The International Journal of
Robotics Research
31(12) 1377–1393

DOI: 10.1177/0278364912460895
ijr.sagepub.com

Fast and accurate scan registration through minimization of the distance between compact 3D NDT representations

Todor Stoyanov, Martin Magnusson, Henrik Andreasson and Achim J Lilienthal

Abstract

Registration of range sensor measurements is an important task in mobile robotics and has received a lot of attention. Several iterative optimization schemes have been proposed in order to align three-dimensional (3D) point scans. With the more widespread use of high-frame-rate 3D sensors and increasingly more challenging application scenarios for mobile robots, there is a need for fast and accurate registration methods that current state-of-the-art algorithms cannot always meet. This work proposes a novel algorithm that achieves accurate point cloud registration an order of a magnitude faster than the current state of the art. The speedup is achieved through the use of a compact spatial representation: the Three-Dimensional Normal Distributions Transform (3D-NDT). In addition, a fast, global-descriptor based on the 3D-NDT is defined and used to achieve reliable initial poses for the iterative algorithm. Finally, a closed-form expression for the covariance of the proposed method is also derived. The proposed algorithms are evaluated on two standard point cloud data sets, resulting in stable performance on a par with or better than the state of the art. The implementation is available as an open-source package for the Robot Operating System (ROS).

Keywords

point set registration, mapping, normal distributions transform

1. Introduction

Point set registration, the task of finding the best fitting alignment between two sets of point samples, is an important problem in mobile robotics. A fast and reliable registration algorithm is key to building accurate 3D geometrical maps, using range data collected by a mobile platform. In some applications, online map building can be vital to the success of autonomous robotic missions in unknown environments. With the development of higher-bandwidth 3D range sensors, the design of faster and more accurate registration algorithms is becoming increasingly important.

Rigid-body registration algorithms can generally be partitioned in two sets, global and local, depending on the search method they employ. Local methods rely on the assumption that a sufficiently good initial guess for the relative position and orientation between the two point sets is available. Global methods, on the other hand, are usually based on local geometrical features that can be uniquely matched, without any assumptions on the initial state (see, for example, Rusu et al. (2009), Pathak et al. (2010), and Bülow and Birk (2011)).

One of the most influential algorithms for point set registration in three dimensions, the Iterative Closest Point (ICP) method (Besl and McKay, 1992), belongs to the local optimization methods class. ICP is simple to implement and often converges to the correct solution. Many modifications of ICP have been proposed (see, for example, Granger and Pennec, 2006; Nüchter et al., 2007; Segal et al., 2009) in order to improve the algorithm's robustness to outliers, its convergence speed and to reduce the bias towards maximizing scan overlap.

An alternative local optimization 3D point set registration method is based on the Three-Dimensional Normal Distributions Transform (3D-NDT), a special case of a Gaussian Mixture Model (GMM) with uniform weights and

Center of Applied Autonomous Sensor Systems (AASS), Örebro University, Sweden

Corresponding author:
Todor Stoyanov, Orebro University, AASS Research Center, Fakultetsgatan 1, Örebro 70281, Sweden.
Email: todor.stoyanov@oru.se

largely disjoint components. This article builds upon previous work in 3D-NDT registration by Magnusson (2009), which demonstrated that the 3D-NDT is a compact spatial representation that can be used for state-of-the-art 3D point set registration. In our previous work (Stoyanov et al., 2012) we proposed an algorithm that minimizes the L_2 distance between 3D-NDT models and demonstrated its utility in performing point set registration. This article presents in details the Distribution-to-Distribution (D2D) 3D-NDT registration algorithm and performs an extensive evaluation on standardized data sets. In addition, this article also details two previously undiscussed aspects of the 3D-NDT registration framework, namely the choice of a good initial starting point for the iterative optimization algorithm and the estimation of a covariance, representing the expected error in a solution obtained by the proposed algorithm.

The next section reviews the Three-Dimensional Normal Distributions Transform and prior work in 3D-NDT registration by Magnusson (2009). Section 3 then overviews the related work in registration of Gaussian Mixture Models and discusses in detail the 3D-NDT D2D registration algorithm. Section 4 reviews prior work on the 3D-NDT histogram: a global descriptor, initially proposed as a tool for loop closure detection in SLAM systems (Magnusson et al., 2009a). An algorithm for finding good initial pose estimates, based on the 3D-NDT histogram is then devised. Section 5 discusses the problem of estimating an error covariance for the solutions obtained by the proposed 3D-NDT D2D registration algorithm. Finally, Section 6 evaluates the performance of the proposed registration algorithm and its components, using standard real-world and simulated data sets.

2. Point set modelling and registration with the 3D-NDT

The Normal Distributions Transform (NDT) was originally developed in the context of 2D laser scan registration (Biber and Straßer, 2003). The central idea is to represent the observed range points as a set of Gaussian probability distributions. NDT has later been extended to three dimensions (Magnusson et al., 2007) and applied to the domains of 3D scan registration and loop detection (Magnusson, 2009), as well as change detection (Andreasson et al., 2007) and path planning (Stoyanov et al., 2010). One of the key advantages of 3D-NDT is the fact that it forms a piecewise smooth spatial representation, resulting in the existence of analytic derivatives. Consequently, standard optimization methods can be employed to produce state-of-the-art registration of 3D point clouds using the 3D-NDT representation (Magnusson et al., 2009b). After registration, a complete 3D-NDT map, which can be used for further tasks, can be created from the individual scans, using a global network-optimization scheme.

Assuming that a set of n point samples $\mathcal{P} = \{\boldsymbol{p}_i = (x_i, y_i, z_i)\}$ with $i = 1, \ldots, n$ has been drawn from a Gaussian distribution $\mathcal{N}(\boldsymbol{\mu}, \Sigma)$, the maximum likelihood estimates of the covariance and mean can be obtained from the observations:

$$\boldsymbol{\mu} = \frac{1}{n} \sum_{i=1}^{i=n} \boldsymbol{p}_i \tag{1}$$

$$M = \left[\boldsymbol{p}_1 - \boldsymbol{\mu} \ldots \boldsymbol{p}_n - \boldsymbol{\mu}\right] \tag{2}$$

$$\Sigma = \frac{1}{n-1} M M^{\mathrm{T}}. \tag{3}$$

The probability density function (pdf) estimated in this manner might or might not be a good representation of the sampled points, depending on the extent to which the assumption of Gaussianity on $\mathcal{P}$ holds. In a typical scenario of points sampled from a physical environment, a single normal distribution would not produce a good global estimate. However, at a sufficiently small scale, a normal distribution can be considered a good estimate of the local surface shape. Thus, the basic principle of the NDT is to represent space using a set of Gaussian probability distributions.

The formulation of a 3D-NDT model M_{NDT}, using a collection of Gaussian pdfs, is in fact analogous to that of a GMM. A GMM M_{GMM}, or alternatively a 3D-NDT model M_{NDT}, reconstructed from the point cloud $\mathcal{P}$ is defined as a set of Gaussian probability distributions:

$$M_{\mathrm{NDT}}(\mathcal{P}) = \{\mathcal{N}(\boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j)\}, \quad j = 1, \ldots, n_{\mathcal{P}}, \tag{4}$$

where $n_{\mathcal{P}}$ is the number of normal distributions in the model (typically $n_{\mathcal{P}} \ll |\mathcal{P}|$). The likelihood of observing a point $\boldsymbol{x}$, generated by the model, can be expressed as a sum of $n_{\mathcal{P}}$ weighted Gaussians:

$$p(\boldsymbol{x}|M_{\mathrm{NDT}}(\mathcal{P})) = \sum_{j=1}^{n_{\mathcal{P}}} w_j \mathcal{N}(\boldsymbol{x}|\boldsymbol{\mu}_j, \Sigma_j). \tag{5}$$

It is important to note that GMMs can be used to model arbitrary smooth density functions, thus making it an ideal candidate for an accurate spatial representation. This desirable property of GMMs comes at a price, however: it is not straightforward to estimate (or learn) a GMM, given a set of input point data. The classical approach taken uses an Expectation Maximization (EM) algorithm in order to find the combination of weights, mean vectors and covariances that produces the smallest residual errors in predicting input data. The EM algorithm however needs as an input the number of components $n_{\mathcal{P}}$, which are not straightforward to guess. In this sense, the 3D-NDT provides a different method to estimate a GMM, based on a grid discretization of space.

Let $\mathcal{F}$ and $\mathcal{M}$ be two point sets, taken from partially overlapping portions of the environment. In addition, we define $\mathcal{F}$ to be a fixed reference, while $\mathcal{M}$ is a moving point set.

The registration task estimates the parameters Θ of a transformation function T, such that $T(\mathcal{M}, \Theta)$ is consistently aligned with respect to $\mathcal{F}$. Previous work by Magnusson (2009) has proposed to first construct the 3D-NDT of the fixed model point set $M_{\text{NDT}}(\mathcal{F})$. The likelihood that a point $\boldsymbol{x}$ is generated from $M_{\text{NDT}}(\mathcal{F})$ is then defined by (5) and rewritten as

$$p(\boldsymbol{x}|M_{\text{NDT}}(\mathcal{F})) = \sum_{i=1}^{n_{\mathcal{F}}} w_i \mathcal{N}(\boldsymbol{x}|\boldsymbol{\mu}_i, \Sigma_i), \tag{6}$$

where $n_{\mathcal{F}}$ is the number of Gaussian components of the 3D-NDT model of the point scan $\mathcal{F}$. The likelihood that the set of points $\mathcal{M}$ is generated by $M_{\text{NDT}}(\mathcal{F})$ can be written as

$$p(\mathcal{M}|M_{\text{NDT}}(\mathcal{F})) = \prod_{j=1}^{|\mathcal{M}|} p(m_j|M_{\text{NDT}}(\mathcal{F})), \tag{7}$$

where m_j iterates over all points in $\mathcal{M}$. The 3D-NDT Point-To-Distribution (P2D) registration algorithm (Magnusson, 2009) minimizes an approximation of the negative log-likelihood function of $p(T(\mathcal{M}, \Theta)|M_{\text{NDT}}(\mathcal{F}))$, over the space of transformation parameters Θ:

$$\tilde{p}(\mathcal{M}|M_{\text{NDT}}(\mathcal{F})) = -\sum_{j=1}^{|\mathcal{M}|} \log p(m_j|M_{\text{NDT}}(\mathcal{F})). \tag{8}$$

For consistency, the weights of all Gaussian components in Equation (6) should be set uniformly, i.e. $w_i = \frac{1}{n_{\mathcal{F}}}$. As noted previously, for each query point $\boldsymbol{x}$ in Equation (6), a reasonable approximation is to only consider the Gaussian component $\mathcal{N}(\boldsymbol{\mu}_x, \Sigma_x)$, closest to $\boldsymbol{x}$. Thus, $w_i = w_x = 1$ when i corresponds to the closest Gaussian component and $w_i = 0$ otherwise. Magnusson (2009) also proposed an interpolated version of the objective function, which takes into account several close Gaussian components (not considered in this work). Finally, after approximation and reorganization of terms, the registration problem is posed as minimizing the objective function:

$$f_{\text{p2d}}(\mathcal{M}, \mathcal{F}, \Theta) = -d_1 \sum_{j=1}^{|\mathcal{M}|} \exp \frac{-d_2}{2} (T(m_j, \Theta) - \boldsymbol{\mu}_m)^{\text{T}} \Sigma_m^{-1} (T(m_j, \Theta) - \boldsymbol{\mu}_m), \tag{9}$$

where d_1 and d_2 are positive regularizing factors, j iterates over all points m_j in the moving scan $\mathcal{M}$ and $(\boldsymbol{\mu}_m, \Sigma_m)$ are the parameters of the normal distribution in $M_{\text{NDT}}(\mathcal{F})$, closest to the transformed current point $T(m_j, \Theta)$. The objective f_{p2d} is doubly differentiable with analytic expressions for the gradient and Hessian.

