Abstract—This paper describes an improved Hue-Assisted Normal Distributions Transform (HANDT), which enhances the accuracy of the registration of point clouds. In our previous work, HANDT was developed to improve the speed and accuracy of registration by utilizing the hue from the hue-saturation-value model. To improve the accuracy, the score function of HANDT is changed from the normal distribution function to the Mahalanobis distance function. In addition, the functions of the hue mean and variance are modified for the circular property of the hue. The performance of normal distributions transform (NDT), HANDT, and improved HANDT are evaluated by benchmark data sets. As a result, the translation and rotation errors of improved HANDT are lower than those of NDT and HANDT.

Index Terms—Hue, color, point cloud, registration, normal distributions transform (NDT).

I. INTRODUCTION
A map built with a range sensor generally provides the structure of the environment. It is not easy to recognize what objects are located on the map. However, if the map is colored, it is possible to recognize objects and textures. Usually, a range sensor and a vision sensor are mounted together on robots. Due to the reliable calibration of the sensors, points scanned by the range sensor are able to be paired accurately with colors scanned by the vision sensor. In addition, low-priced depth cameras such as Microsoft Kinect and Asus Xtion were launched for the past years. It stimulated researchers to study the colored point cloud registration.

The registration of images or scans is an essential process of 3D graph Simultaneous Localization and Mapping (SLAM). Due to the good extensibility of Iterative Closest Point (ICP) algorithm, most of color-supported registration algorithms are based on ICP. The early algorithm is Colored ICP (cICP) [1]. Y, I, and Q from the YIQ color model were integrated to the distance function [2]. Similarly, 4D ICP integrated the hue variable from the hue-saturation-lightness (HSL) model to the distance function [2]. Color-constrained ICP uses six classes corresponding to six intervals of the hue range. Points would be classified to classes according to hue. The points in the highest-scored class would be registered by ICP [3]. The modified Colored ICP quantizes I and Q to classify reference points. The corresponding cell of a new point on the IQ plane is decided by I and Q of the new point, and it searches for the closest point in the cell [4]. Color-supported generalized ICP is a color-supported variant of the generalized ICP (gICP), which is a variant of ICP [5]. L*, a*, b* from the L*a*b* color model are integrated to the distance function of gICP.

NDT is a competitive algorithm with ICP. NDT is known for the fast registration of large point clouds. It is also known to be robust against outliers and noise. Moreover, since hundreds of figures in a voxel are reduced to twelve figures, it is much more efficient in saving data. Hence, NDT is the most suitable algorithm for the registration of a large number of points. There is a color-supported variant called Color NDT. Color NDT is improved by applying color weights to the score function [6]. In spite of the increasing popularity of NDT, no more color-supported variants of NDT have appeared so far.

In our previous work, Hue-assisted NDT (HANDT) was proposed [7]. We utilized the hue from the hue-saturation-value (HSV) model because of its light-invariant and viewpoint-invariant properties. After the reference point cloud is divided by the octree structure, the points in each unit voxel are classified according to their hue values. In addition, the score function of NDT was modified to be weighted by hue coefficients.

We found that errors occurred because of the simple mean and variance of the hue variable. In this paper, the circular mean and variance functions of the hue variable are described. In addition, the form of the score function is changed from the normal distribution function to the Mahalanobis distance function.

II. BACKGROUND
A. The Hue Variable from the HSV Model
We get color information by a vision sensor. The RGB values of an object are varying with the viewpoint of the sensor and the luminous intensity. On the contrary, the hue value from the HSV model is invariant to those effects. Two photographs are taken at different brightness levels as shown in Fig. 1. The distributions of RGB of photographs are as shown in Fig. 2. The distributions in Fig. 2(a) are widely distributed while those of Fig. 2(b) are concentrated at the left side. The hue distributions as shown in Fig. 3 are very close. Averages of four peaks in Fig. 3(a) are 32.58, 114.80, 114.80, 114.80, 114.80, 114.80, 149.30, and 240.26. Due to the brightness-invariant property of the hue, the hue is possible to support the registration of two colored point clouds scanned at different brightness level.
where \( \Theta \) is a vector transformed from the rotational matrix \( R \) and translational vector \( T \).

