2012 IEEE/RSJ International Conference on
Intelligent Robots and Systems
October 7-12, 2012. Vilamoura, Algarve, Portugal

Scan Registration with Multi-Scale K-Means Normal Distributions Transform

Arun Das* and Steven L. Waslander†
University of Waterloo, Waterloo, ON, Canada, N2L 3G1

Abstract— **The normal distributions transform (NDT) scan registration algorithm has been shown to produce good results, however, has a tendency to converge to a local minimum if the initial parameter error is large. In order to improve the convergence basin for NDT, a multi-scale k-means NDT (MSKM-NDT) variant is proposed. This approach divides the point cloud using k-means clustering and performs the optimization step at multiple scales of cluster sizes. The k-means clustering approach guarantees that the optimization will converge, as it resolves the issue of discontinuities in the cost function found in the standard NDT algorithm. The optimization step of the NDT algorithm is performed over a decreasing scale, which greatly improves the basin of convergence. Experiments show that this approach can be used to register partially overlapping scans with large initial transformation error.**

I. Introduction

Scan registration algorithms play an important role in mapping techniques for mobile robots. Many sensors such as RGBD cameras, LIDAR and stereo cameras, provide information about the robot's environment in the form of point-clouds. In applications where the point cloud data has overlapping geometry, scan registration techniques can be used to determine their relative transformation. This assists in the aggregation of point cloud data, which is commonly required for autonomous operation in unknown environments and allows for a wide range of applications such as autonomous mining, surveillance, search, rescue, and exploration.

Many SLAM solutions rely on algorithms such as iterative closest point (ICP) to estimate the relative transformation between two overlapping point clouds. ICP was independently introduced by Besl and McKay [1], Chen and Medioni [2], and Zhang [3]. The ICP algorithm attempts to find transform parameters that minimize the Euclidean distance of corresponding points, which are assumed to be the nearest neighbour points.

Standard ICP does not take into account the underlying surface structure of the point cloud. To address this, Segal et al. introduced the idea of generalized-ICP [4]. This approach calculates surface structure using local neighbourhoods of points and uses this additional information in the optimization process to generate higher quality scan matches. One of the shortcomings of these ICP approaches is the generation of the nearest neighbour correspondences. This step is generally computationally expensive and has been shown to be the bottleneck of the ICP algorithm [5].

The normal distributions transform (NDT) is a relatively new approach, first suggested by Biber and Strasser for scan registration and mapping [6]. It was later expanded into 3D for mine-mapping tasks by Magnusson et al. [7]. This method is desirable because it does not require the computation of explicit point correspondences. The NDT algorithm represents the underlying scan as a set of Gaussian distributions that locally model the surface as a probability density function (PDF). It is an optimization based approach which scores a parameter estimate, p, by transforming a point cloud by p and evaluating it at the PDF. In order to register two point clouds, NDT seeks the transformation parameters such that this score is maximized.

The standard formulation of NDT results in a nonlinear optimization which is susceptible to local minima. A projection of the 3D cost function onto the x and θ plane for NDT is presented in Figure 1. In this work, a method to improve the convergence basin of NDT is presented. The grid based division of the scan is replaced with a multi-scale k-means clustering technique. This clustering over scales captures coarse-to-fine details in the scan. The k-means clustering of the scan removes discontinuities in the cost function caused by the cell boundaries found in the standard NDT algorithm. This guarantees that the nonlinear optimization is well defined and all standard nonlinear optimization algorithms will converge to a solution satisfying necessary conditions for optimality [9]. In order to increase the basin of convergence, the optimization step is performed across varying scales, as the optimization for the current scale is initialized with the solution from the previous scale. This course-to-fine method avoids local minima and is experimentally shown to reliably converge to the global minimum for large initial transformation errors. K-means clustering of the point clouds has been attempted previously by Magnusson, but at a single scale, and only evaluating the scan points at the closest Gaussian cluster [10]. The selection of the cluster size is also an issue, as too small or too large cluster sizes results in poor registration results.

