2014 IEEE International Conference on Robotics & Automation (ICRA)
Hong Kong Convention and Exhibition Center
May 31 - June 7, 2014. Hong Kong, China

Multi Channel Generalized-ICP

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

Abstract— **Current state of the art scan registration algorithms which use only position information often fall victim to correspondence ambiguity and degeneracy in the optimization solutions. Other methods which use additional channels, such as color or intensity, often use only a small fraction of the available information and ignore the underlying structural information of the added channels. The proposed method incorporates the additional channels directly into the scan registration formulation to provide information within the plane of the surface. This is achieved by calculating the uncertainty both along and perpendicular to the local surface at each point and calculating nearest neighbour correspondences in the higher dimensional space. The proposed method reduces instances of degenerate transformation estimates and improves both registration accuracy and convergence rate. The method is tested on the Ford Vision and Lidar dataset using both color and intensity channels as well as on Microsoft Kinect data obtained from the University of Waterloo campus.**

I. Introduction

Simultaneous localization and mapping (SLAM) is one of the keystone components of autonomous systems operating in an unknown environment. With many modern sensors, such as LIDAR, RGBD cameras, and stereo cameras, providing robots with reliable 3D point information of their environment, scan registration techniques have become a prevailing solution to the SLAM problem. Scan registration aligns consecutive scans to obtain the rotation and translation of the system relative to its environment and allows for the aggregation of point cloud data. These aggregate maps provide detailed environmental information which can be used for path planning and obstacle avoidance.

One of the first methods proposed for solving the scan registration problem was the Iterative Closest Point (ICP) method introduced by Besl and McKay [1]. The ICP algorithm minimizes the Euclidean distance between nearest neighbour points in the two scans to find the relative transform. Taking advantage of the locally planar nature of most environments, Chen and Medioni [2], proposed a point to plane based variant of ICP which penalizes the cost only normal to the surface of the environment. This approach mitigates the sampling error seen in point to point ICP which assumes that points correspond exactly between scans. Recently, Segal et al. developed Generalized-ICP (GICP) [3] which, using a probabilistic framework, generalized the ICP method and introduced a plane to plane approach with improved performance over the previous versions.

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

An alternative approach, called the normal distribution transform (NDT), was first suggested by Biber and Strasser in [4] in 2D and extended to 3D by Magnusson et al. [5]. NDT segments the scan into fixed size voxels and calculates a normal distribution of the points within each cell. Scans are then registered by using the point to distribution or distribution to distribution [6] error metric.

The classic scan registration formulations use only the 3D point information to calculate point correspondences, distributions, and to perform the registration. However, many sensors, or combinations of sensors, can provide additional information for each point such as intensity or color. The additional information can help to improve registration accuracy, convergence rate, and solve many structural ambiguities.

With laser scanners becoming common place several authors have proposed not only using the 3D points returned by the scanner but also the intensity values which many scanners also produce. Levinson and Thrun [7] use 2D probabilistic intensity maps and a histogram filter to localize in an urban environment. This method, although successful in certain environments, completely ignores geometric information and assumes sufficient intensity information is always available.

Color information has also become commonly applied to 3D scans. Cameras and depth sensors can be extrinsic calibrated such that points can be associated with corresponding image pixels and colorized. This sensor combination has been used by many authors to attempt to improve the performance various SLAM algorithms.

Color ICP [8] attempts to improve ICP performance by using the additional color channels to perform the nearest neighbour search in a higher dimension. This method showed improved point correspondence results but did not change the underlying scan registration method. A colorized version of NDT has also been proposed by Huhle et al. [9], which uses color based kernel functions to generate a Gaussian mixture model such that each voxel then contains a color based mixture of Gaussians.

Another class of methods which incorporate color into scan matching use image features to find corresponding points and limit the size of the problem. The most well known method of this class is known as RGBD-SLAM and was developed by Endres et al. [10]. RGBD-SLAM uses SIFT or SURF features to match pairs of images and uses RANSAC to estimate the 3D transformation. Other methods include, [11], which initialize GICP using image features and [12] which augments NDT using a small set of image feature correspondences as a secondary error function.

