step. When the grid map is updated

using the single-ND representation, the similarity drops at time step k_1 and keeps decreasing since the existing map ND becomes less and less similar to the scan ND as the robot sees the other part of the corner more. The corresponding ND to the scan ND after time step k_1 are not appropriate, which may cause inaccurate matching between the scan ND and the map ND. Conclusively, the multi-ND representation in each cell can increase the accuracy of the scan-to-map matching.

After the grid map is updated by 43 scans using the multi-ND representation, the jth cell contains 4 map NDs. Having the grid map updated, Table 3 shows the similarities between the scan NDs and map NDs when the robot is placed at three different locations. When the robot is on Position (a) and obtains the new scan, the first map ND, updated by the first 22 scans, is selected as the most similar map ND to the scan ND with the similarity of -1.72. The last map ND, updated by the last 10 scans, has the minimum similarity to the scan ND. When the robot observes the corner from Position (b), both the second map ND, updated by the 23rd to 27th scans, and the third map ND, updated by the 28th to 33rd scans, are quite similar to the scan ND. Although the third map ND can be a good matching map ND for the scan ND, the second map ND is selected as the matching map ND, which enables the proposed technique to match the new scan more accurately to the grid map. When the robot is on Position (c), the result seems to be opposite to the first case, where the last map ND has the maximum similarity and the first map ND has the minimum value. Due to the multi-ND representation, the scan ND can find the exact matching map ND regardless of the pose of the robot. Since the potential strength of the proposed technique is its high accuracy by the multi-ND representation rather than the resolution of the grid space, this subsection finally identifies the accuracy of the proposed technique with respect to the threshold of similarity and the grid size. The experimental data with non-linear motion shown in Figure 6(b) was used with a different set of similarity threshold and grid size. The scholar error quantified was the average of the left and right error in (41). Figure 12 shows the error surface created with different sets of similarity threshold and grid size. It is first seen that the error is consistently small regardless of the grid size when the similarity threshold is small. This is because more map NDs can be created in each grid with small similarity threshold. Seen next is that the error is consistently small regardless of the threshold when the grid size is small. This is due to the fact that the grid size is small enough to be able to better capture the distribution of scan points. These results clearly indicate that the multi-ND representation yields the high accuracy strength of the proposed technique.
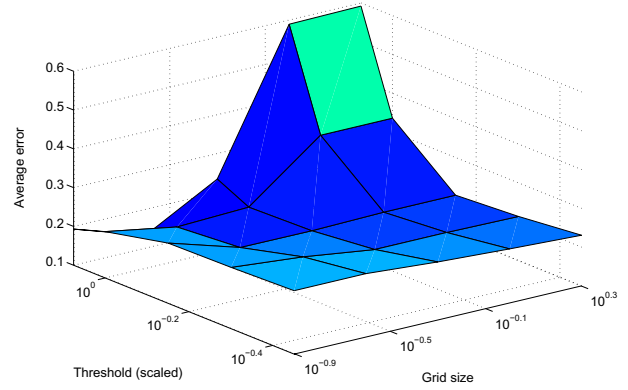
Figure 12. Represents the change in average error as per (41), with respect to the changes in grid size and threshold for multi-NDs.

4.3. Application: mapping large environments

This subsection demonstrates high accuracy and versatility of the proposed technique in solving the SLAM problem in indoor environments shown in Figure 13. The environments were partially structured or unstructured with static and mobile objects including walking people. The environments had long corridors (Environment 2 and 3), a large number of random shaped objects such as chairs and desks (Environment 1, 2, and 3), and large loops (Environment 1,2 and 3) to be closed. During the experiment the robot took 2489, 2426, and 3873 scans for Environment 1, 2, and 3, respectively. Similar to the first experiment, since the autonomous exploration was out of the scope of this paper, the robot was manually driven following pre-determined paths (i.e. a-e-b-c-d-e for Environment 1, a-b-c-d-c-b-a for Environment 2, a-b-a-c-a-d-e-d-f-g-d for Environment 3).