II. Problem Formulation

A. 2D Registration

Scan registration algorithms seek to find the optimal transformation between two point cloud scans, a *reference scan* and a *scene scan*. The goal is to determine transformation parameters which best align the two, such that the two scans

* M.A.Sc. Candidate, Mechanical and Mechatronics Engineering, University of Waterloo; adas@uwaterloo.ca.
† Assistant Professor, Mechanical and Mechatronics Engineering, University of Waterloo; stevenw@uwaterloo.ca

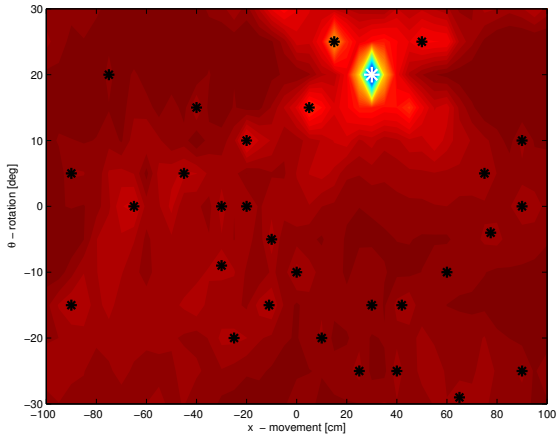


Fig. 1. Cost surface for NDT registration, for varying x and θ parameters. The black markers denote the local minima and the white marker denotes the global minimum.

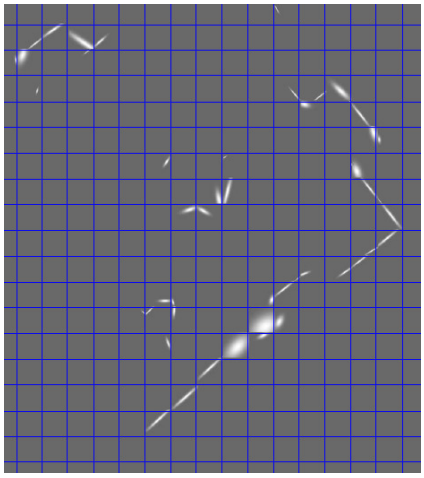


Fig. 2. Probability density of the laser scan from NDT for each cell. Lighter areas indicate higher probabilities of sample points

overlap as much as possible. Denote the reference scan as the set of points $Y = \{y_1, \ldots, y_{N_Y}\}$ where $y_i \in \mathbb{R}^2$ for $i \in \{1, \ldots, N_Y\}$ and the scene scan as $X = \{x_1, \ldots, x_{N_X}\}$ where $x_j \in \mathbb{R}^2$ for $j \in \{1, \ldots, N_X\}$. The transformation that maps a point x from the scene scan to the coordinate frame of the reference scan using a parameter estimate, $p = [t_x, t_y, \theta]^T \in \mathbb{R}^2 \times [0, 2\pi]$, is given as

$$T_Y(p, x) := \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} x + \begin{bmatrix} t_x \\ t_y \end{bmatrix} \quad (1)$$

where θ is the rotation between the two frames and $[t_x, t_y]^T$ is the translation.

B. Normal Distributions Transform

The Normal Distributions Transform (NDT) is a method by which sections of a point cloud are mapped and represented as Gaussian distributions within a grid structure. The NDT registration algorithm begins by subdividing the space occupied by the reference scan into a set of grid cells c (grids in 2D and cubes in 3D). Denote the collection of the reference scan points in cell c_i as $Y^{c_i} = \{y_1^{c_i}, \ldots, y_{N_{c_i}}^{c_i}\}$ such that $y_k^{c_i} \in c_i$ for $k \in \{1, \ldots, N_{c_i}\}$. For each cell c_i, the points of the reference scan occupying that cell are used to generate a mean μ_{c_i} and a covariance matrix Σ_{c_i} for a representative Gaussian distribution $\mathcal{N}(\mu_{c_i}, \Sigma_{c_i})$. This PDF can be interpreted as a generative process that models the local surface points Y^{c_i} within the cell. Assuming that the locations of the reference scan surface points are drawn from this distribution, the likelihood of having measured a point y, within cell c_i, can be modelled as $\rho(y) = \exp\left(-\frac{(y-\mu_{c_i})^T \Sigma_c^{-1} (y-\mu_{c_i})}{2}\right)$, where μ_{c_i} and Σ_{c_i} are the sample mean and covariance for cell c_i. An NDT probability distribution plot of a 2D laser scan is given in Figure 2.