This approach of registering points to 3D Gaussian distributions produces good registration results, using standard non-linear optimization methods. The P2D registration algorithm described above operates on a number of (possibly subsampled) data points from the moving scan $\mathcal{M}$. It is, however, possible to approach the registration problem differently: by directly comparing the 3D-NDT model of $\mathcal{M}$ with the model of the fixed scan $M_{\text{NDT}}(\mathcal{F})$, resulting in a D2D registration approach, discussed in the next section.

As shown previously, the spatial indexing used has an effect on the representation accuracy of the 3D-NDT (Stoyanov et al., 2011), as well as on the speed and accuracy of the registration algorithms that utilize it (Magnusson, 2009). The registration algorithm proposed by Magnusson et al. (2007) uses a regular grid discretization, but performs several consecutive registrations, using different discretization levels, a strategy which was also used in the algorithm presented in this article.

3. D2D 3D-NDT registration

3.1. Background

This section presents a different perspective on the point set registration problem. Instead of maximizing the likelihood of a discrete set of points $\mathcal{M}$ as in the previous section, the registration problem is interpreted as minimizing the distance between two 3D-NDT models $M_{\text{NDT}}(\mathcal{F})$ and $M_{\text{NDT}}(\mathcal{M})$. As mentioned earlier, a 3D-NDT model can be viewed as a pdf, signifying the likelihood of observing a point in space, belonging to an object surface (as in Equation (6)). This formulation is equivalent to a GMM with uniform weights and largely disjoint components. Several possible distance metrics between probability distributions can be used to formulate the D2D objective function. This section reviews some of the related work in GMM registration and motivates the choice of a distance metric.

Several statistical metrics are commonly used to compute the distance between probability distributions. The Kullback–Leibler (KL) divergence $KL(R||S)$ measures the expected additional information needed to explain the information carried by the random variable S, using the random variable R. The symmetric KL divergence

$$SKL(R||S) = KL(R||S) + KL(S||R) \tag{10}$$

is commonly used to compare two probability distributions in an unbiased way. Although approximations of the KL divergence for GMMs exist (Hershey and Olsen, 2007; Goldberger et al., 2008), most of them are based on sampling and cannot be used for analytical purposes. Hershey and Olsen (2007) propose three different closed-form variational bounds to the KL divergence between GMMs. An approach based on the proposed approximations was implemented, but resulted in unsatisfactory performance, due to the too general nature of the proposed bounds. Other distance metrics include the Bhattacharyya and closely related Hellinger distance, both of which do not have analytic solutions for GMMs, and were therefore not be considered.

A different approach is to use mutual information $\mathcal{I}(M_{\text{NDT}}(\mathcal{M}), M_{\text{NDT}}(\mathcal{F}))$ as a distance metric. Wells et al.

(1996) propose an algorithm that maximizes the mutual information between kernel density estimates of MRI and CT images. A well-known way to express mutual information of two random variables is through a re-formulation using the Shannon entropy H_s:

$$\begin{aligned}\mathcal{I}(T(M_{\text{NDT}}(\mathcal{M}),\Theta),M_{\text{NDT}}(\mathcal{F})) &= H_s(M_{\text{NDT}}(\mathcal{F}))\\ &+H_s(T(M_{\text{NDT}}(\mathcal{M}),\Theta)\\ &-H_s(T(M_{\text{NDT}}(\mathcal{M}),\Theta),M_{\text{NDT}}(\mathcal{F}))\end{aligned} \tag{11}$$

where $T(M_{\text{NDT}}(\mathcal{M}),\Theta)$ is a parametrized transformation of the 3D-NDT model of $\mathcal{M}$, and the Shannon entropy of a random variable generated from a pdf f is defined as

$$H_s(x) = E\{-\log f(x)\} = -\int f(x)\log f(x)\,dx. \tag{12}$$

Notably, when minimizing $\mathcal{I}$ as a function of Θ, the term $H_s(M_{\text{NDT}}(\mathcal{F}))$ can be disregarded, as it is constant in Θ. The entropy of the 3D-NDT model of the moving scan $M_{\text{NDT}}(\mathcal{M})$ is invariant under rigid transformations and thus $H_s(T(M_{\text{NDT}}(\mathcal{M}),\Theta)$ also does not have an influence on the objective function. Accordingly, the registration problem can be posed as minimizing the Shannon entropy of the joint pdf of the two models $H_s(T(M_{\text{NDT}}(\mathcal{M}),\Theta),M_{\text{NDT}}(\mathcal{F}))$.

While the entropy of a single Gaussian component has a closed-form solution, no analytic expression can be obtained for a GMM. The most popular approach, also applied by Wells et al. (1996), is to compute the entropy explicitly using Monte Carlo sampling. This, however, is slow and forbids the computation of analytic derivatives. In a recent work, Huber et al. (2008) derive several closed-form approximations of GMM entropy, based on Taylor series expansions. An approach based on these results is feasible, but would encounter the serious difficulty of estimating the joint density of the two 3D-NDTs. Several approaches can be used to estimate the joint density. Most straightforward would be to estimate a new 3D-NDT model from the joint set of model and transformed scan points. This approach would be similar to that used by Wells et al. (1996) who draw samples from the joint distribution and construct a new kernel density estimate. Unfortunately, this method would prohibit further analytical derivations and also introduce a significant overhead of estimating a 3D-NDT model on every iteration of the registration algorithm. Another option is to try to estimate the joint 3D-NDT analytically. Since the distributions are not generally independent random variables, the 3D-NDT components cannot be correctly fused together using the standard Kalman filter update step. As Julier (2007) noted, a Chernoff information-based method is necessary when fusing estimates of random variables with unknown correlation. Finding the optimal fusion parameters is, however, a time-consuming task, as it involves solving an optimization problem for each pair of components to be merged.

3.2. Registration based on L_2 distance

A different approach to GMM registration is to minimize the L_2 norm between the pdfs, over all possible inputs. A point set registration algorithm that utilizes this strategy was proposed and formalized by Jian and Vemuri (2010), building on previous related work on kernel correlation by Tsin and Kanade (2004). The work by Jian and Vemuri (2010) proposes to compute and register a fixed-width spherical Gaussian kernel density estimate of each point cloud. The rest of this section extends their approach to the more general, variable kernel size 3D-NDT representation.

Using the definitions from Section 2, the L_2 distance between two 3D-NDTs (or GMMs) is defined as

$$\begin{aligned}D_{L_2}(\mathcal{M},\mathcal{F},\Theta) = \int (&p(\boldsymbol{x}|M_{\text{NDT}}(\mathcal{F}))\\ &-p(\boldsymbol{x}|T(M_{\text{NDT}}(\mathcal{M}),\Theta)))^2\,d\boldsymbol{x}\end{aligned} \tag{13}$$

or, equivalently,

$$\begin{aligned}D_{L_2}(\mathcal{M},\mathcal{F},\Theta) = &\int p(\boldsymbol{x}|M_{\text{NDT}}(\mathcal{F}))^2\,d\boldsymbol{x}\\ &+\int p(\boldsymbol{x}|T(M_{\text{NDT}}(\mathcal{M}),\Theta))^2\,d\boldsymbol{x}\\ &-2\int p(\boldsymbol{x}|M_{\text{NDT}}(\mathcal{F}))\,p(\boldsymbol{x}|T(M_{\text{NDT}}(\mathcal{M}),\Theta))\,d\boldsymbol{x},\end{aligned} \tag{14}$$

where $T(X,\Theta)$ is a parametrized transformation function, as in the previous section. It is useful to interpret this expression using Renyi's Quadratic Entropy (RQE) H_{rqe}, defined as

$$H_{\text{rqe}}(P(\boldsymbol{x})) = -\int P^2(\boldsymbol{x})\,d\boldsymbol{x}. \tag{15}$$

Noting that any transformation of an arbitrary 3D-NDT model $T(\mathcal{M}_{\text{NDT}},\Theta)$ is still a 3D-NDT model, we can simplify Equation (14). Evidently, the first two terms can be interpreted as the sum of the RQEs of the 3D-NDT models of the fixed and the moving scans. The last term relates to the mixing of the models of the two point scans. The two entropy terms are invariant over rigid transformations and can be disregarded for the objective function. As Jian and Vemuri (2010) note, the third term can be simplified by expanding the Gaussian components $\mathcal{N}$ and applying the following identity:

$$\int \mathcal{N}(\boldsymbol{x}|\boldsymbol{\mu}_i,\Sigma_i)\mathcal{N}(\boldsymbol{x}|\boldsymbol{\mu}_j,\Sigma_j)\,d\boldsymbol{x} = \mathcal{N}(0|\boldsymbol{\mu}_i-\boldsymbol{\mu}_j,\Sigma_i+\Sigma_j). \tag{16}$$

Thus, a final expression for the L_2 distance to be minimized is obtained as

$$\begin{aligned}D_{L_2}(\mathcal{M},\mathcal{F},\Theta) \sim \sum_{i=1}^{n_{\mathcal{M}}}\sum_{j=i}^{n_{\mathcal{F}}}&\mathcal{N}(0|T(\boldsymbol{\mu}_i,\Theta)\\ &-\boldsymbol{\mu}_j, T(\Sigma_i,\Theta)+\Sigma_j),\end{aligned} \tag{17}$$

with $n_{\mathcal{M}}$ and $n_{\mathcal{F}}$ being the number of Gaussian components in $M_{\text{NDT}}(\mathcal{M})$ and $M_{\text{NDT}}(\mathcal{F})$, respectively.

The transformation function $T(\cdot, \Theta)$ can be parametrized using different rotation representations. The parameters of the transformation are accumulated in a vector $\boldsymbol{p} = (t_x, t_y, t_z, r_x, r_y, r_z)$, with translation vector $\boldsymbol{t} = (t_x, t_y, t_z)$, and x–y–z Euler angles r_x, r_y, r_z. Further discussion on the chosen rotation representation are deferred to Section 3.3. The objective function of the D2D registration algorithm is defined as

$$f_{\text{d2d}}(\boldsymbol{p}) = \sum_{i=1}^{n_{\mathcal{M}}} \sum_{j=i}^{n_{\mathcal{F}}} -d_1 \exp\left(-\frac{d_2}{2} \boldsymbol{\mu_{ij}}^{\text{T}} (R^{\text{T}} \Sigma_i R + \Sigma_j)^{-1} \boldsymbol{\mu_{ij}}\right), \tag{18}$$

where $\boldsymbol{\mu_{ij}} = \boldsymbol{\mu_{ij}}(\boldsymbol{p})$ is the transformed mean vector distance:

$$\boldsymbol{\mu_{ij}}(\boldsymbol{p}) = R\boldsymbol{\mu}_i + \boldsymbol{t} - \boldsymbol{\mu}_j. \tag{19}$$

The gradient vector of the D2D NDT registration function can then be derived, componentwise, as

$$\frac{\partial}{\partial p_a} f_{\text{d2d}}(\boldsymbol{p}) = \frac{d_1 d_2}{2} \left(\boldsymbol{\mu_{ij}}^{\text{T}} B \boldsymbol{j_a} - \boldsymbol{\mu_{ij}}^{\text{T}} B Z_a B \boldsymbol{\mu_{ij}}\right) \exp\left(-\frac{d_2 \boldsymbol{\mu_{ij}}^{\text{T}} B \boldsymbol{\mu_{ij}}}{2}\right), \tag{20}$$

with p_a being a component of the pose vector $\boldsymbol{p}$ and the following component expressions

$$B = (R^{\text{T}} \Sigma_i R + \Sigma_j)^{-1} \tag{21}$$

$$\boldsymbol{j_a} = \frac{\partial}{\partial p_a}\left(R\boldsymbol{\mu}_i + \boldsymbol{t} - \boldsymbol{\mu}_j\right) \tag{22}$$

$$Z_a = \frac{\partial}{\partial p_a}\left(R^{\text{T}} \Sigma_i R\right). \tag{23}$$