Figure 14 shows trajectories of the robot in the three environments estimated by the proposed technique. In order to emphasize the effect of the global correction made by the proposed technique, trajectories of the robot estimated by the two scan-to-scan matching techniques without the global correction are also plotted in the same figure. As shown in the figure when the pose of the robot is computed by the proposed technique the robot successfully closes loops in the environments, though this is not an active loop closure. For all the environment the total distances the robot traveled are 186 m, 258.2 m, and 374 m, and accumulated orientation changes of the robot are 3646°, 1949°, and 3729°, respectively. On the other hand, when the pose of the robot is estimated by the scan-to-scan matching techniques without the global correction, there exist deviations in the trajectories of the robot ending up with the failure in closing the loops.

In Figure 15 maps of all three environments built by the proposed technique are shown. The localization error can be qualitatively analyzed by seeing the quality
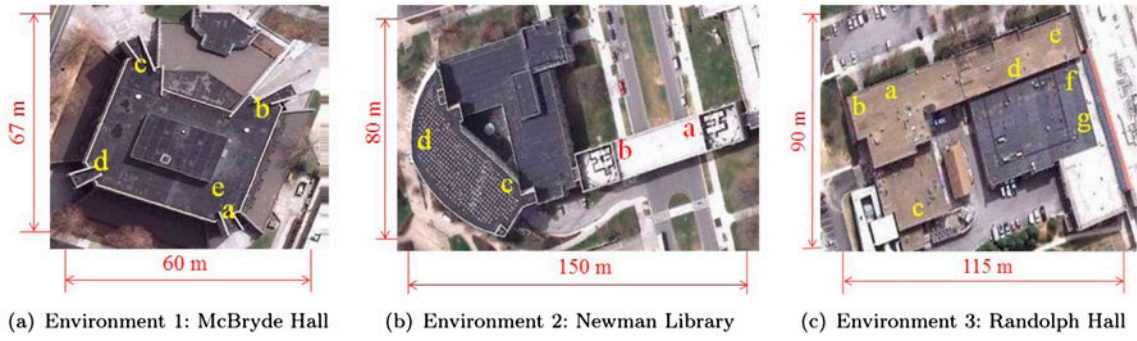
(a) Environment 1: McBryde Hall   (b) Environment 2: Newman Library   (c) Environment 3: Randolph Hall

Figure 13. Test Environments.