Since the point cloud is now modelled by a piecewise continuous and piecewise differentiable summation of Gaussians, numerical optimization tools can be used in order to register the scene scan with the reference scan. A fitness cost, $s : \mathbb{R}^2 \times [0, 2\pi] \rightarrow \mathbb{R}$, can be calculated which quantifies the measure of overlap between scans X and Y, as $s(p) = -\sum_{k=1}^{N_x} \rho_Y(T(p, x_k))$. This denotes the sum of probabilities that each point in scan X is drawn from the NDT representation of Y, when transformed by parameter guess p. Using this cost, the goal of the NDT scan registration algorithm is to find a parameter estimate, p, such that the cost, s, is minimized. It should be noted that the cost function, $s(p)$, has an analytic gradient, g and Hessian, H, contribution for each point.

While performing the optimization, arbitrarily small changes in the parameter estimate could cause the scene scan points to cross cell boundaries, resulting in discontinuities in the overall cost function. This implies that the cost function is non-smooth and the gradient and Hessian do not exist in the transformation space at these cell boundaries. Magnusson et al. address this using a tri-linear interpolation scheme which takes the influence of neighbouring cells into account when calculating the gradient and Hessian contribution from a single scene point [8]. However, this does not solve the issue, as a boundary crossing could still cause a finite change in the cost function for an arbitrarily small change in the transformation parameters.

In the standard NDT algorithm, the optimization is initialized with a parameter estimate p_0 and is executed until some convergence criteria such as $|g| < \delta$, where δ is a user defined threshold for the desired norm of the gradient at the function minimum. The number of iterations are limited to $iter_{max}$, since convergence is not guaranteed. Denote the functions $W_g(x_k, p)$ and $W_H(x_k, p)$ which use a point in the scene scan and the transformation parameters to calculate an associated gradient and Hessian contribution for point x_k. Biber uses a Newton method with line search to perform the optimization, and full details of the optimization and the derivation of $W_g(x_k, p)$ and $W_H(x_k, p)$ can be found in [6]. The NDT method is summarized in Algorithm 1.

C. K-means Clustering

K-means clustering is a data partitioning technique which, given a set of n data points, $D = \{d_1, d_2, \ldots, d_n\}$, seeks to divide the data points D into k sets, $S = \{s_1, s_2, \ldots, s_k\}$.

Algorithm 1 Register scene scan X to reference scan Y using NDT

Initialization of NDT:
allocate NDT cells C, for reference scan Y
$c_i \leftarrow \{y_i : y_i \in c_i\}$
for all $c_i \in C$ **do**
  $\mu_{c_i} \leftarrow \frac{1}{N_{c_i}} \sum_{k=1}^{N_{c_i}} y_k^{c_i}$
  $\Sigma_{c_i} \leftarrow \frac{1}{N_{c_i}-1} \sum_{k=1}^{N_{c_i}} (y_k^{c_i} - \mu_{c_i})(y_k^{c_i} - \mu_{c_i})^T$
end for
Registration of scene scan:
$p \leftarrow p_0$
$iter_{cur} \leftarrow 0$
while $\mid g \mid > \delta$ and $iter_{cur} < iter_{max}$ **do**
  $s \leftarrow 0$
  $g \leftarrow 0$
  $H \leftarrow 0$
  for all $x_k \in X$ **do**
    $\bar{x}_k \leftarrow T(p, x_k)$
    $c_{cur} = \{c_i \in C : \bar{x}_k \in c_i\}$
    $s \leftarrow s + \rho_{c_{cur}}(\bar{x}_k)$
    $g \leftarrow g + W_g(\bar{x}_k, p)_{c_{cur}}$
    $H \leftarrow H + W_H(\bar{x}_k, p)_{c_{cur}}$
  end for
  $\Delta p = (H)^{-1}(-g)$
  $p \leftarrow p + \Delta p$
  $iter_{cur} \leftarrow iter_{cur} + 1$
end while

The associated clustering metric is the squared distance between the points in the data set and their associated closest cluster mean. K-means seeks to divide the data points into k sets such that the function $f_{km} = \sum_{i=1}^{k} \sum_{d_j \in s_i} \|x_j - \mu_i\|^2$, is minimized. The term μ_i is the mean of the points assigned to set s_i. The k-means algorithm is typically performed using an iterative method and has two main steps. Given a set of initial k mean locations, the first step assigns each of the data points in D to the cluster with the closest mean value. In the second step, the means are re-calculated as the mean value of the data points assigned to each cluster. These steps are repeated until the movement of the mean values between t and $t-1$ iterations, α, falls below a certain threshold, δ. The k-means algorithm for clustering is summarized in Algorithm 2.