Similarly, the Hessian matrix is

$$\frac{\partial^2}{\partial p_a \partial p_b} f_{\text{d2d}}(\boldsymbol{p}) = d_1 d_2 \left(\boldsymbol{j_a}^{\text{T}} B \boldsymbol{j_a} - 2\boldsymbol{\mu_{ij}}^{\text{T}} B Z_a \boldsymbol{j_a} + \boldsymbol{\mu_{ij}}^{\text{T}} B H_{ab} - \boldsymbol{\mu_{ij}}^{\text{T}} B Z_a B Z_b B \boldsymbol{\mu_{ij}} - \frac{1}{2} \boldsymbol{\mu_{ij}} B Z_{ab} B \boldsymbol{\mu_{ij}} - \frac{d_2}{4} \boldsymbol{q}^{\text{T}} \boldsymbol{q}\right) \times \exp\left(-\frac{d_2 \boldsymbol{\mu_{ij}}^{\text{T}} B \boldsymbol{\mu_{ij}}}{2}\right), \tag{24}$$

where

$$H_{ab} = \frac{\partial^2}{\partial p_a \partial p_b}\left(R\boldsymbol{\mu}_i + \boldsymbol{t} - \boldsymbol{\mu}_j\right) \tag{25}$$

$$Z_{ab} = \frac{\partial^2}{\partial p_a \partial p_b}(R^{\text{T}} \Sigma_i R) \tag{26}$$

$$\boldsymbol{q} = \boldsymbol{\mu_{ij}}^{\text{T}} B \boldsymbol{j_a} - \boldsymbol{\mu_{ij}}^{\text{T}} B Z_a B \boldsymbol{\mu_{ij}}. \tag{27}$$

The first- and second-order derivatives involving components of the transformation function $T(\boldsymbol{p})$ are easily calculated for each component of the pose vector $\boldsymbol{p}$ and used to obtain the expressions $\boldsymbol{j_a}, Z_a, H_{ab}, Z_{ab}$. The precise derivative expressions in terms of $(t_x, t_y, t_z, r_x, r_y, r_z)$ can readily be calculated and are not reported in this work. The so-defined objective and derivatives are used as inputs to a Newton's method optimization routine with a More–Thuente line search for step control. As mentioned earlier, the current version of the registration algorithm only considers the pairwise closest components from both 3D-NDTs. As the component distributions of each 3D-NDT are largely disjoint, this presents for a reasonable approximation of the objective function and for fast evaluation. Thus, the objective (18) and its derivatives are evaluated only at each pair of closest components from the models $M_{\text{NDT}}(\mathcal{F})$ and $T(M_{\text{NDT}}(\mathcal{M}), \boldsymbol{p}))$.

3.3. Rotation representation

While deriving the objective function (18), a choice to represent the pose of the scan was made. Namely, the pose vector considered is represented as $\boldsymbol{p} = (t_x, t_y, t_z, r_x, r_y, r_z) \in \mathbb{R}^6$, with x–y–z Euler angles r_x, r_y, r_z. The pose $\boldsymbol{p}$, however, consists of a translation and rotation component and should rather be viewed as $\boldsymbol{p} \in SE(3) = \mathbb{R}^3 \times SO(3)$, where $SO(3)$ is the special orthogonal group, representing orientations in three dimensional space. As the space of possible poses is not equivalent to the Euclidean space $\mathbb{R}^6$, it is important to devise the optimization scheme in accordance with this fact.

Euler angles are one of the possible parametrizations of orientations in $SO(3)$, using three consecutive rotations around the unit axes. Unfortunately, if they are used for representing orientation globally over all of $SO(3)$, singularities in the parametrization may be reached. This result is also a direct corollary from the well-known hairy ball theorem in topology, that prohibits the existence of smooth charts on even-dimensional unit spheres. Thus, a common approach for representing orientation is to embed the problem in a higher-dimensional space using unit quaternions and operating on the unit sphere S^3. This approach, however, has the weakness that when fusing locally obtained, incremental changes in orientation, the result might not be a unit quaternion and would require re-normalization.

Recent work by Hertzberg et al. (2011) proposes an approach to elegantly handling optimization problems defined on a manifold M (such as $SE(3)$). Two operations, $\boxplus : M \times \mathbb{R}^n \to M$ and $\boxminus : M \times M \to \mathbb{R}^n$, are defined and used as a map between variables on a manifold and algorithms operating locally in a Euclidean space. These operations can be thought of as performing a linearization around a state $\boldsymbol{q}_i \in M$ and performing operations in a locally Euclidean space. Given that the local operations deal with states that are sufficiently close, this scheme would guarantee the consistency of the final results.

Applying the approach of Hertzberg et al. (2011) directly to the registration problem considered, however, would break some of the nice properties of the algorithm considered, namely the existence of analytic derivatives. Note

that (18) (and (9) as well) operate on states that are embedded in a manifold. Thus, the objective function for the 3D-NDT D2D algorithm operates on poses that are not defined in Euclidean space, but rather belong to $SE(3)$. Following Hertzberg et al. (2011), special care would have to be taken when computing the derivatives:

$$\frac{\partial}{\partial \boldsymbol{p}} f_{\text{d2d}} = \lim_{\epsilon \to 0} \frac{f_{\text{d2d}}(\boldsymbol{p}) - f_{\text{d2d}}(\boldsymbol{p} \boxplus \epsilon e)}{\epsilon}, \tag{28}$$

where $e \in \mathbb{R}^6$ is the identity element. As the $\boxplus$ operator involves two projections and depends on the current state, analytic expressions for the derivatives will be infeasible to compute. Thus, a simpler approach to ensuring the stability of the objective function was taken in devising the registration algorithm presented here.

At each step of the optimization algorithm, a pose increment $\Delta_{\boldsymbol{p}}$ is computed, and subsequently used to increment the current pose $\boldsymbol{p}_i \in \mathbb{R}^6$. By controlling for the maximum step size in the line search method, a constraint can be placed on $\Delta_{\boldsymbol{p}}$, such that the increment is sufficiently small and can be reliably computed in a linearized rotation space. Next, instead of computing $\boldsymbol{p}_{i+1} = \boldsymbol{p}_i \boxplus \Delta_{\boldsymbol{p}}$, the whole 3D-NDT model $M_{\text{NDT}}(\mathcal{M}_i)$ is transformed using the current pose increment: $M_{\text{NDT}}(\mathcal{M}_{i+1}) = T(M_{\text{NDT}}(\mathcal{M}), \Delta_{\boldsymbol{p}})$. The next step of the optimization process then operates on the transformed 3D-NDT model, but uses the same pose $\boldsymbol{p}_i$ for computing the derivatives of the objective. The incremental updates $\Delta_{\boldsymbol{p}}$ are fused together into a homogeneous transformation matrix H, that can correctly represent $SE(3)$. When the optimization algorithm converges to a (locally) optimal $f^{\star}_{\text{d2d}}$, the resulting homogeneous transformation matrix H constitutes the final output of the registration algorithm.

This strategy is in fact very similar to that used in the ICP algorithm, which transforms the moving scan $\mathcal{M}$ at every iteration. The key difference is that the ICP algorithm uses a closed-form solution of the known point correspondence registration problem (an often used method is that proposed by Horn (1987)) on every iteration. Thus, ICP does not calculate derivatives on $SE(3)$, but instead relies on point correspondences, computed using the current pose. As the algorithm iterates, correspondence selection and closed form pose optimization are performed in an alternating manner.

The proposed modifications to the optimization algorithm ensure that the topology of the state space of f_{d2d} is taken into account while computing the final solution. Incidentally, this also greatly simplifies the computation of the Jacobian and Hessian matrices. As the state $\boldsymbol{p}_0$ around which the derivatives are computed never changes, $\boldsymbol{p}_0$ can be set to zero without loss of generality, and thus $\boldsymbol{t} = 0$ and $R = I$. This results in simpler expressions for the derivatives (20) and (24). The results for evaluating these expressions at $\boldsymbol{p} = 0$ can be easily calculated by evaluating the expressions, obtained in Section 3.2. This simplification comes at the added cost of transforming all components of $\mathcal{M}$ by $\Delta_{\boldsymbol{p}}$. This operation can, however, be done fast and efficiently, by transforming the mean and covariance of each Gaussian component at the cost of two matrix–matrix and one matrix–vector multiplications, as well as several additions, per component.

3.4. Relation to ICP methods

It may be worth noting the relationship of the algorithm proposed in the previous sections to prior work on ICP registration. Two notable similar works were proposed by Segal et al. (2009) and Jian and Vemuri (2010). Both of these works present generalizations of the ICP algorithm through the use of a modified distance metric function. In this respect, the objective functions presented in these works are similar to the objective proposed in Equation (18). Jian and Vemuri propose the use of a covariance-norm-based distance function. The key difference to our approach lies in the construction of the covariance matrix Σ, which Jian and Vemuri assume to be spherical and parametrized only by a variance σ along the main diagonal. Moreover, σ is a stationary parameter: it does not change for any location in $\mathbb{R}^3$. In the algorithm presented in this work, the covariance Σ is non-stationary and is estimated, based on the local structure of the registered point clouds.

The method proposed by Segal et al. (2009), Generalized ICP, formulates an objective function based on the error between points that are assumed to be drawn from independent Gaussian distributions. Thus, the objective function is similar to Equation (18). There are, however, several important differences between 3D-NDT D2D and Generalized ICP. First of all, the way the covariance matrix is estimated differs: the application of Generalized ICP discussed by Segal et al. (2009) assumes locally planar patches around each point, then calculates the normal direction to the local surface and uses it to bias the orientation of the covariance matrix. Second, the way that the optimal transformation parameters $T(\cdot, \Theta)$ are estimated is different, namely Generalized ICP relies on a conjugate gradient descent with numerical evaluation of derivatives, while 3D-NDT D2D uses a Newton's method with analytical derivative expressions. Last but not least, 3D-NDT D2D operates only on the estimated Gaussian components of the 3D-NDT models $M_{\text{NDT}}(\mathcal{F})$ and $M_{\text{NDT}}(\mathcal{M})$, as opposed to the full sets of point samples $\mathcal{F}$ and $\mathcal{M}$, thus allowing for a significantly reduced number of corresponding components considered in the evaluation of the objective function. The reduction in components, due to the compactness of the 3D-NDT representation, leads to lower runtimes, discussed in detail in the results (Section 6.2).