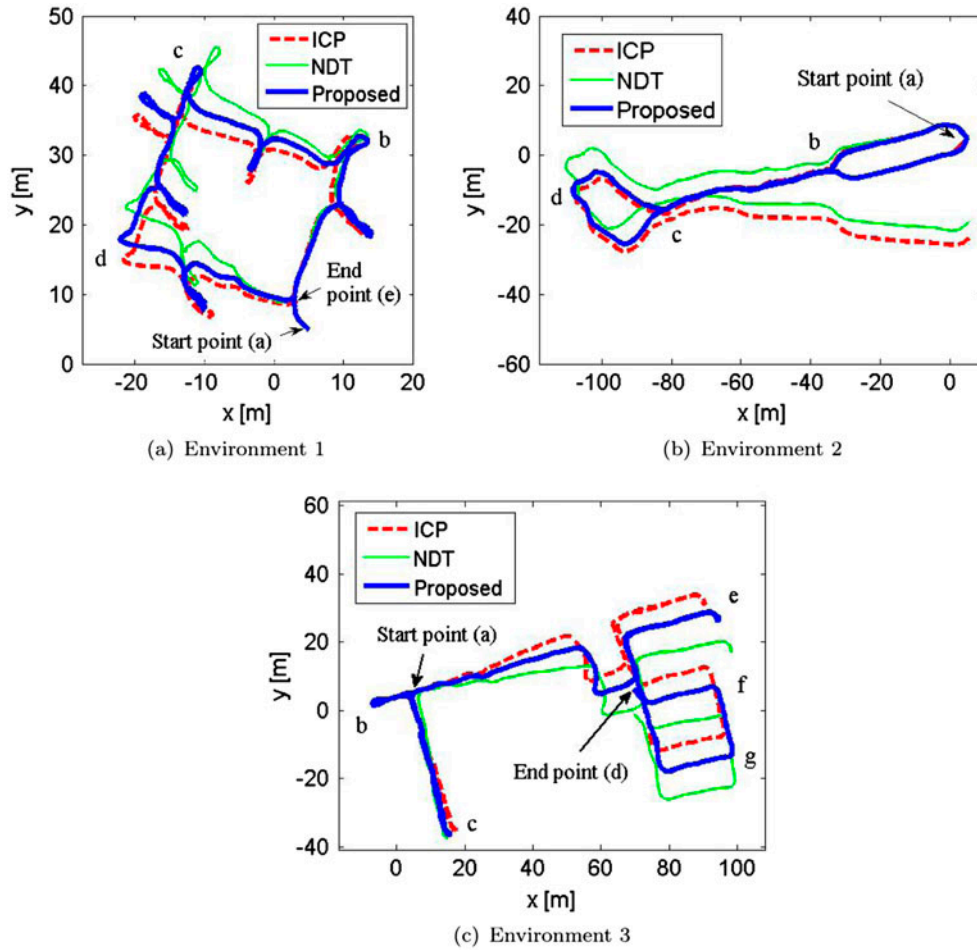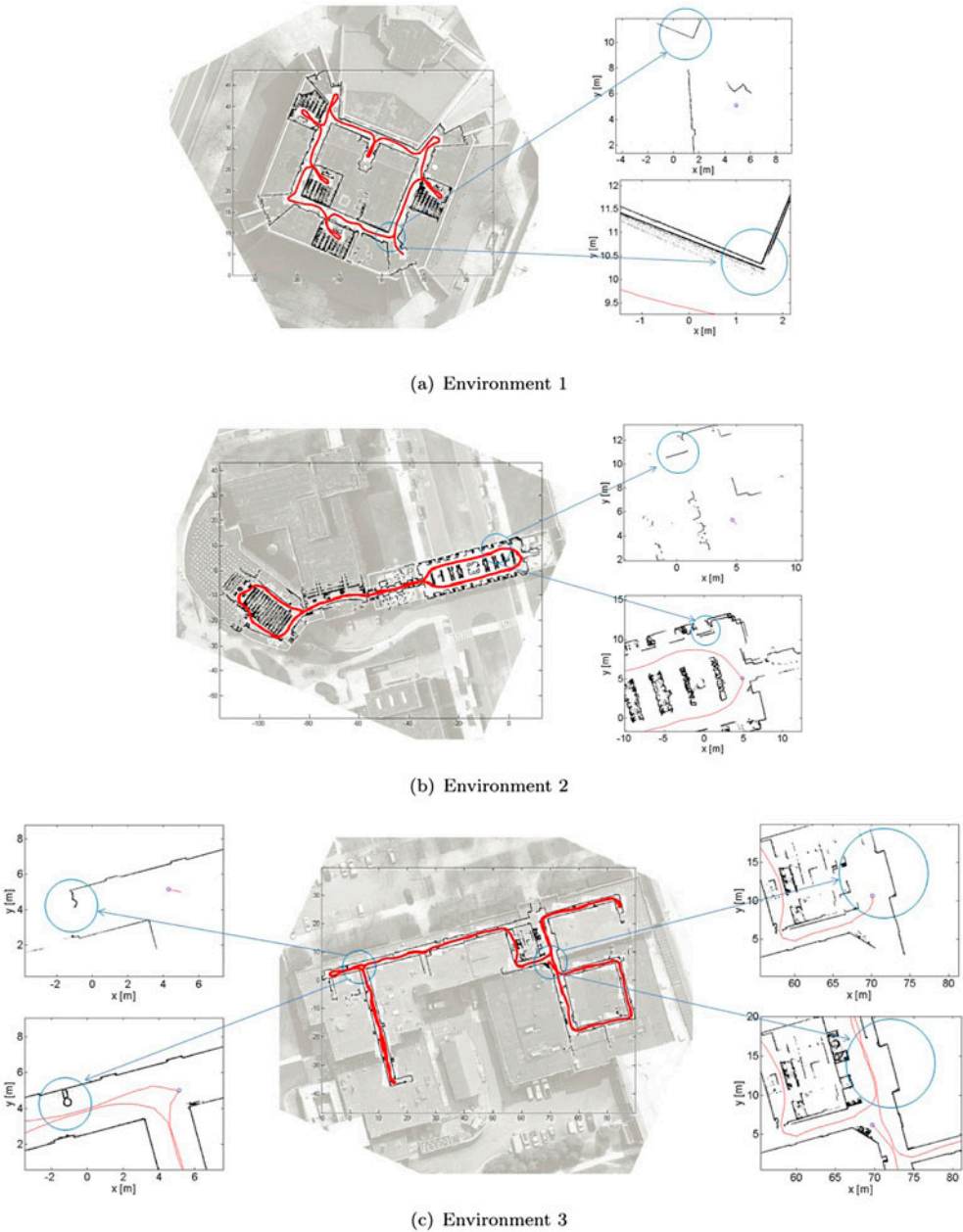(a) Environment 1   (b) Environment 2