III. Proposed Method: Multi-Scale K-Means NDT (MSKM-NDT)

In order to overcome the poor convergence basin of the NDT algorithm, a multi-scale registration approach is proposed. This approach uses k-means clustering to cluster the scan at different scales and perform the optimization at successively decreasing scales. The multi-scale optimization improves the convergence basin with a course-to-fine approach that avoids local minima. To accommodate these, the

Algorithm 2 Algorithm for k-means clustering

Randomly choose k mean locations:
$M_0 \leftarrow \{\mu_1^0, \mu_2^0, \ldots, \mu_k^0\}$
while $\alpha > \delta$ **do**
  Perform assignment step:
  for all $s_i \in S$ **do**
    $s_i \leftarrow \{x_d \in D : \|x_d - \mu_i\| \leq \|x_d - \mu_j\| \ \forall \ j \in \{1, \ldots, k\}$
  end for
  Perform mean update step:
  for all $i \in \{1, \ldots, k\}$ **do**
    $\mu_i^t \leftarrow \frac{1}{|s_i|} \sum_{x_j \in s_i} x_j$
  end for
  Check for convergence:
  $M_t \leftarrow \{\mu_1^t, \mu_2^t, \ldots, \mu_k^t\}$
  $\alpha \leftarrow \|M_{t-1} - M_t\|$
end while

standard NDT algorithm is modified in the following ways:

1) The grid based approach in NDT for the division of the scan is replaced with the k-means algorithm for clustering the scan into k clusters. Let the set of all clusters be $\Gamma = \{\gamma_1, \gamma_2, \ldots, \gamma_k\}$, where each cluster has N_{γ_j} points and $\sum_{i=1}^{k} N_{\gamma_j} = N_y$. Where standard NDT models the distribution of points within each cell as a Gaussian distribution, MSKM-NDT models each k-means cluster as a Gaussian distribution, ρ_{γ_j}.
2) The standard NDT algorithm performs the optimization step for one fixed grid spacing, while MSKM-NDT performs the optimization step at decreasing scales, increasing the number of clusters as the scale is decreased. The set of scales $\Phi = \{\phi_1, \phi_2, \ldots, \phi_N\}$ is selected in descending order. This results in a course-to-fine optimization scheme which aims to converge to the global minimum of the cost function.
3) To accommodate this style of clustering of the reference scan, the standard NDT cost calculation is modified. Standard NDT evaluates the Gaussian distribution within the cell that a transformed point x_d falls into. For MSKM-NDT, each transformed point, x_d, is evaluated with respect to all cluster distributions in the current scale, ϕ_i, with the strongest contributions coming from the most likely clusters. This results in a continuous and differentiable cost function at each stage of the multi-scale optimization.

An example of a reference scan clustered using k-means at two different scales is presented in Figure 3. The ellipses represent one standard deviation for the Gaussian distributions of the cluster points. The figures demonstrate that clustering at larger scales captures very coarse features within the scan and clustering at smaller scales captures the finer details. At larger scales, the cost function has a wider basin of convergence, however due to the coarseness of the clustering, the minimum at that scale may not correspond to

the true transformation parameters. At smaller scales, the cost function has a much narrower basin of convergence, however the global minimum is likely to correspond to the true transformation parameters. This idea is depicted in Figure 4. In this multi-scale approach, the optimization for a scale, ϕ_i, is initialized with the parameter solution from the previous scale, ϕ_{i-1}. The MSKM-NDT algorithm is summarized in Algorithm 3.