978-1-4799-3684-7/14/$31.00 ©2014 IEEE

Feature based methods can have several draw backs as demonstrated in [9]. Since feature based methods rely on matching a small number of points, noise in the 3D location can cause significant errors. Additionally, a small number or even a single false correspondence can cause catastrophic failure. This can make feature based methods unreliable particularly in dynamic environments where a single feature could move and distort the entire map.

This work proposes the Multi Channel Generalized-ICP method. The Multi Channel GICP method is an extension of the GICP method which incorporates additional channels of information for each point. The additional channels are used to generate a spacial covariance in the plane of the surface which is used to compliment the existing plane to plane matching and allow the planes to be aligned not only in the normal direction, but also perpendicular to it. The in-plane covariance is calculated using a kernel weighted covariance based on the additional channels and is normalized by the unweighed population covariance of the points. The in-plane covariance is added to the GICP archetypical covariance in the planar directions and is then rotated into 3D space as done in GICP. Point correspondences are also changed to use the Color ICP method, which leverages a higher dimensional weighted kD tree. The error function is unchanged from the original GICP, however the modified covariance will induce non-trivial error terms in the planar directions.

The proposed method was evaluated using the Ford Campus Vision and Lidar Dataset [13], as well as Microsoft Kinect data taken on the University of Waterloo campus.

II. GICP

The Generalized Iterative Closest Point (GICP) method was developed by Segal et al. in [3] as a unifying framework of the previously proposed ICP methods. The GICP formulation uses a probabilistic framework to determine the error function and proceeds as follows.

First, it is assumed that the nearest neighbour correspondences have been calculated and scan $A = \{a_i\}$, where $a_i \in \mathbb{R}^3$ for $i \in \{1, ..., N\}$, and scan $B = \{b_j\}$, where $b_j \in \mathbb{R}^3$ for $j \in \{1, ..., N\}$, are indexed with corresponding points having the same indices and non corresponding points being removed. Using the probabilistic model it is assumed that the point clouds A and B are generated from an underlying set of distributions, where $a_i \sim \mathcal{N}(\hat{a_i}, C_i^A)$ and $b_i \sim \mathcal{N}(\hat{b_i}, C_i^B)$. Therefore given perfect correspondences and the correct transform, T^*,

$$\hat{b_i} = T^* \hat{a_i} \tag{1}$$

The difference between samples a_i and b_i is then defined as $d_i = b_i - Ta_i$. Given that a_i and b_i are drawn from independent Gaussian distributions, and given the correct transformation, d_i can be written as:

$$d_i \sim \mathcal{N}(0, C_i^B + T^* C_i^A (T^*)^T) \tag{2}$$

The transform is then solved for using maximum likelihood estimation (MLE) and simplified to the form

$$T = \arg\min_T \sum_i d_i (C_i^B + T C_i^A T^T) d_i^T \tag{3}$$

This formulation can be used to represent any of the standard forms of ICP including basic point to point as well as point to plane ICP. However GICP proposes a plane to plane model in which it is assumed that points are sampled from surfaces which are locally planar. In this model the covariance of a point is assumed to be small in the direction of the normal at that point and large in all other directions. This assumes that the points have little information to offer in the directions tangent to the plane. The covariance at every point, $w_i \in A \cup B$, in both A and B is calculated using an archetype covariance, C^G, define as

$$C^G = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \epsilon \end{bmatrix}$$

where ϵ is a constant representing the covariance along the normal. The covariance at a point is then calculated as

$$C_i^W = (R_i^W) C^G (R_i^W)^T \tag{4}$$

where R^W is the rotation matrix which rotates ϵ to align with the surface normal, at point w_i.

The local covariance, C_i^L, is calculated using the k nearest points to the query point w_i found using [14]. The local covariance approximates the model covariance in the region around the query point.The surface normal information for this method is then computed using principal component analysis (PCA) on the local covariance, C_i^L. The component with the lowest eigenvalue corresponds to the surface normal. In practice the model covariance at a given point can be calculated using the singular value decomposition (SVD) of the local covariance,

$$C^L = USV^T \tag{5}$$

where the singular values are the diagonal elements of $S \in R^{n \times n}$ sorted in descending order and U and V are orthonormal matrices. In the singular value decomposition decomposition U is equivalent to the rotation matrix R_i^W and therefore S can be replaced by C^G, to compute, C_i^W.