(c) Environment 3

Figure 14. Trajectory of the robot in the three environments.

of the map since the localization and the map building are directly related to each other. In order to have better understanding of the accuracy, each map is overlapped on top of the satellite view of each environment where a red line is the trajectory of the robot. No other additional techniques such as active loop closure are used in the proposed algorithm. For all the environments the results demonstrate the high accuracy of the proposed scan-to-map matching technique in real SLAM scenarios. When the robot visits the same areas in the environments more than once, it observes the same features more than once. The accumulated pose error is then roughly calculated by comparing the positions of the same features observed at different time steps. As shown in the figure the intersections that the robot has visited more than once are zoomed in and the features in the intersections are highlighted. In each environment there exists small amount of a position error (Environment 1), both a posi-

**Table 4.** Position and orientation error produced by the proposed technique in Environment 1, 2 and 3.

		Proposed method	
Environment	Total	Error	Error rate
1	186 m	12.2 cm	0.000656
	3646°	0.64°	0.000176
2	258 m	25 cm	0.00097
	1594°	1.17°	0.00073
3 (Point a)	112 m		
	1204°	Loop successfully closed, therefore error not measurable	
3 (Point d)	182 m		
	1770°		

(a) Environment 1

(b) Environment 2

(c) Environment 3

Figure 15. Maps of test environments by the proposed technique overlapped on satellite view of the environments and critical intersections the robot visited multiple times.
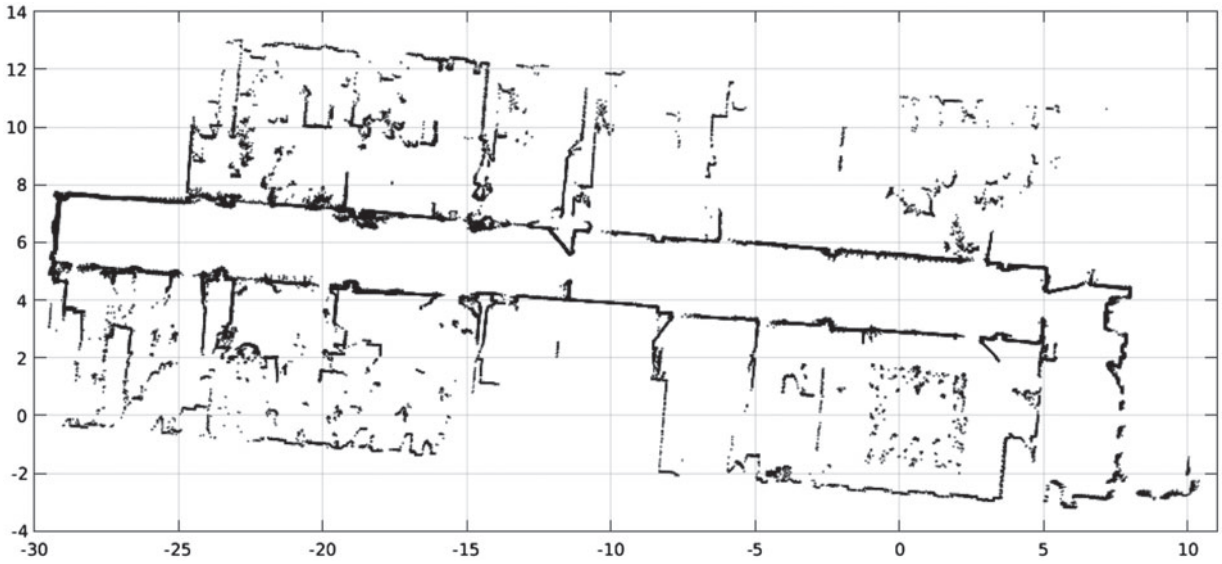
Figure 16. Map of the Freiburg University building 079 public domain dataset,[21] created using the proposed algorithm.
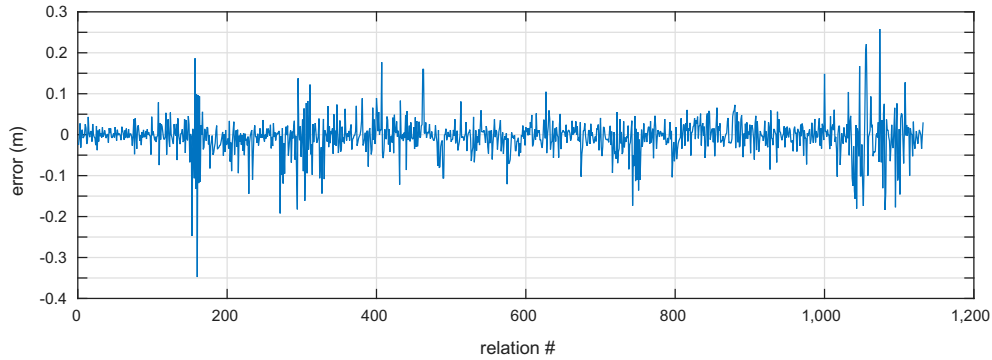
Figure 17. Shows the error in relative displacement in the nominated poses against the near-ground truth given in the dataset.