Algorithm 3 Register scene scan X to reference scan Y using MSKM-NDT

$p \leftarrow p_0$
for all $\phi_i \in \Phi$ **do**
  Perform k-means clustering of scan:
  $\Gamma \leftarrow \text{kmeans}(Y, \phi_i)$
  for all $\gamma_j \in \Gamma$ **do**
    $\mu_{\gamma_j} \leftarrow \frac{1}{N_{\gamma_j}} \sum_{k=1}^{N_{\gamma_j}} y_k^{\gamma_j}$
    $\Sigma_{\gamma_j} \leftarrow \frac{1}{N_{\gamma_j}-1} \sum_{k=1}^{N_{\gamma_j}} (y_k^{\gamma_j} - \mu_{\gamma_j})(y_k^{\gamma_j} - \mu_{\gamma_j})^T$
  end for
  Registration of scene scan:
  while $\mid g \mid > \delta$ **do**
    $s \leftarrow 0$
    $g \leftarrow 0$
    $H \leftarrow 0$
    for all $x_k \in X$ **do**
      $\bar{x}_k \leftarrow T(p, x_k)$
      for all $\gamma_j \in \Gamma$ **do**
        $s \leftarrow s + \rho_{\gamma_j}(\bar{x}_k)$
        $g \leftarrow g + W_g(\bar{x}_k, p)_{\gamma_j}$
        $H \leftarrow H + W_H(\bar{x}_k, p)_{\gamma_j}$
      end for
    end for
    $\Delta p = (H)^{-1}(-g)$
    $p \leftarrow p + \Delta p$
  end while
end for

IV. Experimental Results

In order to test the convergence basin of the MSKM-NDT approach, an experiment is conducted using scans of an indoor environment collected with a *Hokuyo UTM-30LX* 2D LIDAR. For the test, a pair of sequential scans are transformed relative to each other by a known amount and a registration algorithm is performed in order to determine the transformation parameters. The true transformation parameters are collected using a *NaturalPoint Optitrack* indoor positioning system, with an translational accuracy of approximately 1 mm and rotational accuracy of approximately 1 degree. If each of the the resulting transformation parameters are within 5% of the true transformation parameters, the registration is deemed successful. To test a wide case of scenarios, the scans are tested over an x-position range of -200 cm to 200 cm (with a 50 cm resolution), a y-position range of -200 cm to 200 cm (with a 50 cm resolution) and a rotation range of -30 degrees to 30 degrees (with a 15 degree resolution). This yields 405 combinations in total.