III. Multi Channel GICP

The proposed method, Multi Channel GICP, is an extension of the GICP algorithm which incorporates additional channels of information. The Multi Channel GICP algorithm uses additional channels such as color, intensity, or any other spectral information, to introduce additional information to the problem. Additionally, Multi Channel GICP uses the added channels directly in the correspondence search to attempt to provide more robust results. The increased problem space not only solves the degeneracy problem but also improves accuracy, convergence, and robustness of the scan registration results.

Multi Channel GICP assumes, as GICP did previously, that the environment is locally planar and that the 3D points

only contain useful information in the direction normal to the surface. However, since the points have at least one additional channel of information the additional channel(s) can be used to define the covariance of a point along the surface plane as well. The added channels will have no effect normal to the plane as the sample must lie on the surface and therefore will complement the positional information well.

First, let all points, $p_i = \{p_i^p, p_i^d\}^T$, have both positional information, $p_i^p \in \mathbb{R}^3$, and n descriptor channels, $p_i^d \in \mathbb{R}^n$, which can include, for example, intensity or RGB colour information. Then for each point in A, and B, the model covariance sets, C^A and C^B, are calculated using both position and descriptor information.

Let $q \in A \cup B$ be the current query point for which the model covariance is to be calculated. The local covariance, C_i^L, of the query point position is calculated using the k nearest neighbour points to q^p using [14]. Let the nearest neighbours be defined as the set of points $L = \{l_j\}$ for $j \in \{1, ..., k\}$, such that $\|l_j^p - q^p\| \leq \|r - q^p\|$, for all $r \in Q \cap \bar{L}$, where Q is the point cloud associated with the current query point, q. Singular value decomposition (SVD) is performed to extract the principal components. The normal of the surface is found as the component with the smallest singular value in S. The neighbourhood points are then projected onto the plane perpendicular to the normal and reduced to $\mathbb{R}^2$. Let $z_j^p \in \mathbb{R}^2$ be the projected point and $Z = \{z_j^p, z_j^d\}^T$ be the new set of points in $\mathbb{R}^{2+n}$. The projection is then given as

$$\begin{aligned} z_j^p &= \begin{bmatrix} U_1{}^T \\ U_2{}^T \end{bmatrix} l_j^p \\ z_j^d &= l_j^d \end{aligned} \tag{6}$$

where U_1 and U_2 are the first and second columns of the SVD matrix U. Note that after the transform, the new population covariance, $\Sigma_w \in \mathbb{R}^{2\times2}$, of the points, Z^p, is the diagonal matrix of the largest two singular values of S.

Now that the 3D points have been projected onto the local surface approximation, the reduction in uncertainty due to the incorporation of descriptor information in the plane can be calculated. To this end a descriptor kernel weighted covariance is calculated using weightings based on a probabilistic model similar to that used in [9]. The descriptor kernel calculates the probability that an arbitrary point corresponds to the query point in descriptor space. The kernel can be defined as a Gaussian distribution, $\mathcal{N}(q^d, \Lambda)$, centered at the query point descriptor, q^d, and with $\Lambda \in \mathbb{R}^{n\times n}$ being the measurement covariance of the descriptor sensor. The kernel weights are then calculated for each point in Z as:

$$\lambda_j = \exp(-\frac{1}{2}(z_j^t - q^t)^T \Lambda^{-1} (z_j^t - q^t)) \tag{7}$$

Using the kernel weights the descriptor kernel weighted covariance and mean, Σ_t and μ^p, can be calculated as

$$\mu^p = \frac{1}{\sum_j \lambda} \sum_j \lambda_j z_j^p \tag{8}$$

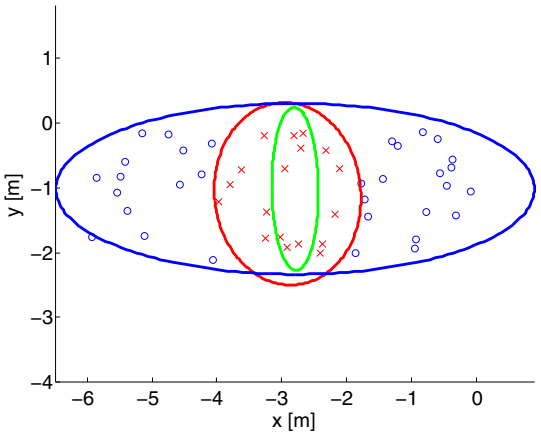


Fig. 1. An example of the population (blue), and descriptor (red), covariances given a biased initial population and the resulting normalized correlation (green).