One way to obtain a rough intuitive visualization of the difference between the considered registration algorithms is through the aid of a simple 1D example (Figure 1). Two sets of 1D points, a fixed point cloud in red and a moving scan in blue, are to be registered by minimizing an objective function. The classical ICP method minimizes the sum

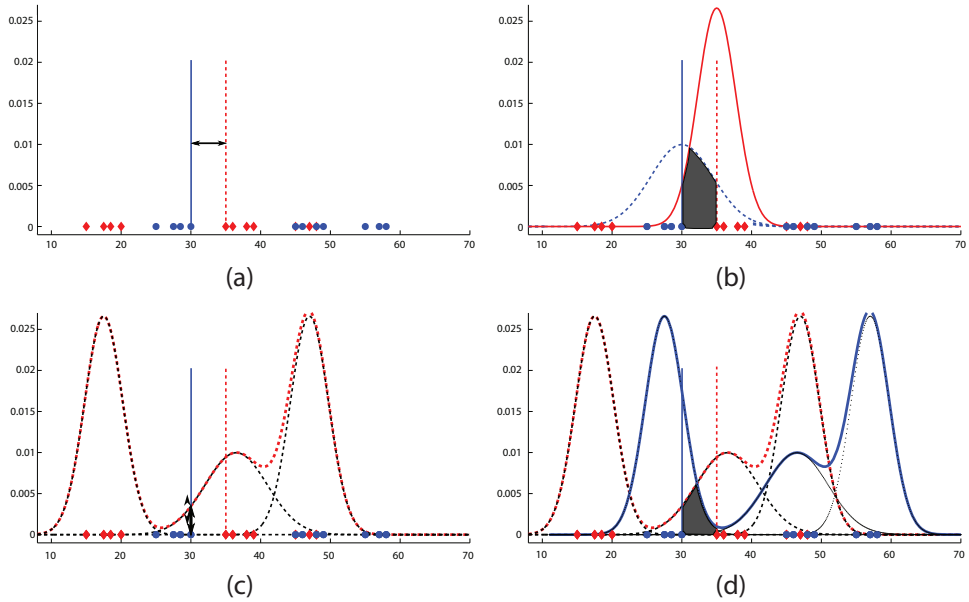

Fig. 1. An example of the objective functions minimized by different registration algorithms: Vanilla ICP in (a), Generalized ICP in (b), NDT-P2D in (c) and NDT-D2D in (d). The contribution of a pair of points to the objective function is proportional to the length of the thick black arrow in (a) and (c) and to the shaded area in (b) and (d).

of Euclidean distances in between associated closest neighbour points (visualized as a black arrow in Figure 1(a)). The generalized ICP algorithm places an associated covariance matrix around each point and proposes a modified distance metric. In the considered example, a Gaussian function is centred on each of the associated points and the overlapping area is added to the objective. The P2D 3D-NDT algorithm (Figure 1(c)) first computes a NDT model of the red point set (shown as a three-component mixture in red). The likelihood of the blue points is then maximized, given the red mixture model. The thick black arrow in this case represents the contribution of the same two points considered in Figure 1(a)–(b) to the objective function. It is important to note that unlike the two ICP algorithms, the 3D-NDT P2D does not explicitly compute this value. The model point (in red) is not used directly, but rather integrated into the 3D-NDT model (red curve). Finally, the 3D-NDT D2D (Figure 1(d)) algorithm computes both mixture models and maximizes their overlap. The contribution of the same two points, though also not explicitly computed by the algorithm, can be thought of as proportionate to the shaded area. The value of the overall objective function (i.e. to be minimized) is inversely proportional to the overlap between the red and blue curves.

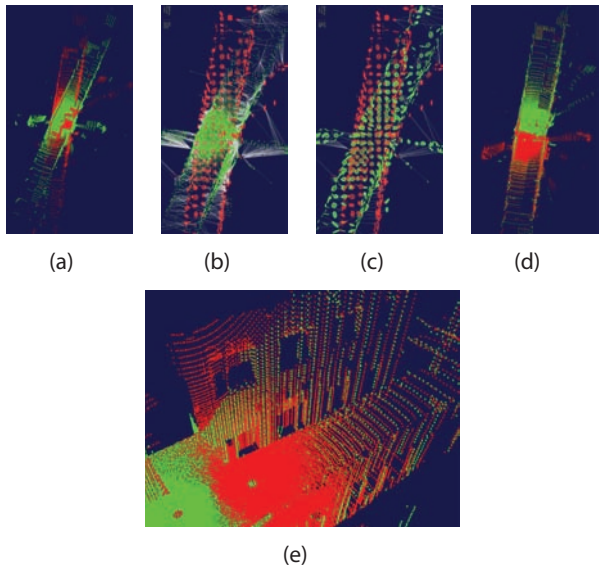

Fig. 2. An illustration of the performance of the registration algorithms. (a) Initial poses of the point sets to be registered. (b) Point-to-Distribution registration first iteration with correspondences drawn as lines. (c) Distribution-to-Distribution (D2D) first iteration, correspondences drawn as lines. (d) Final pose from 3D-NDT D2D registration. (e) A detail from the registered scan, different point clouds shown in red and green (colour refers to the online version of this article).

4. Estimating an initial guess

As mentioned in the introduction, the goal of the registration task is to estimate the relative pose between two physical locations, based on two sets of point samples. When considering iterative methods to solving the registration problem, several conditions have to be met in order to achieve a good final result. First of all, the objective function should be formulated in such a way that its minimum coincides with the exact pose offset between the two scans. As this condition practically cannot be guaranteed,

small errors in registration are always made, resulting in an associated covariance of the final estimate (discussed in Section 5). In addition, iterative methods rely on the existence of an initial guess $\boldsymbol{p}_0$, close enough to the global minimum of the objective function. There are three distinct ways of obtaining an initial estimate: from an external pose estimation system (for example, odometric or inertial sensors), through a pre-determined value or through the use of a global alignment strategy. The first method traditionally produces good results, but might not always be available (for example, on airborne or underwater vehicles). The second option, using a pre-determined value, usually refers to starting from the zero pose and relying on the physical proximity of the scanned locations. Both of these alternatives can lead to failure, in odometry deprived environments or in cases of large offsets between individual point scans. Thus, it is important to investigate the third option: global alignment methods for finding good initial offsets and bringing the registration algorithm inside the convergence zone of its optimal solution.

Several solutions have been proposed to achieve good initial alignment of point sets. As poses close to the optimal alignment $\boldsymbol{p}^*$ tend to be good initial starting points, this problem is in fact equivalent in formulation to the registration task. The key difference is that the proposed solution cannot be based on iterative optimization, but should instead directly provide a global solution. Two dominant classes of approaches to non-iterative 3D registration can be identified in the literature, based on local or global feature descriptors. Johnson and Hebert (1999) define spin images, local accumulators around a defining axis, and use them to perform registration and object recognition. Gelfand et al. (2005) define the volume integral descriptor and use it to register arbitrary 3D shapes. In a more recent series of works, Rusu et al. (2008) define the Point Feature Histogram (PFH) and its accelerated version the Fast PFH (FPFH) (Rusu et al., 2009) and use them on a series of registration problems. Steder et al. (2011) propose a different local feature, the Normally Aligned Radial Feature (NARF), and use it for place similarity identification. These local feature approaches can result in good initial alignments, but are usually prohibitively slow due to the large number of feature extraction/matching operations performed (except for the NARF features all other systems report runtimes for a single comparison well above 1 second). Approaches based on global appearance on the other hand are much faster: Makadia et al. (2006) and Pathak et al. (2010) both report fast and reliable registration results. Makadia et al. (2006) make use of an extended Gaussian image to represent point set data and perform efficient correlation. Pathak et al. (2010) extract planes in point sets and propose a noise-aware registration method, that filters through multiple correspondence candidates using a modified Random Sampling Consensus (RANSAC)-like procedure. Bülow and Birk (2011) propose another global registration method for 3D scan data, using spectral images and convolution in the Fourier domain. The solution proposed in this section is most closely related to these registration methods, but uses the 3D-NDT as an underlying spatial representation.

4.1. 3D-NDT histogram

In prior work, Magnusson et al. (2009a) propose the use of the 3D-NDT histogram as a place similarity metric for loop detection in SLAM systems. The basic idea of the 3D-NDT histogram is to describe a point set $\mathcal{P}$ by examining the shapes of the individual components in the 3D-NDT model $M_{\text{NDT}}(\mathcal{P})$. Each Gaussian component (μ_i, Σ_i) can be categorized, based on the properties of its covariance matrix Σ_i. Let $(\lambda_1, \lambda_2, \lambda_3)$ be the eigenvalues of Σ_i, with $\lambda_1 \geq \lambda_2 \geq \lambda_3 > 0$. If all three eigenvalues are roughly equal, the normal distribution, defined by Σ_i is roughly spherical. If, on the other hand, one of the eigenvalues is much smaller than the other two ($\lambda_1 \sim \lambda_2 \gg \lambda_3$), the distribution is disc shaped. Finally, if the largest eigenvalue is much greater than the other two values ($\lambda_1 \gg \lambda_2 \sim \lambda_3$), the distribution is line shaped. Using these observations, a threshold t_f can be defined to partition all distributions from $M_{\text{NDT}}(\mathcal{P})$ into three subsets: spherical, disc shaped and linear.

The 3D-NDT histogram h is then calculated, based on the number of linear, flat (disc-shaped) and spherical normal distributions in $M_{\text{NDT}}(\mathcal{P})$. Each of the three classes of distributions are further split into a number of bins. Thus,

$$h(M_{\text{NDT}}(\mathcal{P})) = \left(h_0^s \dots h_{|s|}^s, h_0^f \dots h_{|f|}^f, h_0^l \dots h_{|l|}^l\right) \quad (29)$$

where h^s, h^f and h^l are bins that accumulate the spherical, flat and linear distributions, respectively. The splitting of each of the three classes into further bins is necessary in order to provide a better discriminative power of the proposed descriptor. In the same way as in the previous work by Magnusson et al. (2009a), a single bin is used for the linear distributions, while the flat distributions are placed into bins according to the orientations of their normals. A preset number of directions (20 in this work, as opposed to 9 in the previous work by Magnusson et al. (2009a)) is computed according to the golden ratio spiral sphere coverage algorithm. An illustration of the construction of a histogram of orientations is shown in Figure 3. In order to obtain more precise results while using few orientation bins, the average direction of the distributions that fall inside each bin is also stored. Finally, the spherical distributions are split into bins according to their distance from the origin of the point set.

In order to better handle the variable density and discretization effects, three different histograms are computed for each 3D-NDT model. The approach used is the same as that proposed by Magnusson (2009): maintaining three histograms for each 3D-NDT, one for low, one for mid and one for high distance from the origin. Thus, the final 3D-NDT

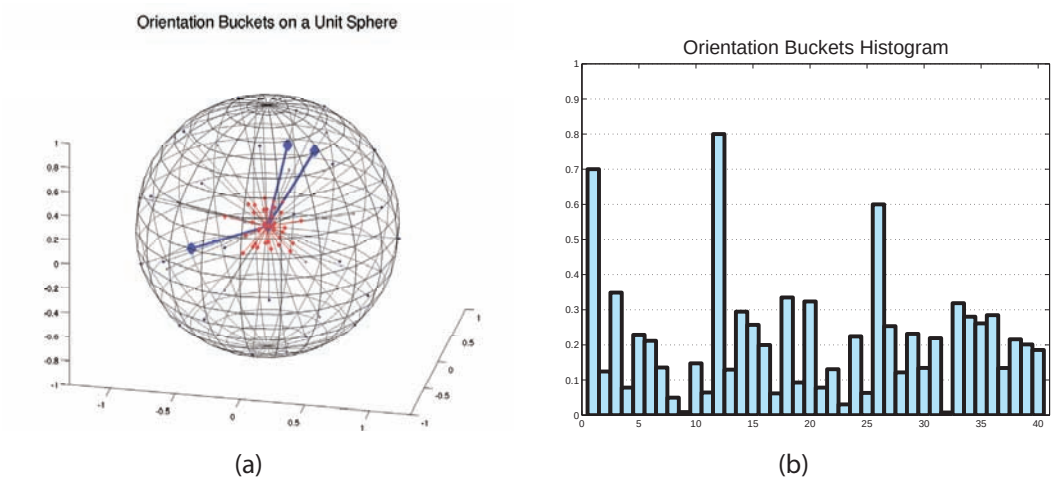

Fig. 3. The orientations of flat patches in a 3D-NDT scan can be split into buckets, depending on their orientation. This Figure presents an illustrative example of how the histogram is constructed. A unit sphere with the prototype bin directions (in black) is shown in Figure 3(a). The frequency of occurrence of flat patches, with a normal within the respective bins, is indicated by a scaled vector in red. The three most dominant directions are shown in blue. The respective histogram vector is constructed by listing the weights (or frequencies) of orientations within each bin and shown in Figure 3(b). Colour refers to the online version.