tion and an orientation error (Environment 2), or nearly no error (Environment 3) after the robot comes back to the same area. In Table 4 numerical results of the position and orientation errors from the proposed technique are listed. For all the environments the error rates defined as the position or the orientation error divided by the total distance or the accumulated orientation changes are less than 0.001, which indicates that the position and the orientation errors are less 10cm and 0.1° after traveling 100m with the accumulated orientation changes of 100°. The results show that the proposed technique is successful in estimating the robot pose and building the map in three different real environments. These errors were obtained manually by measuring distance between prominent features present in the map that were identified visually (Eg.: corners).

Furthermore we used the popular and publicly available Freiburg 079 dataset [21] with the proposed algorithm to create the map shown in Figure 16. This dataset contains near-ground truth relations data that had been

obtained by authors by manually matching scans at the given timestamps.[24] The plot in Figure 17 shows the error $\varepsilon(\delta_{ij})$, which is obtained using (42),[25] against the relation number from the dataset.

$$\varepsilon(\delta_{ij}) = \delta_{ij} \ominus \delta_{ij}^* \qquad (42)$$

Were δ_{ij} is the relative displacement between two poses i and j and δ_{ij}^* is the corresponding near-ground truth displacement from the dataset. These poses i and j are matched to be within $\pm 0.1s$ from the timestamps that are given in the dataset.

5. Conclusion

In this paper a grid-based scan-to-map matching technique for accurate mapping has been proposed. The proposed technique performs the local scan-to-scan matching and corrects the error from the matching by the use of a global scan-to-map matching technique. The

map to match is a grid map which may hold multiple NDs within each grid cell. Due to the scan-to-map matching with the multi-ND representation, the proposed technique exhibits little errors in the scan-to-map matching, and does not further accumulate the errors. The new scan is also represented by NDs enabling the novel ND-to-ND scan matching mechanism. The equal treatment of cells in the proposed ND-to-ND matching could further contribute to the accurate global scan-to-map matching.

The proposed scan-to-map matching technique was applied to three different experiments. The first experiment investigated the proposed scan-to-map matching technique in terms of the position and orientation error while the robot moved and sequentially obtained new scans. The experiment showed that the accumulated orientation error was negligible and that the position error stayed within 4 cm after traveling around 10 m. The second experiment, investigating the effectiveness of maintaining multiple NDs within a cell, showed that the scans were matched better to the map with higher similarity when the cell had multiple NDs than one ND. The experiment also demonstrated the robust effect of the use of multiple NDs. Finally, the proposed technique was used to create the maps in four real environments to demonstrate its applicability to real world problems. The resulting maps showed that the proposed technique, without any post processes such as the loop closure, generated position errors in the order of ten centimeters with very small orientation errors for the first three environments after traveling around 200 m with large orientation changes. Furthermore the results from the public domain Freiburg dataset when compared to the near-ground truth data given in the dataset shows to have a root mean square error of 0.0446 m, a mean absolute error of 0.0285 m and a maximum absolute error of 0.3477 m.