$$\Sigma_t = \frac{1}{\sum_j \lambda} \sum_j \lambda_j (z_j^p - \mu^p)(z_j^p - \mu^p)^T \tag{9}$$

This gives the spacial distribution of points based on their similarity to the query point. This distribution models the uncertainty of the descriptor information along the wall locally, however it can be biased if the original sample population was itself already biased. Figure 1 shows an example of a population and descriptor covariance with a biased initial population. To compensate for this potential bias the distribution is normalized by the population covariance such that

$$\Omega = \Sigma_w^{-\frac{1}{2}} \Sigma_t \Sigma_w^{-\frac{1}{2}} \tag{10}$$

The correlation coefficient matrix, $\Omega \in \mathbb{R}^{2\times2}$, shows the correlation of the descriptor weighted data compared to that of the population. A value less than one indicates that the descriptor data increased the data certainty in that direction, equal to one shows no change, while a value less than one indicates an increase in uncertainty. Directions which have a low correlation coefficient are more likely to capture correct correspondences in descriptor space in that direction. Cases which have correlation coefficients equal to or great than one indicate areas of low descriptor correspondence certainty, such as a wall of a single continuous color.

To use this information in the GICP framework, Ω is used along the planar directions. Therefore the resulting covariance used in the MC-GICP algorithms is

$$C^D = \begin{bmatrix} \Omega & 0 \\ 0 & \epsilon \end{bmatrix}$$

In addition to the covariance changes the calculation of corresponding points is also changed to reflect the higher dimensionality of the information. An $n + 3$ dimensional weighted kD tree is used to incorporate all of the information into the search as first shown by Johnson and Kang [8]. A weighting vector, $\alpha = \{\alpha_1, ..., \alpha_{n+3}\}$ where $\alpha_1 = \alpha_2 = \alpha_3 = 1$ are the weights of the position data, is used to weight the descriptor information relative to the positional information.

The error function and minimization of the Multi Channel GICP method remain unchanged from the original GICP, presented in Equation (3), which means that all current methods for solving the GICP optimization are still valid and no changes are necessary to the optimizer.

IV. Implementation

The proposed method is a generalized framework which can be used with a variety of different sensor combinations. The number of additional channels which can be added to the points is not limited by the algorithm but in practice is only limited by diminishing returns on the usefulness of the information. Three possible sensor configurations are discussed below. First, a system which uses only a LIDAR sensor with intensity information, second, a typical colourized point cloud generated from a camera and range sensor combination, and finally a configuration incorporating a LIDAR sensor with intensity information with a camera setup to provide four additional information channels.

A. Laser Intensity Descriptor

A single LIDAR, such as the Velodyne HDL-64E, can add an additional channel in the form of laser intensity, which provides useful information to the registration. In the single channel case, the descriptor covariance reduces to a single value, and is found by computing the variance, $\Lambda \in R$, of the laser scan intensity values. Finally the weighting, α_4, of the intensity channel in the nearest neighbour search to scale the influence the intensity channel will have on finding nearest neighbours.

B. Color Descriptor

The combination of a camera and a range sensor is a common setup on many robotic systems. Color provides three channels to incorporate into the model. In this case $\Lambda \in \mathbb{R}^{3\times3}$ is a covariance matrix of three variables. However, it can be assumed that the color channels are independent and therefore Λ will be a diagonal matrix consisting only of the intra-channel variances. The exact values of Λ will depend not only on the sensor used but also on the color space which is chosen.

In this paper the RGB space is used. In the RGB space, the color variances and weightings can be set equal to each other to represent equal uncertainty in each of the colors. The option of alternative color spaces is also possible as each space provides different benefits and considerations. For an HSV or YUV space, the variance of the value and illumination channels can be larger and the nearest neighbour weighting smaller to distinguish a higher uncertainty in illumination which is common in real world scenes. The choice of color space does not have a direct impact on the algorithm but does change the desired values of Λ and $\{\alpha_4, \alpha_5, \alpha_6\}$.