Algorithm 1 *Align*($\mathcal{M}$.$\mathcal{F}$)

$\boldsymbol{h}_1 \leftarrow h(M_{\text{NDT}}(\mathcal{M}))$
$\boldsymbol{h}_2 \leftarrow h(M_{\text{NDT}}(\mathcal{F}))$
non-max-suppress($\boldsymbol{h}_1, \boldsymbol{h}_2$)
$\mathcal{D}_1 = \{\text{dominantAvgDirections}(\boldsymbol{h}_1)\}$
$\mathcal{D}_2 = \{\text{dominantAvgDirections}(\boldsymbol{h}_2)\}$
for for all permutations *pm* of $(\mathcal{D}_1, \mathcal{D}_2)$ **do**
  $R_i \leftarrow$ closed-form solution using *pm*
  $\tilde{\boldsymbol{h}}_1 \leftarrow T(\boldsymbol{h}_1, R_i)$
  $s_i \leftarrow dist(\tilde{\boldsymbol{h}}_1, \boldsymbol{h}_2)$
return R_i associated with the largest similarity s_i

histogram computed contains 26 bins for each range (1 linear, 20 flat and 5 spherical bins), resulting in 78 bins/values that represent the whole 3D-NDT model.

4.2. Matching of 3D-NDT histograms

In this section, a method for finding a rotation matrix, which aligns two 3D-NDT histograms, is discussed. When using the 3D-NDT histogram for place recognition, Magnusson (2009) perform a step to make the histogram rotationally invariant. The authors assume that the dominant plane normal in a histogram is aligned with the *z*-axis, while the second most common orientation is in the *xy*-plane. In the context of this work, however, rotational invariance is undesirable. Instead, the algorithm attempts to align the dominant directions of a histogram $h(M_{\text{NDT}}(\mathcal{M}))$ to the dominant directions in $h(M_{\text{NDT}}(\mathcal{F}))$.

A simple algorithm to align two histograms is devised and described in Algorithm 1. In order to find the best alignment between two histograms $\boldsymbol{h}_1$ and $\boldsymbol{h}_2$, a pre-processing routine is first used to suppress bins that are not local maxima. This step is necessary, in order to counter discretization effects and to find independently dominant bins in each histogram. Next, the bins in the histograms $\boldsymbol{h}_1$ and $\boldsymbol{h}_2$ are sorted and the top n dominant bins selected. Representative dominant direction vectors $\mathcal{D}_1 = \{\boldsymbol{d}_1^1 \dots \boldsymbol{d}_n^1\}$ and $\mathcal{D}_2 = \{\boldsymbol{d}_1^2 \dots \boldsymbol{d}_n^2\}$ are then computed as the average of the normal vectors, collected in each dominant bin. Next, all possible permutations associating two directions $\boldsymbol{d}_a^1, \boldsymbol{d}_b^1$ from $\mathcal{D}_1$ with two directions $\boldsymbol{d}_a^2, \boldsymbol{d}_b^2$ from $\mathcal{D}_2$ are computed. As the number of directions considered is capped at n (with $n = 3$ used here), the number of permutations tested is limited (6 possible permutations here). Next, a closed-form solution for the relative rotation R_i can be easily computed. The selected pairs of dominant vectors implicitly define a plane, based on their cross products. Thus, the rotation R_i is chosen as the smallest rotation necessary in order to transform $\boldsymbol{d}_a^1 \times \boldsymbol{d}_b^1$ to $\boldsymbol{d}_a^2 \times \boldsymbol{d}_b^2$. The resulting rotational matrix is then used to transform $\boldsymbol{h}_1$ to $\tilde{\boldsymbol{h}}_1$. Finally, the histogram $\tilde{\boldsymbol{h}}_1$ that is most similar to $\boldsymbol{h}_2$ is selected, resulting in a purely rotational initial guess R_i.

In order to choose the best-matching alignment between $h(M_{\text{NDT}}(\mathcal{M}))$ and $h(M_{\text{NDT}}(\mathcal{F}))$, the similarity between two histograms is used. Given $\boldsymbol{h}_1$ and $\boldsymbol{h}_2$, the Euclidean distance $dist(\tilde{\boldsymbol{h}}_1, \boldsymbol{h}_2)$ can be used as a similarity metric. Since the spherical and linear bins h^s and h^l are rotation-invariant, it is not necessary to include them in the similarity metric for the purposes of the presented algorithm. Thus, in Algorithm 1 it is enough to transform the average directions of $\tilde{\boldsymbol{h}}_1$ and then compute the distance $\|T(\boldsymbol{h}_1, R_i) - \boldsymbol{h}_2\|_2$. In this manner, the candidate transformations can be sorted, according to the similarity between the transformed and target histogram.

5. Estimating registration covariance

As mentioned in the previous section, when an iterative registration algorithm converges to the optimal solution of its objective function, this does not guarantee a perfect match. Indeed, a small offset from the real pose is to be expected in all cases, and may accumulate in a significant error over multiple registrations. Thus, it is common to think of the solution returned by the iterative solver as the expected value of a random variable, which also has an associated covariance. Subsequently, this covariance matrix can be used in a probabilistic framework to achieve a globally consistent registration estimate.

Estimating what kind of covariance is to be expected for the solutions of different registration algorithms is thus an important task. One popular approach is to use the second-order Taylor expansion of the objective function as a model for the objective. Using this model, a least-squares solution can be computed and a closed-form expression for the covariance can be calculated based on the Hessian of the error function. This method was used by Bengtsson and Baerveldt (2003) to calculate the covariance of the 2D scan registration algorithm IDC. Censi (2007), however, convincingly argues that the shape of the error function close to the solution is not the only factor that influences the covariance of the result. In fact, it is important to also account for the variance of the objective function, with respect to the noise in the input point scan data. The rest of this section applies the conclusions of Censi (2007) to the problem of estimating the covariance of the poses produced with the 3D-NDT D2D algorithm.

5.1. Covariance equations

Given an estimate $\hat{\boldsymbol{p}}$ of the relative pose between two point scans, the covariance of the solution is given by Censi (2007, Equation (4)):

$$cov(\hat{\boldsymbol{p}}) \sim \left(\frac{\partial^2 f_{\mathrm{d2d}}}{\partial \boldsymbol{p}^2}\right)^{-1} \frac{\partial^2 f_{\mathrm{d2d}}}{\partial \boldsymbol{p} \partial \boldsymbol{z}} cov(\boldsymbol{z}) \frac{\partial^2 f_{\mathrm{d2d}}}{\partial \boldsymbol{p} \partial \boldsymbol{z}} \left(\frac{\partial^2 f_{\mathrm{d2d}}}{\partial \boldsymbol{p}^2}\right)^{-1} \tag{30}$$

where z denotes the sensor measurements used in computing the error function f_{d2d}. The component $\left(\frac{\partial^2 f_{\mathrm{d2d}}}{\partial \boldsymbol{p}^2}\right)^{-1}$ is just the inverse of the Hessian of the objective function and has already been computed in Equation (24). Thus, in order to compute the covariance $cov(\hat{\boldsymbol{p}})$, the covariance of the sensor measurements $cov(\boldsymbol{z})$ and the partial derivative of the Jacobian of f_{d2d} from Equation (20) with respect to $\boldsymbol{z}$ need to be computed.

In the work of Censi (2007), the measurements $\boldsymbol{z}$ are represented by the location of each sensor point in a polar coordinate system. The error function minimized by the 3D-NDT D2D algorithm, however, operates directly on the 3D-NDT representations of the point sets and thus reacts differently to noise. Let $\mathcal{X}$ be the set of points used to estimate a single component of the 3D-NDT. Then for each $\hat{\boldsymbol{x}} \in \mathcal{X}$ we have

$$\hat{\boldsymbol{x}} = \boldsymbol{x} + \boldsymbol{v} : \boldsymbol{v} \sim \mathcal{N}(0, \Sigma_s) \tag{31}$$

with $\boldsymbol{x}$ being the ground-truth point position and $\boldsymbol{v}$ white Gaussian noise with expected covariance Σ_s. Then the mean of the component $\boldsymbol{\mu}_i = E[\hat{\boldsymbol{x}}] = E[\boldsymbol{x}] + E[\boldsymbol{v}]$. Since $\boldsymbol{v}$ is zero mean, the value of $\boldsymbol{\mu}_i$ is statistically unaffected by the noise: $\boldsymbol{\mu}_i = E[\boldsymbol{x}]$. The covariance of the component Σ_i, however, depends on the covariance of the noise. Assuming that the noise is independent of the physical position of the points ($cov(\boldsymbol{x}, \boldsymbol{v}) = 0$), an expression for the covariance is achieved: $\Sigma_i = \Sigma_{\boldsymbol{x}} + \Sigma_s$. Therefore, in order to estimate $cov(f_{\mathrm{d2d}})$, the sensor observations considered are the covariances of each Gaussian component Σ_i. Thus, $cov(\boldsymbol{z}) = diag(\Sigma_s)$ and

$$\frac{\partial^2 f_{\mathrm{d2d}}}{\partial \boldsymbol{p} \partial \boldsymbol{z}} = \frac{d_1 d_2}{2} \left(\boldsymbol{\mu_{ij}}^{\mathrm{T}} B(BZ_a + Z_a B) B \boldsymbol{\mu_{ij}} - \boldsymbol{\mu_{ij}}^{\mathrm{T}} BB \boldsymbol{j_a} + \frac{d_2}{2} \boldsymbol{\mu_{ij}}^{\mathrm{T}} BB \boldsymbol{\mu_{ij}} \right) \exp\left(-\frac{d_2 \boldsymbol{\mu_{ij}}^{\mathrm{T}} B \boldsymbol{\mu_{ij}}}{2} \right) \tag{32}$$

is the expression for the partial derivatives of Equation (20) with respect to Σ_i. Thus, the only parameter to be set is the expected sensor-dependent point variance Σ_s.

6. Evaluation

This section presents the results from evaluating the proposed 3D-NDT D2D registration algorithm, including the initial guess and covariance estimation presented. The 3D-NDT histogram approach to initial alignment is evaluated first in a stand-alone manner in Section 6.1 and together with an iterative registration algorithm in Section 6.2. The latter section also serves to evaluate the 3D-NDT D2D algorithm implementation on two standard data sets and compare its performance against that of the 3D-NDT P2D and ICP implementations. Finally, the covariance estimation of the 3D-NDT D2D algorithm is evaluated in a manner similar to that proposed by Censi (2007).