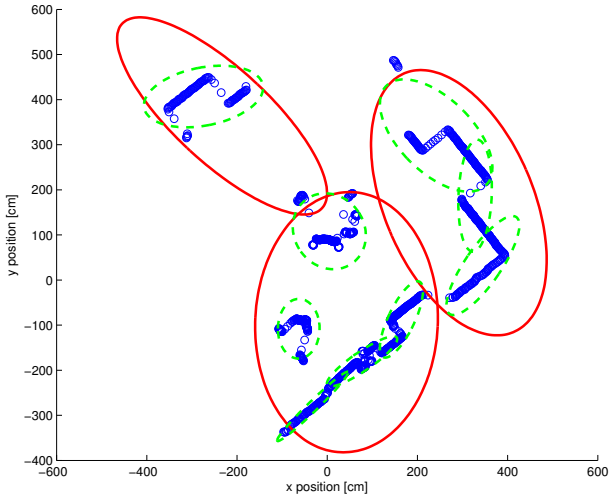


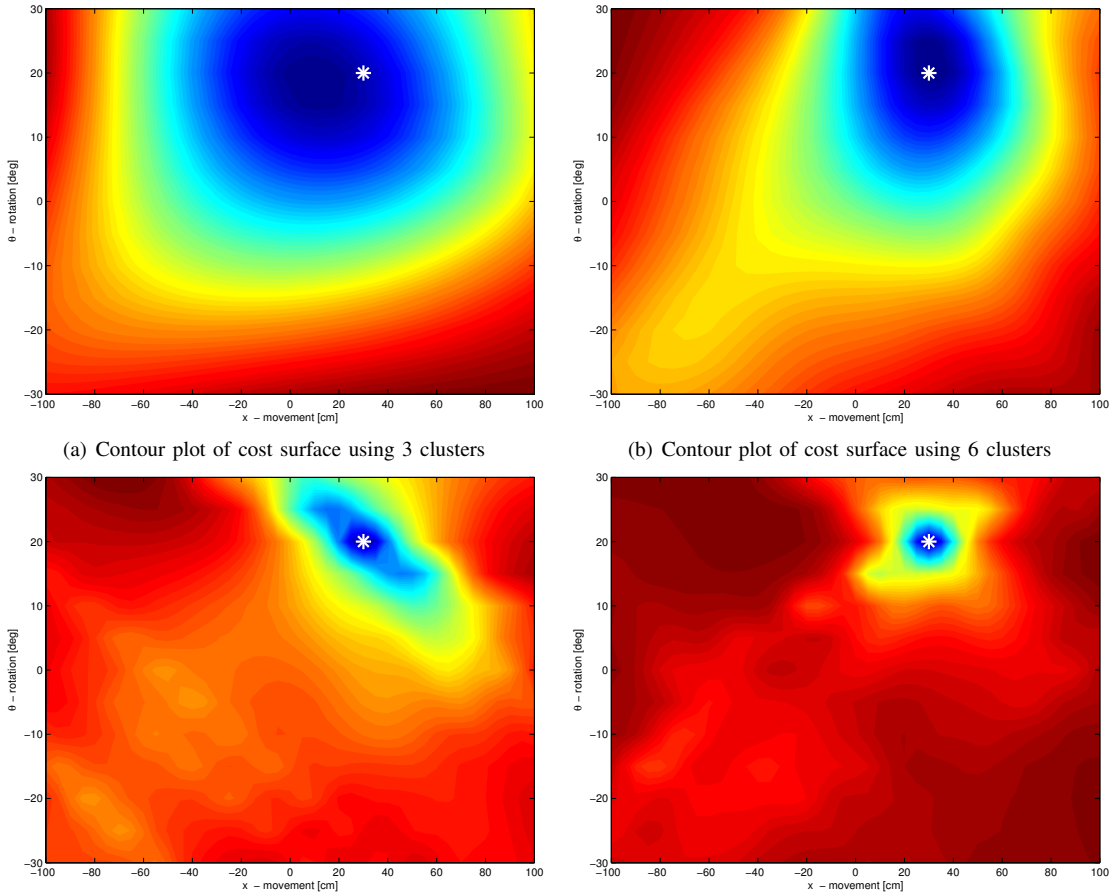
Fig. 3. Reference scan clustering at scale $\phi = 3$ (red ellipses) and at scale $\phi = 9$ (green dashed ellipses). The ellipses represent one sigma of the Gaussian clusters.

	NDT	MSG-NDT	MSKM-NDT
Full overlap	4.9%	22.2%	94.3%
Partial overlap	3.0%	20.0%	75.9%

TABLE I
Summary of success rates for registration tests

Three scan registration approaches are compared: the standard NDT algorithm (NDT) with a grid resolution of 50 cm, a multi-scale grid discretization NDT (MSG-NDT) and the proposed method, a multi-scale k-means clustering NDT (MSKM-NDT). The multi-scale grid discretization approach is an intuitive scale based optimization approach suggested by Magnusson et al. [7]. For this method, a large scale denotes large grid spacing, while a small scale denotes finer grid spacing. For the experiment a cluster scale setting of $\Phi = \{3, 6, 9, 15\}$ for the MSKM-NDT is selected and a grid scale setting of $\Phi = \{400, 200, 100, 50\}$ (distance between grid cells in [cm]) for MSG-NDT, is selected. For all algorithms, the *MATLAB* optimization toolbox is used to optimize the cost function. The experiment is performed for scans which partially and are fully overlapping. Fully overlapping scans are taken from the same vehicle pose, and partially overlapping scans are taken from different vehicle poses. The results of the convergence tests are shown in Figure 5. The blue dot indicates the x, y displacement between scans. The arrow shows the rotations which were tested at that x, y location. This describes a full parameter set for the transformation. A red arrow indicates the registration failed for that parameter set, while a green arrow indicates the registration succeeded.