C. Combined Color and Intensity

The combination of both color and laser intensity information presents and interesting configuration which is not typically leveraged in current algorithms. Although the channels of the combined descriptor could be considered to be independent it has been shown in [15] that laser intensity and color intensity are in fact positively correlated. This can be incorporated into the algorithm by setting the inter-channel covariance terms of Λ to non zero values. The covariance matrix can be determined experimentally using a set of known training data. The weighting values for the nearest neighbour search are typically inversely proportional to the variance of that particular channel and are therefore dependent on the specific sensors being used.

V. Experimental Results

The Multi Channel GICP method is evaluated using two sets of data. The first set is the Ford Campus Vision and Lidar Dataset [13]. The Ford dataset contains LIDAR and omnidirectional image data as well as ground truth and is used to evaluate the quantitative accuracy of the scan registration results against the provided ground truth. The method is evaluated using the laser intensity, color and combined descriptors as described in Section IV. The second set of data was obtained using a Microsoft Kinect sensor on the University of Waterloo campus. This dataset is used to evaluate the method on data with limited geometric structure. The method is evaluated qualitatively based on the reconstruction of a flat textured surface. Finally the convergence rate of each algorithm is compared. In all cases the method is compared to both the original GICP and Color ICP algorithm as implemented in the Point Cloud Library (PCL) [16] and the RGB color space is used.

A. Ford Dataset Absolute Error

The Ford Campus Vision and Lidar Dataset was generated using a Ford F-250 pickup truck equipped with a Velodyne HDL-64E laser scanner, a Point Grey Ladybug3 omnidirectional camera, and a Applanix POS-LV 420 INS with Trimble GPS used for ground truth data. A series of 50 frames from the dataset were used to perform pairwise registration on each consecutive pair of scans. The resulting transforms were then compared to the ground truth data to calculate the mean and standard deviation of the translational and rotational errors. The results are summarized in Table I.

TABLE I
Summary of translation and rotation errors on the Ford dataset

	Color ICP	GICP	MC-GICP Intensity	MC-GICP Color	MC-GICP Combined
mean error [m]	0.5106	0.3082	0.3857	0.2941	0.2891
std. dev.	0.3280	0.1446	0.1756	0.1409	0.1387

The error distributions from Table I show that Multi Channel GICP using either color or combined descriptors has increased accuracy and reduced uncertainty over that of GICP and Color ICP. Of the Multi Channel GICP varients, the combined descriptor produced the best results followed by the color descriptor while intensity alone produce poor

results. This is expected as the combined descriptor provides the most robust information while the intensity alone has minimal distinctive variation in value. The combined descriptor increases the information in each point while maintaining consistancy. This results in increased correlation and correspondence certanty. Intensity pro This shows that the Multi Channel GICP method is dependent on the use of an accurate and distinct descriptor space.

The aggregated maps generated by Color ICP, GICP and, Multi Channel GICP using the combined descriptor are shown in Figure 2. In the aggregate maps, it can be seen that Multi Channel GICP creates more accurate results. This is evident by the blurring which can be seen in the Color ICP and GICP maps but is reduced in the Multi Channel GICP results.

B. Kinect Sparse Geometry Data

The Kinect dataset was obtained using the Microsoft Kinect RGB-D sensor mounted to a mobile robotics test platform.

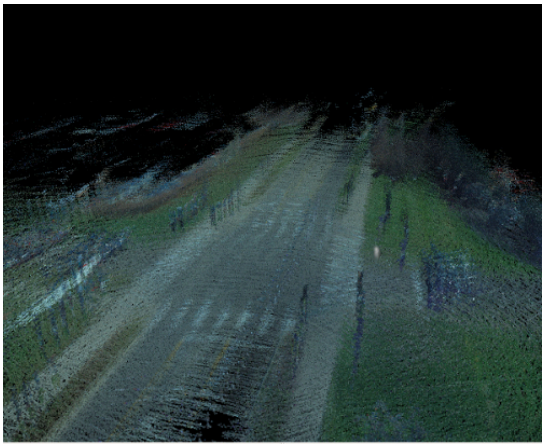
The sequence is of a flat wall which is covered in posters. This sequence contains no distinct geometric surfaces other than the flat wall and therefore produces degenerate solutions in the x and y directions when only geometry is considered. The sequence consists of 200 frames traversing from right to left over a $10m^2$ area. The aggregated results of the pairwise registration of the three methods are presented in Figure 3.