6.1. Initial guess estimation

The evaluation presented in this section, as well as in Section6.3 is based on the use of a simulation environment. However, before starting the analysis of performance on simulated measurements, it is important to discuss one of the weaknesses of simulated experiments: namely the modelling of noise. A large variety of systematic and non-systematic errors can occur in different range sensor models. As the real-world data sets evaluated in the next section were acquired with SICK laser scanners, the subsequent analysis will focus on this particular class of devices. Ye and Borenstein (2002) characterize the noise present in the SICK LMS 200 when scanning a flat target mounted on

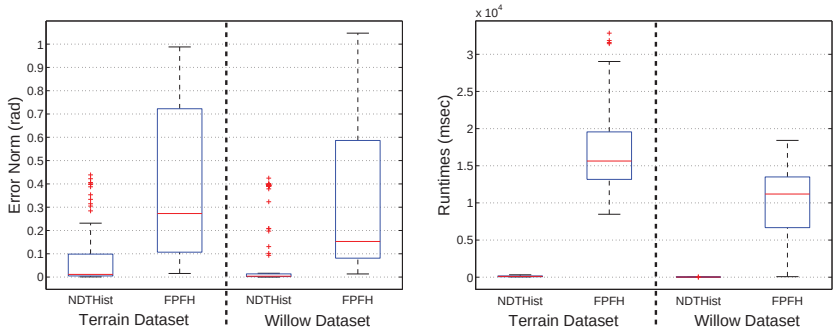

Fig. 4. Results of tests of initialization procedures over 160 scan pairs for each data set. Figure 4(a) shows the orientation errors for the 3D-NDT histogram and FPFH matching methods. Figure 4(b) shows the respective runtimes.

a high-precision linear motion table. The authors conclude that a small linear error is present, as well as a random noise offset. The noise variance depends on the grey level intensity of the target, as well as on the beam angle with respect to the normal of the scanned surface, but not on the measured distance. As Tuley et al. (2005) point out, another common error source occurs due to the so-called mixed measurements: range values that correspond to a reflection from multiple objects. The authors identify four modes of mixed measurements pertaining to the SICK LMS 200. The most common problem by far occurs when scanning points at edges of objects. Owing to the physical characteristics of the laser beam, light reflects both from the foreground and background, resulting in a measurement placed somewhere in between the two objects. Thus, the error models proposed by Ye and Borenstein (2002), as well as Tuley et al. (2005), were implemented and used to collect point set data in the ROS/Gazebo simulation engine.

Two simulated environments were used in order to collect evaluation data — the first is the standard ROS/Gazebo Willow garage office world, while the second is a custom-made outdoor environment representing an asphalt mill. The Willow world is a typical office environment with varying room sizes and corridors, but little clutter from furniture or objects. The second model features gravel piles, large open spaces and material collection bins and was developed for prototyping in the ALL-4-eHAM project.[1]

In order to test the 3D-NDT histogram initialization, 20 different positions were chosen in each environment. At each position 10 point clouds were acquired, spaced 15° apart. Skipping over one scan, this results in eight scan pairs, 30° apart at each position. The simulated 3D sensor used has a field of view of 180 × 120°, thus resulting in an overlap of about 80% between different scans. These scans were used as a evaluation set: the final pose $\hat{\boldsymbol{p}}$, returned by the initialization algorithm, is compared with the known ground truth orientation $\boldsymbol{p}$. As the discussed 3D-NDT histogram initial guess only corrects for orientation errors, the evaluation was only performed using scans taken from the same position (i.e. no translational component). Finally, the angular error of the estimate $\epsilon(\hat{\boldsymbol{p}})$ can be defined and measured as the smallest angle of an arbitrary-axis rotation that brings the predicted pose to the ground truth pose.

Plots of the angular error ϵ obtained by the 3D-NDT histogram and FPFH on the two data sets are shown in Figure 4(a), while the respective runtimes are shown in Figure 4(b). The FPFH implementation from the Robot Operating System (ROS) (Quigley et al., 2009) was used as a baseline for comparison against a feature-based approach. As expected, the runtime of the 3D-NDT histogram initialization is substantially lower, with average runtimes on the order of 150 ms, as opposed to about 15 seconds for FPFH. In addition, both initialization algorithms perform much better on the indoor data set 'Willow' compared with the outdoor set. The poorer performance in outdoor environments is often due to the lower number of geometric features available, resulting in more difficulties in disambiguation. On average, the 3D-NDT histogram-based initialization results in more accurate and faster computation of an initial orientation $\hat{\boldsymbol{p}}$.

6.2. *Registration*

The P2D and D2D versions of the 3D-NDT registration algorithm were implemented and integrated as a package for ROS (Quigley et al., 2009), and the source code developed is published online.[2] The versions used for comparison in this work evaluate the likelihood of only the closest components, exploiting the sparsity of the registered 3D-NDT models. The parameters d_1 and d_2 for 3D-NDT P2D are computed in the same way as in previous work by Magnusson (2009), while those for 3D-NDT D2D are set to $d_1 = 1$ and $d_2 = 0.05$ respectively, as in Stoyanov et al. (2012). All tests were performed with iterative optimization at 3D-NDT grid sizes of 4, 2, 1 and 0.5 m. As a baseline for comparison, a standard off-the-shelf ICP implementation from the Point Cloud Library (PCL)[3] was used. The implementation used is representative of an average ICP performance, with speed up achieved through

Table 1. Properties of the considered data sets.

Data set	AASS	Hannover
Number of point clouds	60	923
Average points per scan	∼ 90,000	∼ 15,000
Average offset		
— translation	1.85 m	1.35 m
— rotation	0.12 rad	0.06 rad

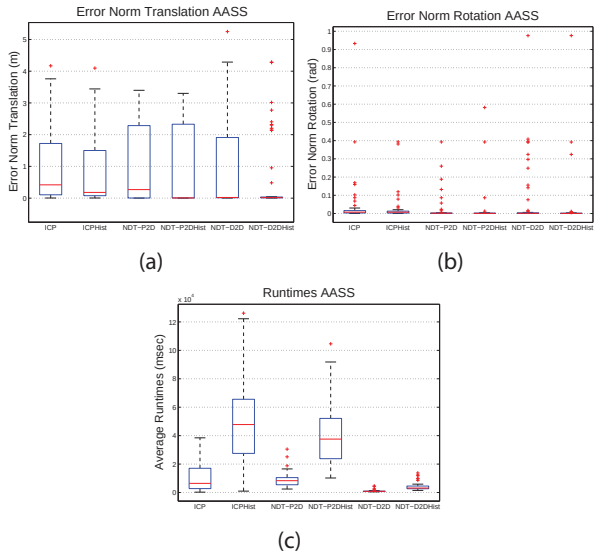

Fig. 5. Results of registration algorithms evaluating the AASS data set, 59 scan pairs. Each figure shows boxplots for six different combinations: three registration algorithms with and without using an initial guess from the 3D-NDT histogram. Boxes are centred at average values, red lines show the median value. (a) Norm of the translation error. (b) Orientation errors. (c) Respective runtimes.

kd-tree nearest-neighbour queries and increased robustness through rejection of inconsistent correspondences. Comparison with other state-of-the-art ICP algorithms (such as Generalized ICP (Segal et al., 2009) or 3DTK[4] fast ICP implementations of Nüchter et al. (2007)) or global registration methods such as spectral registration (Bülow and Birk, 2011) or plane matching (Pathak et al., 2010) have not been performed at present and will be investigated as future work.

In our prior work (Stoyanov et al., 2012) the 3D-NDT D2D registration algorithm was evaluated on several typical range scan pairs from an indoor environment. These evaluations used five point clouds from the AASS loop data set (available online[5]) and generated for each a set of 343 initial test positions, away from the optimal solution. The purpose of this evaluation was to test the robustness of the proposed registration algorithm to a large set of possible disturbances. For completeness, the results and analysis of this evaluation are included at the end of this subsection. In addition, the 3D-NDT P2D, D2D and the ICP algorithm implementations were evaluated on two large data sets with many input point clouds: namely the full AASS loop and Hannover-2 data sets from the 3D scans online repository. The reference ICP and 3D-NDT P2D implementations both operate on point clouds, acquired from real-world range sensors. Thus, in order to avoid problems incurred by variable point sample density, the input point sets for these algorithms were regularly subsampled, using a 0.1 m grid.

Both of the data sets used include points sampled using a rotating SICK laser, from two different environments. In both cases, a relatively reliable ground truth data exists for the real positions at which the scans were taken. However, as mentioned earlier, due to the nature of the pose space $SE(3)$ it is not possible to reliably compare a test and reference trajectories. It is possible, however, to compute a relative metric for each pair of registered point sets. Thus, the registration algorithms evaluated in this section are evaluated on a set of scan pairs, and not on the full final trajectory for the complete data sets. The properties of the two data sets are summarized in Table 1.

The results of running the three registration algorithms (with and without NDT histogram initialization) are shown in Figure 5 for the AASS and Figure 6 for the Hannover2 data set. In each case, separate boxplots are generated for the average translational and rotational errors. The box plots generated show a box that encloses a 75th percentile of the data, with the median value marked by a red line. Whiskers extend to the furthest point not considered an outlier, i.e. all data within three standard deviations.

The results on the AASS loop data set support the claim that the 3D-NDT D2D registration algorithm returns final alignments comparable to those of ICP and 3D-NDT-P2D. Unfortunately, the translational errors of all three algorithms (without initial alignment) exhibit a large variance, suggesting that the optimization procedure often falls into local minima. The rotational errors on the other hand are much lower, suggesting that local minima are likely occurring for point scan pairs located in corridors or other areas with low translational features. Indeed upon inspection of the registration results, two major failure modes are observed: one due to locally similar geometries and one due to feature poor environments, such as corridors. In the second case, in particular, rotational errors remain low, while translation along the corridor walls can vary substantially, producing larger translational errors. The error plots also support the claim that the 3D-NDT Histogram initial guess results in improved results for all three iterative registration algorithms. It is important to also note that the histogram initialization results were used in a consecutive manner, i.e. the top three initial poses and the zero pose were all used as initial points for a registration, with the best-matching (lowest error score) registered pose reported as a final result. Thus, the initialization can cost up to four times more in runtime, as four registrations are performed instead of one. The runtimes of the evaluated algorithms (Figure 5(c) follow the expected pattern, with 3D-NDT-D2D performing

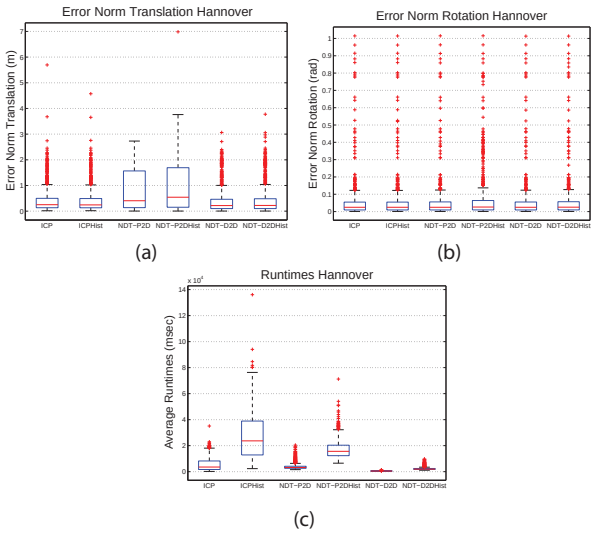

Fig. 6. Results of registration algorithms evaluating the Hannover data set, 922 scan pairs. (a) Norm of the translation error. (b) Orientation errors. (c) Respective runtimes.

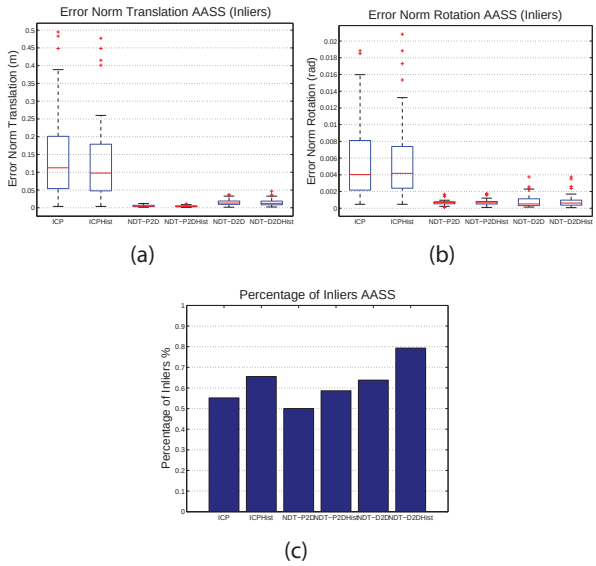

Fig. 7. Results of registration algorithms evaluating the AASS data set, 59 scan pairs. Considering only inliers. (a) Norm of the translation error. (b) Orientation errors. (c) Fraction of scan combinations considered inliers.

almost an order of a magnitude faster than both 3D-NDT D2D and ICP.