Table I summarizes the results of the registration tests. The NDT algorithm has the smallest valley of convergence, as it is essentially only able to converge to correct solutions for small initialization errors for both the fully and partially overlapping case. The MSG-NDT algorithm has a wider valley



(a) Contour plot of cost surface using 3 clusters

(b) Contour plot of cost surface using 6 clusters

(c) Contour plot of cost surface using 9 clusters

(d) Contour plot of cost surface using 15 clusters

Fig. 4. Contour plots for cost surfaces with differing cluster amounts. This example shows the cost functions for two scans which have only been transformed in x and θ. The white marker indicates the true transformation parameters, $x = 30$ and $\theta = 20$. Using successive optimizations at decreasing scale, the overall optimization of the cost function avoids the local minima and converges at the global minimum

of convergence, but still generally fails for large translations in the x and y directions. The MSKM-NDT algorithm had the highest success rate and suggests the widest valley of convergence. Furthermore, there is no visible trend to when the MSKM-NDT algorithm fails. The failures in this case are most likely geometry specific to the environment, and could be rectified by performing the registration with additional scales. Computationally, a naive implementation of MSKM-NDT showed only a limited performance degradation. The multi-scale methods had comparable run times, while the standard NDT algorithm was approximately 4 times faster. Although the k-means clustering increases the computation time, less time is spent evaluating the cost function. This is because in general, there are fewer Gaussian clusters for scan point evaluation with MSKM-NDT, compared to an NDT implementation with a discrete, fixed grid size. A complete analysis of computational complexity is an area of future work.

V. Conclusion

In this work, a modified NDT algorithm known as the multi-scale k-means NDT (MSKM-NDT) is presented. In contrast to the standard NDT algorithm, the proposed method divides the scan into clusters at multiple scales using k-means clustering. The optimization step is then performed for multiple scales. The k-means clustering of the scan removes discontinuities of the cost function found in the standard NDT algorithm caused by the cell boundaries. The result is a smooth and differentiable cost function which guarantees convergence of the optimization at each scale. The coarse-to-fine approach has been experimentally shown to avoid the local minima associated with the standard NDT algorithm and improve the convergence basin. Future work for this area includes extending the results to 3D, optimizing the implementation in order to achieve real-time performance, comparison to other state-of-the-art methods, and evaluation of the algorithm on a larger class of environments and standard datasets.