This paper is focused on the development of a new technique for scan-to-map matching, and much work is still left open to demonstrate its practical usefulness. Ongoing work primarily includes the investigation into computational efficiency as the proposed ND-to-ND matching technique is supposed to be significantly superior to the point-to-X matching techniques. The extension of the proposed technique to provide a full SLAM solution which would include propagation of uncertainty of the robot pose and loop closures is also being investigated. The potential of this method in being applied to 3D mapping is also currently investigated. Future works also include the applications of the proposed technique to different multi-task robotic problems in combination with other works of the authors.[26–28]

Disclosure statement

No potential conflict of interest was reported by the authors.

Notes on contributors

Kunjin Ryu received the MSc degree in electrical engineering from the University of Pennsylvania, Philadelphia, PA, USA, in 2008, and the PhD degree in mechanical engineering from Virginia Polytechnic Institute and State University, Blacksburg, VA, USA, in 2012. He is currently a senior research engineer at Samsung Electronics. His research interest includes mobile robot localization and mapping, search, tracking, navigation, and cooperative control for multi-robot systems.

Lakshitha Dantanarayana is a PhD student at the Centre for Autonomous Systems, University of Technology, Sydney (UTS), Australia. He obtained his BSc of Eng.(Hons.) degree in Electronics and Telecommunications Engineering from the University of Moratuwa, Sri Lanka in 2010. His research is focused on localization, mapping, SLAM, and Navigation for assistive robots. This work was conducted while he was a visiting researcher at the Department of Mechanical Engineering at Virginia Polytechnic Institute and State University (Virginia Tech), USA.

Tomonari Furukawa received PhD in Quantum Engineering and Systems Science from University of Tokyo in 1996. Since 2011, he has been the Professor of Mechanical Engineering at Virginia Polytechnic Institute and State University (Virginia Tech), USA. He also holds a joint professorship appointment at University of Technology, Sydney (UTS), Australia. His research work focuses on inverse analysis and optimization methods in experimental/computational mechanics and robotics where he currently has his particular interests in Bayesian estimation and control of autonomous systems, stochastic material characterization and structural health monitoring and multiphysics modeling and simulation. He also led teams for several international robotics competitions and has just led Virginia Techs Team VALOR at the DARPA Robotics Challenge.

Gamini Dissanayake is the James N. Kirby Professor of Mechanical and Mechatronic Engineering at University of Technology, Sydney (UTS), Australia. He has expertise in a broad range of topics in robotics including sensor fusion, localization, mapping, SLAM and human-robot interaction. He graduated in Mechanical/Production Engineering from the University of Peradeniya, Sri Lanka in 1977. He received his MSc in Machine Tool Technology and PhD in Mechanical Engineering (Robotics) from the University of Birmingham, England in 1981 and 1985 respectively.

ORCID

Lakshitha Dantanarayana 🔟 http://orcid.org/0000-0002-8887-329X
Tomonari Furukawa 🔟 http://orcid.org/0000-0003-2811-4221
Gamini Dissanayake 🔟 http://orcid.org/0000-0002-7992-0680

References

[1] Castellanos JA, Tardos JD. Mobile robot localization and map building: a multisensor fusion approach. Norwell, (MA): Kluwer Academic Publishers; 2000.

[2] Bailey T. Mobile robot localisation and mapping in extensive outdoor environments [PhD thesis]; 2002. Sydney: University of Sydney.

[3] Smith RC, Cheeseman P. On the representation and estimation of spatial uncertainty. Int. J. Robot. Res. 1986;5:56–68.

[4] Dissanayake G, Newman P, Clark S, et al. A solution to the simultaneous localization and map building (SLAM) problem. IEEE Trans. Robot. Autom. 2001;17:229–241.

[5] Lu F, Milios E. Robot pose estimation in unknown environments by matching 2D range scans. J. Intell. Robot. Syst. 1997;18:249–275.

[6] Diosi A, Kleeman L. Laser scan matching in polar coordinates with application to SLAM. In: 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS; Canada: Monton; 2005. p. 1439–1444.

[7] Nüchter A, Elseberg J, Schneider P, et al. Linearization of rotations for globally consistent n-scan matching. In: Proceedings – IEEE International Conference on Robotics and Automation; Anchorage, AK, USA. 2010. p. 1373–1379.

[8] Olson CF, Matthies LH. Maximum likelihood rover localization by matching range maps. In: IEEE International Conference on Robotics and Automation; Leuven, Belgium. 1998. p. 272–277.