The results of running the evaluated algorithms on the Hannover data set are shown in Figure 6. The results of all six algorithm combinations are noticeably worse on this data set although the average translation and rotation difference is smaller. The fact that Hannover2 is an outdoor data set with substantially fewer and sparser point data makes this a more challenging test case. Although the translational errors recorded are comparable to those in the AASS data set, large numbers of outliers are also present. The same is true for the rotational errors observed, which also show a higher average error. As the large number of outliers obscures the performance of the evaluated algorithms, the results for only the successfully registered scans were also plotted.

Figures 7 and 8 show the translational/rotational error plots only for the scan pairs that were considered successfully registered. The criteria for a successful registration were chosen as a translational error of 0.5 m or less and a rotational error of 0.2 rad or less. Notably these criteria are rather relaxed, but nevertheless sufficient to accent on the behaviour of each registration algorithm in the vicinity of the true solution. Figure 7 shows that when an inlier solution is reported, all six evaluated algorithms reliably converge to the immediate neighbourhood of the ground-truth solution. Figure 7(c) also shows the ratio of successfully registered point clouds. The trend in this data set is clearly visible: the largest percentage of successfully registered scans also exhibit the lowest errors and are obtained by 3D-NDT D2D with histogram initialization. The results of successfully registered scans for the Hannover data set are shown in Figure 8. In this case as well, it is easier to interpret the graph of successful registrations. Both the translational and

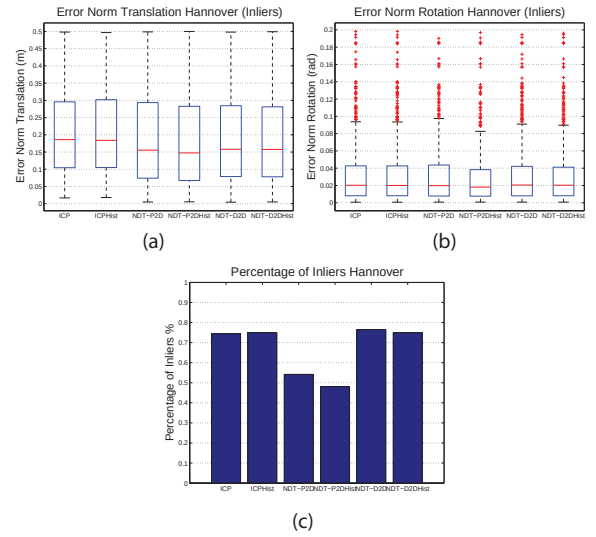

Fig. 8. Results of registration algorithms evaluating the Hannover data set, 922 scan pairs. Considering only inliers. (a) Norm of the translation error. (b) Orientation errors. (c) Fraction of scan combinations considered inliers.

rotational errors on this data set are comparable over all six algorithms tested, with almost identical median value offset. This suggests a systematic bias, that may be due to many factors, ranging from a offsets in the ground truth data to an inherently difficult registration problems and discretization/sparsity issues. The histogram initialization does not actually improve results on this data sets, in fact in the case of 3D-NDT P2D it leads to a slight drop in successfully registered scans. The most likely cause of this poor performance is the low average rotational offset between

scans. Thus, additionally refining the initial rotation guess does not result in an improvement and might even result in suboptimal solutions.

As mentioned earlier, the 3D-NDT D2D algorithm consistently performs faster than both ICP and the 3D-NDT P2D implementations tested. These results are evident in both the AASS (Figure 5(c)) and Hannover (Figure 6(c)) data sets, and thus it is important to comment on the reasons behind this speed-up. First and foremost, the speed of the D2D algorithm is due to the compactness of the 3D-NDT as a spatial representation. A typical 3D scan in the AASS data set contains of the order of 90,000 points. Even after sub-sampling, both ICP and the P2D algorithms have to iterate over a large number of points, in order to evaluate the objective functions minimized. All three of the evaluated algorithms spend time looking for closest neighbours: a single evaluation of ICP's objective takes $\mathcal{O}(n \log n)$ operations, with n being the number of points in the point cloud $\mathcal{P}$. The P2D algorithm needs $\mathcal{O}(n \log q)$ operations, with $q \ll n$ being the number of Gaussian components in the 3D-NDT model of the fixed point set. On a typical point cloud from the AASS data set, about 1,500 Gaussian components are computed at the finest discretization level of 0.5 m. The D2D registration algorithm only requires $\mathcal{O}(q \log q)$ operations per evaluation of the objective. In fact, the time for registering two point sets with 3D-NDT D2D is usually dominated by the time necessary to construct the 3D-NDT representations of the input point sets.

Finally, we discuss the results for the sensitivity of the registration algorithms to different initial positions, as presented in Stoyanov et al. (2012). Using the ground-truth relative poses and five test scan pairs from the AASS data set, test scenarios for the registration algorithms were generated. One of the point sets in the pair was transformed away from the ground-truth position and then used as an input to the evaluated registration algorithm. For each of the five test pairs, offsets of between -1.5 and 1.5 m in steps of 0.5 m (along the x- and y-axes) were introduced. In addition, the scans were rotated around the z-axis by an angle varying from -30° to 30° (in steps of 10°). Thus, a total of 343 test transformations (combinations of seven distinct points along t_x, t_y and r_z) were generated and applied to the scan point set $\mathcal{S}$. Registration was judged to be successful if the output transformation brought the scan $\mathcal{S}$ within 0.2 m and 0.05 rad from its ground-truth position.

The performance of the discussed registration algorithms is summarized in Figure 9. Overall, the newly proposed 3D-NDT D2D algorithm slightly outperforms the Point-to-Distribution implementation tested in this work. The performance of the baseline ICP algorithm is very good on scan pairs 2, 3 and 4 and extremely poor on scan pairs 1 and 5. It is worth noting that the problematic scan pairs represent a typical failure case for ICP: a corridor environment with largely overlapping, feature poor scans. Figure 9(d) and (e) summarize the mean translational and rotational errors over all successfully registered point set pairs.

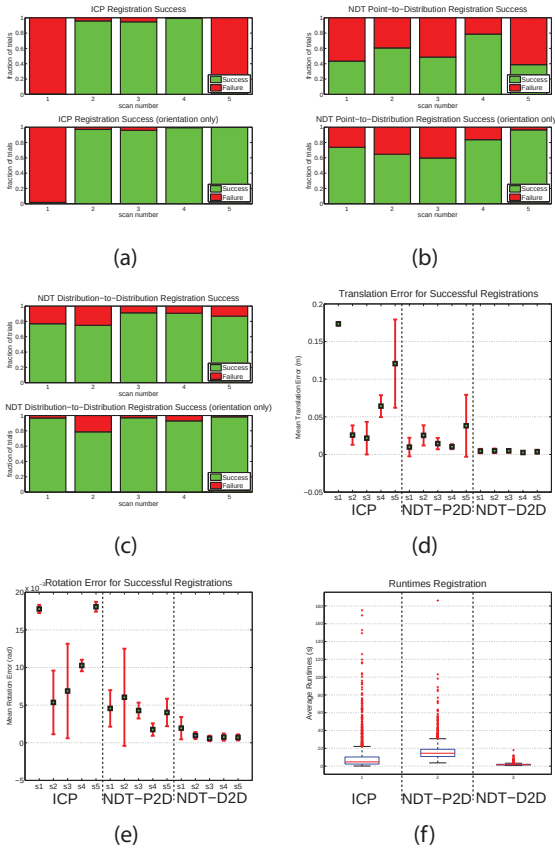

Fig. 9. Summary of results of the evaluated registration methods, over all five scan pairs. (a) Point-to-Distribution 3D-NDT success rates with (top) and without (bottom) considering translation errors. (b) Distribution-to-Distribution success rates. (c) ICP success rates. (d) Mean translation and (e) rotation error for the successfully registered scans. (f) Runtimes distribution for the three algorithms.

The D2D algorithm proposed exhibits a stable performance over all tests, with a small variance in the error scores. Note that for the P2D and especially for the ICP algorithm, the results for scan pairs 1 and 5 were obtained from a much smaller sample of successful registrations.

6.3. *Covariance estimation*

The covariance estimation procedure, proposed in Section 5 was evaluated in the same way as in the original work by Censi (2007). The Gazebo simulation in ROS and the Willow garage office map were used. A path of 15 poses was defined, and 100 scans were taken, with noise parameters set as in Section 6.1. Next, for each pair of poses, an initial guess $\boldsymbol{p}_i \sim \mathcal{N}(0, \Sigma_o)$ was sampled and used to model an odometry system. The variances of the initial guess were set to 0.3 m in each translational component and 0.1 rad for each rotational component. The output registration pose of ICP, 3D-NDT P2D and 3D-NDT D2D were recorded and