Figure 3 demonstrates the ability of the Multi Channel GICP algorithm to use the additional channels to compensate for the lack of geometric information. Color ICP is also able to partially compensate but has much lower accuracy observed by the increased blurring of the posters. Color ICP has trouble correctly aligning the scans due to noise in the images causing incorrect correspondences and skewing the results. The GICP results fail catastrophically in this case due to the lack of geometric information along the wall and therefore all the scans are incorrectly aligned.

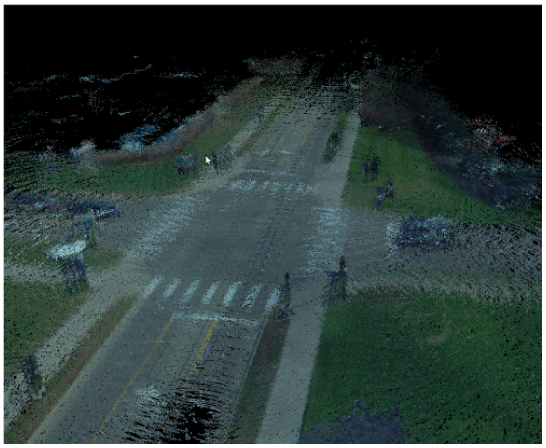
C. Convergence Rate

The final evaluation compares the convergence rates of the three algorithms. Given an example frame from the Ford Data Set, the error residual is plotted versus the iteration in Figure 4. As all three algorithms use computationally similar cost functions, iterations are proportional to convergence time.

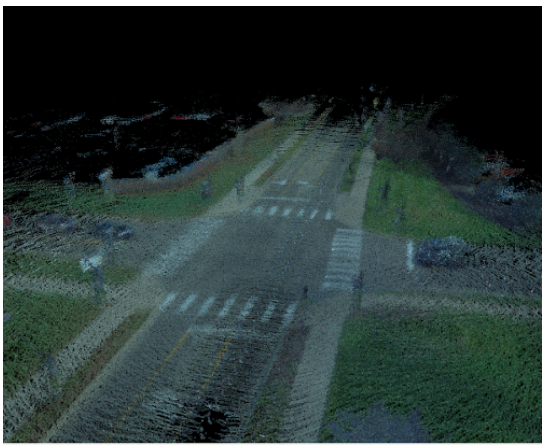
Figure 4 clearly shows that MC-GICP converges significantly faster than the original GICP algorithm. Color ICP has been shown to converge faster than standard ICP due to the fact that is acquires the correct correspondences faster using the higher dimensional search space. However, GICP, is shown to converge faster than Color ICP. This is because GICP does not rely as heavily on correct point correspondences to converge and only needs corresponding points to lie on the same surface. MC-GICP combines the beneficial properties of both to acquire the correct surface correspondences and converges most rapidly.



(a) Color ICP



(b) GICP



(c) MC-GICP Combined

Fig. 2. Aggregated point cloud maps generated from a subsection the Ford Vision and Lidar dataset.

VI. Conclusion

This work presents the Multi Channel Generalized-ICP method for robust scan matching. The proposed method incorporates the additional sensor channels directly into the GICP formulation to provide additional information in the plane parallel to the local surface. The GICP algorithm relies solely on surface normal information at each point,

(a) GICP

(b) Color ICP

(c) MC-GICP

Fig. 3. Aggregated point cloud maps generated from the poster wall dataset using the Kinect sensor.

and requires surface normals from the point set to span all of $\mathbb{R}^3$ to properly determine the transformation and avoid degeneracy of the solution. However, many real world environments do not provide sufficient information using surface normals alone, such as hallways, or flat open spaces. The Multi Channel GICP method modifies the model covariance planar to the surface normal and calculates nearest neighbour correspondences in a higher dimensional space, thereby exploiting the additional information available in the point cloud from secondary sensor channels to avoid this shortcoming. The proposed method demonstrates improved registration accuracy and convergence rate as well as robustness to degenerate geometric cases. Future work includes incorporating the proposed method with a graph-SLAM backend for loop closure and exploring other possible descriptor space combinations.