[9] Burguera A, González Y, Oliver G. Probabilistic sonar scan matching for robust localization. In: Proceedings – IEEE International Conference on Robotics and Automation. Vol. 2007; Roma, Italy. 2007. p. 3154–3160.

[10] Gutmann JS, Konolige K. Incremental mapping of large cyclic environments. In: Proceedings 1999 IEEE International Symposium on Computational Intelligence in Robotics and Automation. CIRA'99 (Cat. No.99EX375); Monterey, CA, USA. 1999. p. 318–325.

[11] Besl PJ, McKay ND. A method for registration of 3-D shapes. IEEE Trans. Pattern Anal. Machine Intell. 1992;14:239–256.

[12] Weiss G, von Puttkamer E. A map based on laserscans without geometric interpretation. In: Proceedings of Intelligent Autonomous Systems 4(IAS-4); Karlsruhe. 1995. p. 403–407.

[13] Biber P, Straßer W. The normal distributions transform: a new approach to laser scan matching. In: Proceedings 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2003) (Cat. No.03CH37453); Vol. 3; Las Vegas: IEEE; 2003. p. 2743–2748.

[14] Lu FLF, Milios E. Robot pose estimation in unknown environments by matching 2D range scans. In: Proceedings CVPR '94, 1994 IEEE Computer Society Conference on Computer Vision and Pattern Recognition; Vol. 18; Seattle, WA. 1994. p. 935–938.

[15] Thrun S, Bugard W, Fox D. A real-time algorithm for mobile robot mapping with applications to multi-robot and 3D mapping. In: International Conference on Robotics and Automation; San Francisco, CA, USA. 2000. p. 321–328.

[16] Bosse M, Newman P, Leonard JJ, et al. Slam in large-scale cyclic environments using the atlas framework. Int. J. Robot. Res. 2004;23:1113–1139.

[17] Newcombe RA, Davison AJ, Izadi S, et al. KinectFusion: Real-time dense surface mapping and tracking. In: 2011 10th IEEE International Symposium on Mixed and Augmented Reality; October. IEEE; Basel. 2011. p. 127–136.

[18] Tomono M. Efficient global scan matching using saliency-based scan point resampling. In: 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems. Edmonton: IEEE; 2005. p. 1856–1861.

[19] Saarinen J, Andreasson H, Stoyanov T, et al. Normal distributions transform occupancy maps: application to large-scale online 3D mapping. In: International I, editor. Proceedings – IEEE International Conference on Robotics and Automation; Karlsruhe. 2013. p. 2233–2238.

[20] Arun KS, Huang TS, Blostein SD. Least-squares fitting of two 3-D point sets. IEEE Trans. Pattern Anal. Machine Intell. 1987;PAMI-9:698–700.

[21] Stachniss C. Freiburg indoor building 079 dataset. Freiburg.

[22] Ula C, Temelta H. 3D multi-layered normal distribution transform for fast and long range scan matching. J. Intell. Robot. Syst.: Theory Appl. 2013;71:85–108.

[23] Takeuchi E, Tsubouchi T. A 3-D scan matching using improved 3-D normal distributions transform for mobile robotic mapping. In: 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems; October. Beijing: IEEE; 2006. p. 3068–3073.

[24] Kümmerle R, Steder B, Dornhege C, et al. On measuring the accuracy of SLAM algorithms. Autonom. Robots. 2009;27:387–407.

[25] Burgard W, Stachniss C, Grisetti G, et al. A comparison of SLAM algorithms based on a graph of relations. In: 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems; October. St. Louis (MO): IEEE; 2009. p. 2089–2095.

[26] Furukawa T, Durrant-Whyte HF, Lavis B. The element-based method – theory and its application to Bayesian search and tracking. IEEE Int. Conf. Intell. Robots Syst. 2007;2007:2807—2812.

[27] Furukawa T, Mak LC, Ryu K, et al. Bayesian search, tracking, localization and mapping: a unified strategy for multi-task mission. In: INFORMS; November; Vol. 2011; Charlotte, USA. p. 13–16.

[28] Ryu K, Furukawa T. A LRF-based teleoperated navigation method. In: The International Conference on Intelligent Unmanned Systems; Chiba, Japan; 2011.