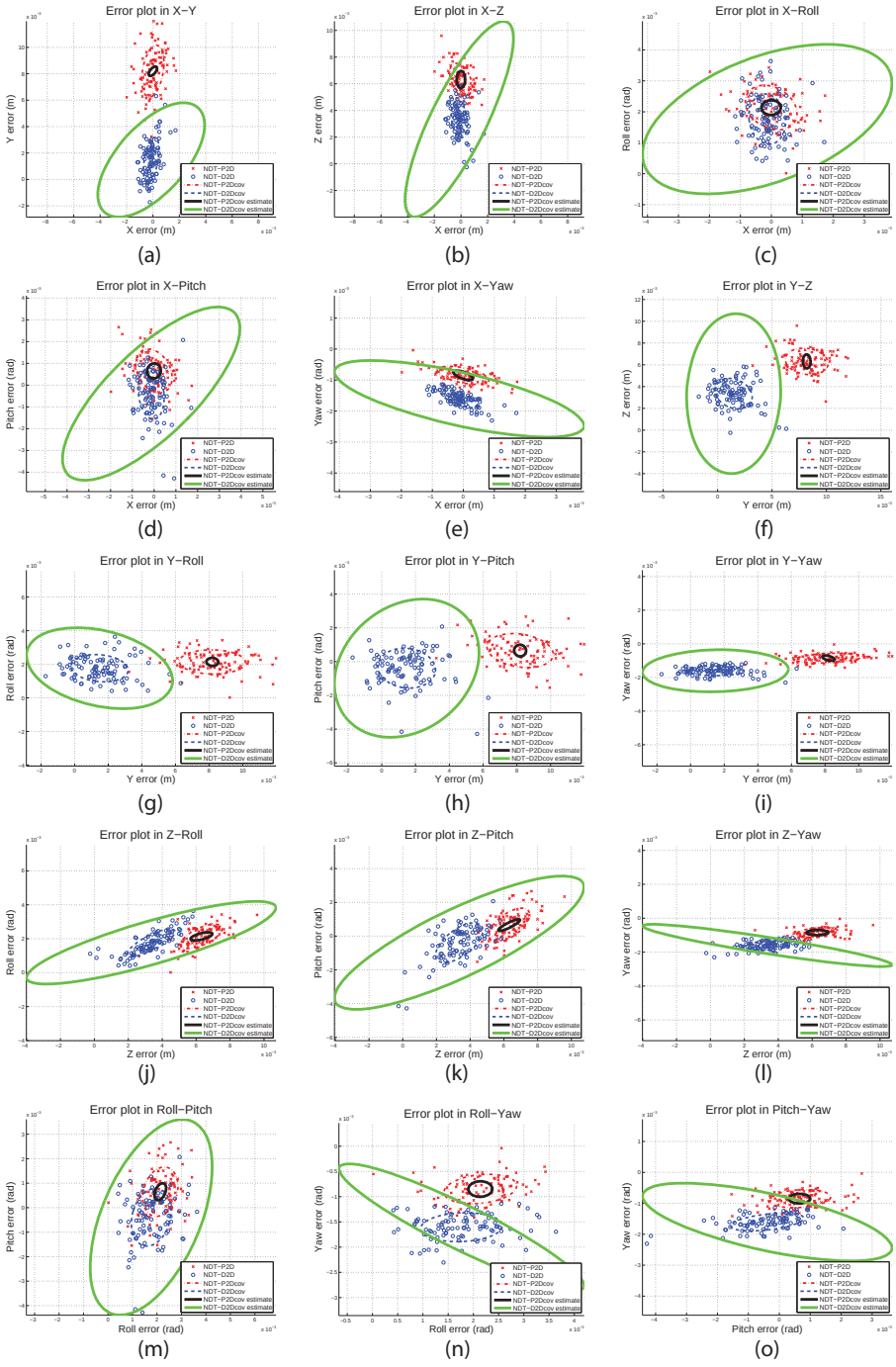

Fig. 10. Projections of the 6D covariance matrix and its estimates in different planes. Individual components of the error poses for 3D-NDT P2D and D2D, marked as red crosses and blue circles, respectively. Sample covariance and covariance estimates drawn as 1σ level sets. Sample covariance of P2D algorithm shown as a red dotted line, Hessian estimate of the covariance as a solid black line. Sample covariance for D2D algorithm shown as a blue dotted line, while estimate is shown as a black dotted line.

used to calculate a set of relative error poses for each scan pair. Finally, the ground truth pose $\boldsymbol{p}$ was used as an input to compute the estimate $cov(f_{\text{d2d}}(\boldsymbol{p}))$ according to Equation (32). The covariance of the 3D-NDT P2D algorithm was set to the inverse of the Hessian of f_{p2d}.

Results of the covariance comparison for one sample pair of scans are shown in Figure 10. Owing to a consistently large offset present in the ICP solution and the fact that no estimate of the covariance of ICP was implemented, the ICP points are not shown in Figure 10. Notably, in contradiction with the observations of Censi, the Hessian-based approach underestimates the covariance of the solutions of f_{p2d}. In contrast, the proposed estimate according to Equation (32) overestimates the observed covariance.

Although this behaviour is unexpected, we argue that at the scale of variance observed in the test examples, the over-estimate does not pose a significant problem. In fact, in the authors' experience slightly over-estimating the covariance matrix is preferable to slight under-estimation, as it allows for more freedom in a possible graph-optimization SLAM application and does not allow for over-confident registrations.

7. Conclusion and future directions

In this article we have presented a novel approach to point cloud registration, based on minimizing the distance between 3D-NDT models. The proposed 3D-NDT D2D registration algorithm was evaluated extensively on standard point cloud data sets and was found to be substantially faster than the P2D version and the standard ICP implementation. Moreover, the accuracy of 3D-NDT D2D was demonstrated to be on a par with and in some cases even substantially higher than that of ICP. This article also presented and evaluated two additional improvements to the 3D-NDT D2D registration algorithm, namely the use of a 3D-NDT-based global descriptor to estimate good initial poses for registration and an estimate of the algorithm's covariance. A fast method to align two 3D-NDT distributions was developed, based on the 3D-NDT histogram. In addition, the approach proposed by Censi (2007) was applied to the 3D-NDT D2D objective function, producing an inexpensive, closed-form method to estimate the match covariance.

Although the presented algorithm exhibits a substantially improved performance on the evaluated test data sets, several limitations still remain and should be addressed. Most notably, being an iterative approach, the proposed algorithm still tends to be susceptible to local minima in the objective function. Thus, additional improvements of the 3D-NDT histogram method of finding initial poses would result in a substantially more stable registration algorithm. Another important future work direction is to test the performance of the proposed algorithm on different sensor data, most notably using noisier sensors with a narrower field of view. Finally, an investigation on the performance of a complete mapping system based on the 3D-NDT D2D registration algorithm will be further pursued. By coupling the three components presented in this work (the D2D registration algorithm with 3D-NDT histogram initialization and covariance estimation) with a loop detection system, such as that proposed by Magnusson et al. (2009a), and a global pose-graph optimization scheme we will strive to achieve a fast, accurate and stable mapping system, based solely on the 3D-NDT.

Notes

1. See http://aass.oru.se/Research/Learning/all4eham/index.html.
2. See http://code.google.com/p/oru-ros-pkg/.
3. See http://pointclouds.org.
4. See http://slam6d.sourceforge.net/.
5. See http://kos.informatik.uni-osnabrueck.de/3Dscans/.

Funding

This research received partial funding from the European Union Framework Program 7 project RobLog (grant number FP7 - 270350) and the Kunskaps och Kompetensutveckling Stiftelsen project SAUNA (grant number 20100315).

References

Andreasson H, Magnusson M and Lilienthal AJ (2007) Has something changed here? Autonomous difference detection for security patrol robots. In *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2007)*, pp. 3429–3435.

Bengtsson O and Baerveldt A-J (2003) Robot localization based on scan-matching—estimating the covariance matrix for the IDC algorithm. *Robotics and Autonomous Systems* 44(1): 29–40.

Besl P and McKay N (1992) A method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 14: 239–256.

Biber P and Straßer W (2003) The normal distributions transform: a new approach to laser scan matching. In *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2003)*, vol. 3, pp. 2743–2748.

Bülow H and Birk A (2011) Spectral registration of noisy sonar data for underwater 3D mapping. *Autonomous Robots* 30: 307–331.

Censi A (2007) An accurate closed-form estimate of ICP's covariance. In *Proceedings of IEEE International Conference on Robotics and Automation (ICRA 2007)*, pp. 3167–3172.

Gelfand N, Mitra NJ, Guibas LJ and Pottmann H (2005) Robust global registration. In *Proceedings Eurographics Symposium on Geometric Processing*, SGP '05, Aire-la-Ville, Switzerland, Switzerland, 2005. Eurographics Association.

Goldberger J, Gordon S and Greenspan H (2008) An efficient image similarity measure based on approximations of KL-divergence between two Gaussian mixtures. In *International Conference on Computer Vision (ICCV 2003)*, pp. 487–493.

Granger S and Pennec X (2006) Multi-scale EM-ICP: a fast and robust approach for surface registration. In *Computer Vision (ECCV 2002)* (*Lecture Notes in Computer Science*, vol. 2353). New York: Springer, pp. 69–73.

Hershey J and Olsen P (2007) Approximating the Kullback Leibler divergence between Gaussian mixture models. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2007)*, volume 4.

Hertzberg C, Wagner R, Frese U and Schröder L (2011) Integrating generic sensor fusion algorithms with sound state representations through encapsulation of manifolds. *Information Fusion*, in press. DOI: 10.1016/j.inffus.2011.08.003.

Horn BKP (1987) Closed-form solution of absolute orientation using unit quaternions. *Journal of the Optical Society of America A* 4: 629–642.

Huber M, Bailey T, Durrant-Whyte H and Hanebeck U (2008) On entropy approximation for Gaussian mixture random vectors. In *International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI 2008)*, pp. 181–188.

Jian B and Vemuri B (2010) Robust point set registration using Gaussian mixture models. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 33(8): 1633–1645.

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

Julier, S (2007) An empirical study into the use of Chernoff information for robust, distributed fusion of Gaussian mixture models. In *IEEE International Conference on Information Fusion, 2006*, pp. 1–8.

Magnusson M (2009) *The Three-Dimensional Normal-Distributions Transform — an Efficient Representation for Registration, Surface Analysis, and Loop Detection*. Ph.D. thesis, Örebro University.

Magnusson M, Andreasson H, Nüchter A and Lilienthal AJ (2009a). Automatic appearance-based loop detection from 3D laser data using the normal distributions transform. *Journal of Field Robotics* 26: 892–914.

Magnusson M, Duckett T and Lilienthal AJ (2007) 3D scan registration for autonomous mining vehicles. *Journal of Field Robotics* 24: 803–827.

Magnusson M, Nüchter A, Lörken C, Lilienthal AJ and Hertzberg J (2009b) Evaluation of 3D registration reliability and speed - a comparison of ICP and NDT. In *Proceedings IEEE International Conference on Robotics and Automation (ICRA 2009)*, pp. 3907–3912.

Makadia A, Patterson A and Daniilidis K (2006) Fully automatic registration of 3D point clouds. In *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 1297–1304.

Nüchter A, Lingemann K, Hertzberg J and Surmann H (2007) 6D SLAM—3D mapping outdoor environments. *Journal of Field Robotics* 24: 699–722.

Pathak K, Birk A, Vaskevicius N and Poppinga J (2010) Fast registration based on noisy planes with unknown correspondences for 3-D mapping. *IEEE Transactions on Robotics* 26: 424–441.

Quigley M, Conley K, Gerkey BP, et al. (2009) ROS: an open-source Robot Operating System. In *ICRA Workshop on Open Source Software*.

Rusu R, Blodow N, Marton Z and Beetz M (2008) Aligning point cloud views using persistent feature histograms. In *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2008)*, pp. 3384–3391.

Rusu RB, Blodow N and Beetz M (2009) Fast Point Feature Histograms (FPFH) for 3D registration. In *IEEE International Conference on Robotics and Automation (ICRA 2009)*, pp. 3212–3217.

Segal A, Haehnel D and Thrun S (2009) Generalized-ICP. In *Proceedings of Robotics: Science and Systems*, Seattle, WA.

Steder B, Ruhnke M, Grzonka S and Burgard W (2011) Place recognition in 3D scans using a combination of bag of words and point feature based relative pose estimation. In *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2011)*.

Stoyanov T, Magnusson M, Almqvist H and Lilienthal AJ (2011) On the accuracy of the 3D normal distributions transform as a tool for spatial representation. In *Proceedings of IEEE International Conference on Robotics and Automation (ICRA 2011)*, pp. 4080–4085.

Stoyanov T, Magnusson M, Andreasson H and Lilienthal AJ (2010) Path planning in 3D environments using the normal distributions transform. In *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2010)*, Taipei, Taiwan, pp. 3263–3268.

Stoyanov T, Magnusson M and Lilienthal AJ (2012) Point set registration through minimization of the distance between 3D-NDT models. In *Proceedings of IEEE International Conference on Robotics and Automation (ICRA 2012)*, pp. 5196–5201.

Tsin Y and Kanade T (2004) A correlation-based approach to robust point set registration. In *Computer Vision (ECCV 2004)* (*Lecture Notes in Computer Science*, vol. XXXX). New York: Springer, pp. 558–569.

Tuley J, Vandapel N and Hebert M (2005) Analysis and removal of artifacts in 3-D LADAR data. In *Proceedings of the 2005 IEEE International Conference on Robotics and Automation, 2005 (ICRA 2005)*, pp. 2203–2210.

Wells W III, Viola P, Atsumi H, Nakajima S and Kikinis R (1996) Multi-modal volume registration by maximization of mutual information. *Medical Image Analysis* 1: 35–51.

Ye C and Borenstein J (2002) Characterization of a 2D Laser Scanner for Mobile Robot Obstacle Negotiation. In *Proceedings IEEE International Conference on Robotics and Automation, 2002 (ICRA '02)*, vol. 3, pp. 2512–2518.