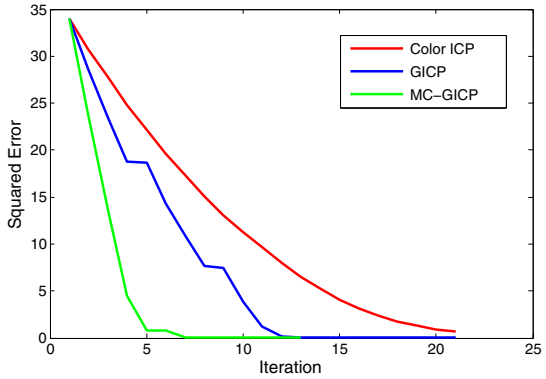


Fig. 4. Comparison of convergence of the Color ICP (red), GICP(blue) and MC-GICP (green) algorithms.

REFERENCES

[1] P. Besl and H. McKay, "A method for registration of 3-D shapes," *IEEE Transactions on Pattern 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 *International Conference on Robotics and Automation (ICRA)*, vol. 3. IEEE, Apr 1991, pp. 2724–2729.

[3] A. Segal, D. Haehnel, and S. Thrun, "Generalized-ICP," in *2009 Robotics: Science and Systems (RSS)*, June 2009.

[4] P. Biber and W. Strasser, "The normal distributions transform: a new approach to laser scan matching," in *International Conference on Robotics and Automation (ICRA)*, vol. 3. IEEE, Oct. 2003, pp. 2743–2748.

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

[6] T. Stoyanov, M. Magnusson, and A. Lilienthal, "Point set registration through minimization of the L2 distance between 3D-NDT models," in *International Conference on Robotics and Automation (ICRA)*. St. Paul, MN, USA: IEEE, May 2012, pp. 5196–5201.

[7] J. Levinson and S. Thrun, "Robust vehicle localization in urban environments using probabilistic maps," in *International Conference on Robotics and Automation (ICRA)*. IEEE, 2010, pp. 4372–4378.

[8] A. E. Johnson and S. Bing Kang, "Registration and integration of textured 3D data," *Image and vision computing*, vol. 17, no. 2, pp. 135–147, 1999.

[9] B. Huhle, M. Magnusson, W. Straßer, and A. J. Lilienthal, "Registration of colored 3D point clouds with a kernel-based extension to the normal distributions transform," in *International Conference on Robotics and Automation*. IEEE, 2008, pp. 4025–4030.

[10] F. Endres, J. Hess, N. Engelhard, J. Sturm, D. Cremers, and W. Burgard, "An evaluation of the RGB-D slam system," in *Robotics and Automation (ICRA), 2012 IEEE International Conference on*. IEEE, 2012, pp. 1691–1696.

[11] G. Pandey, J. McBride, S. Savarese, and R. M. Eustice, "Visually bootstrapped generalized ICP," in *International Conference on Robotics and Automation (ICRA)*. IEEE, 2011, pp. 2660–2667.

[12] B. Huhle, P. Jenke, and W. Straßer, "On-the-fly scene acquisition with a handy multi-sensor system," *International Journal of Intelligent Systems Technologies and Applications*, vol. 5, no. 3, pp. 255–263, 2008.

[13] G. Pandey, J. R. McBride, and R. M. Eustice, "Ford campus vision and lidar data set," *International Journal of Robotics Research*, vol. 30, no. 13, pp. 1543–1552, November 2011.

[14] M. Muja and D. G. Lowe, "Fast approximate nearest neighbors with automatic algorithm configuration," in *International Conference on Computer Vision Theory and Application (VISSAPP)*. INSTICC Press, 2009, pp. 331–340.

[15] G. Pandey, J. R. McBride, S. Savarese, and R. Eustice, "Automatic targetless extrinsic calibration of a 3d lidar and camera by maximizing mutual information." in *AAAI*, 2012.

[16] R. B. Rusu and S. Cousins, "3D is here: Point Cloud Library (PCL)," in *IEEE International Conference on Robotics and Automation (ICRA)*, Shanghai, China, May 9-13 2011.