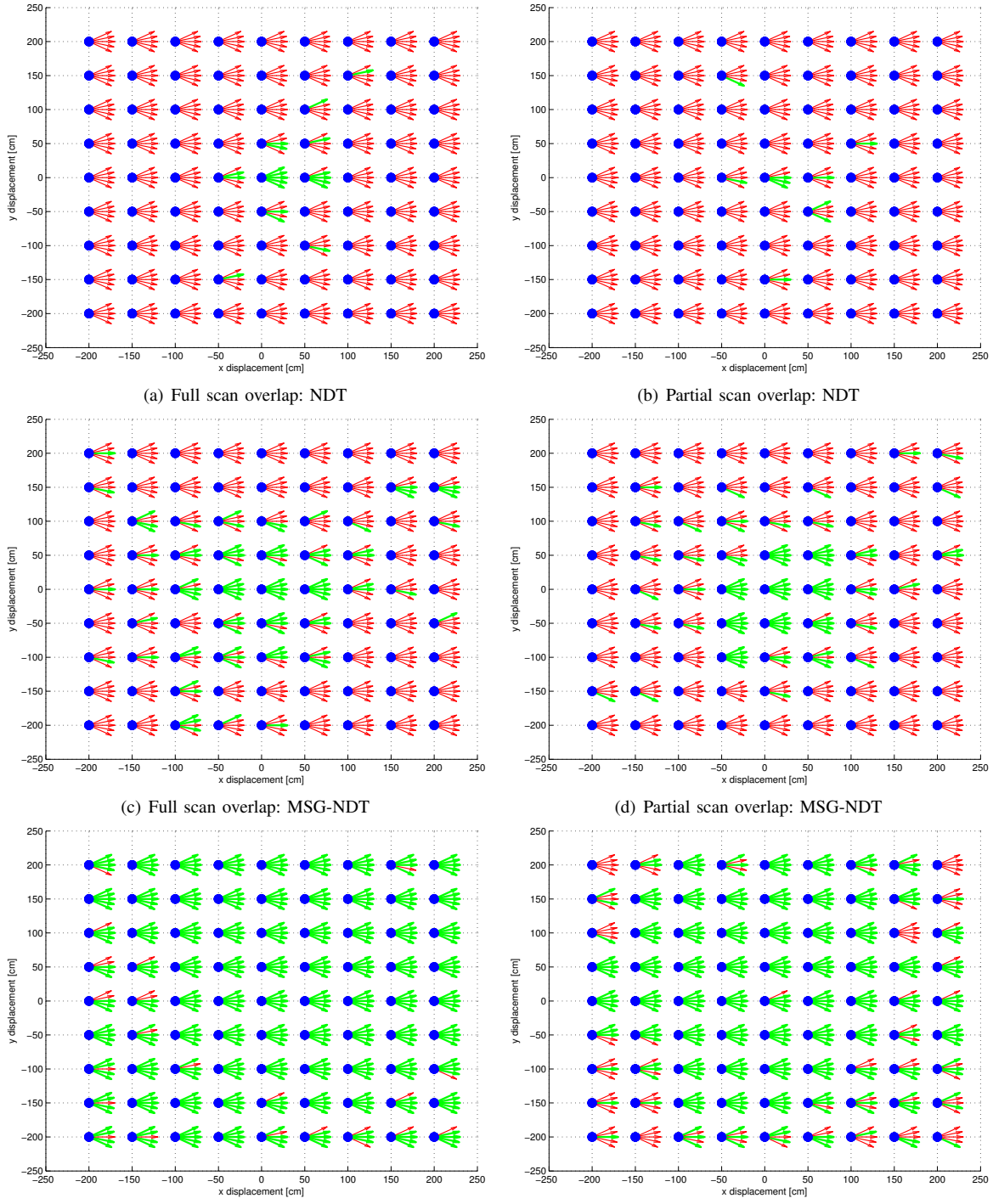
References

[1] P. Besl and H. McKay, "A method for registration of 3-d shapes," *IEEE Transactions onPattern Analysis and Machine Intelligence*, vol. 14, no. 2, pp. 239 –256, Feb 1992.

[2] Y. Chen and G. Medioni, "Object modeling by registration of multiple range images," in *Robotics and Automation, Proceedings., 1991 IEEE International Conference on*, Sacramento, CA , USA, Apr 1991, pp. 2724 –2729 vol.3.

[3] Z. Zhang, "Iterative point matching for registration of free-form curves and surfaces," INRIA-Sophia Antipolis, Tech. Rep., 1992.

[4] A. Segal, D. Haehnel, and S. Thrun, "Generalized-ICP," in *Proceedings of Robotics: Science and Systems*, Seattle, USA, June 2009.



(a) Full scan overlap: NDT

(b) Partial scan overlap: NDT

(c) Full scan overlap: MSG-NDT

(d) Partial scan overlap: MSG-NDT

(e) Full scan overlap: MSKM-NDT

(f) Partial scan overlap: MSKM-NDT

Fig. 5. Plots illustrating the registration results for the NDT, MSG-NDT and MSKM-NDT algorithms using a pair of scans from various initial transformations (green - thick arrows, red - thin arrows).

[5] D. Borrmann, J. Elseberg, K. Lingemann, A. Nüchter, and J. Hertzberg, "Globally consistent 3d mapping with scan matching," *Robotics and Autonomous Systems*, vol. 56, no. 2, pp. 130–142, 2008.

[6] P. Biber and W. Strasser, "The normal distributions transform: a new approach to laser scan matching," in *Intelligent Robots and Systems, Proceedings. 2003 IEEE/RSJ International Conference on*, vol. 3, Las Vegas, NV, USA, oct. 2003, pp. 2743 – 2748 vol.3.

[7] M. Magnusson, A. Lilienthal, and T. Duckett, "Scan registration for autonomous mining vehicles using 3D-NDT," *Journal of Field Robotics*, vol. 24, no. 10, pp. 803–827, 2007.

[8] M. Magnusson, A. Nuchter, C. Lorken, A. Lilienthal, and J. Hertzberg, "Evaluation of 3D registration reliability and speed-a comparison of ICP and NDT," in *IEEE International Conference on Robotics and Automation*, Kobe, Japan, 2009, pp. 3907–3912.

[9] D. Bertsekas, *Nonlinear Programming*. Athena Scientific, 1999.

[10] M. Magnusson, "The three-dimensional normal-distributions transform — an efficient representation for registration, surface analysis, and loop detection," Ph.D. dissertation, Örebro University, Dec. 2009, Örebro Studies in Technology 36.