At the second step, the transformation vector \( \xi = \left[ \alpha \beta \gamma \tau_x \tau_y \tau_z \right]^T \) is updated iteratively to get the best for \( \xi^* \) which maximizes a score function as

\[
S_{sp} = \sum_{j=1}^{N} \exp \left( -\frac{1}{2} (p'_{new, j} - M_j) \Sigma_j^{-1} (p'_{new, j} - M_j) \right)
\]

where \( S_{sp} \) is a spatial score, \( N \) is the number of points in \( C_{new} \), \( M_j \) and \( \Sigma_j \) are the corresponding mean and covariance of \( p'_{new, j} \) respectively. The corresponding ND of \( p'_{new, j} \) is the voxel where \( p'_{new, j} \) is. Blue balls in Fig. 4(a) are \( C_{new} \). Each ball is connected to an ellipsoid in the voxel where the ball is. To maximize \( S_{sp} \), it is common to use Newton’s method or Levenberg-Marquardt Algorithm.

C. The Review of the HANDT Algorithm

As well known, HANDT is a color-supported variant of NDT utilizing the hue from HSV model. HANDT is in the same frame work of original NDT which are dividing, computing, and registering.

\[ p'_{new} = Rp_{new} + T \]  
\[ T = \left( T_x \quad T_y \quad T_z \right)^T \]
\[ R = \begin{pmatrix} 1 & 0 & 0 \\ \cos \beta & 0 & \sin \beta \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \cos \gamma & -\sin \gamma & 0 \\ \sin \gamma & \cos \gamma & 0 \\ 0 & 0 & 1 \end{pmatrix} \]

where \( s \) is the score of \( p'_{new} \), which is a vector transformed from \( p'_{new} \) by rotational matrix \( R \) and translational vector \( T \). \( p'_{new, j} \) is a vector transformed from the \( j \)th newly scanned point \( p_{new, j} \) by (4). \( S_j \) is computed by substituting \( p'_{new, j} \) to (3).

At the second step, it computes the mean \( M_i \) and covariance \( \Sigma_i \) of \( C_{ref, j} \) by

\[
M_i = \frac{1}{N_i} \sum_{j=1}^{N_i} p_{ref, ij}
\]
\[
\Sigma_i = \frac{1}{N_i - 1} \sum_{j=1}^{N_i} (p_{ref, ij} - M_i)^2
\]

where \( p_{ref, ij} = (p_{x, ij}, p_{y, ij}, p_{z, ij}) \) is the position of the \( j \)th point in the cloud \( C_{ref, j} \), and \( N_i \) is the number of points in \( C_{ref, j} \). \( M_i \) and \( \Sigma_i \) are used to compose the \( i \)th normal distribution (ND) function in the form of (3).

\[
s = \exp \left( -\frac{1}{2} (p'_{new} - M_i)^T \Sigma_i^{-1} (p'_{new} - M_i) \right)
\]
extra mean $\mu_{ij}$ and variance $\sigma_{ij}$ for hue as

$$\mu_{ij} = \frac{1}{N_j} \sum_{k=1}^{N_j} p_{\text{ref},k,h}$$

(8)

$$\sigma_{ij} = \frac{1}{N_j} \sum_{k=1}^{N_j} (p_{\text{ref},k,h} - \mu_{ij})^2$$

(9)

where $N_j$ is the number of points in $g_{\text{hue},ij}$, and $p_{\text{ref},k,h}$ is the hue of $k$th point $p_{\text{ref},k}$. $\mu_{ij}$ and $\sigma_{ij}$ compose a normal distribution weight function as

$$w = \exp\left(\frac{-(p_{\text{new},h} - \mu_{ij})^2}{2\sigma_{ij}}\right)$$

(10)

where $p_{\text{new},h}$ is the hue of a new point $p_{\text{new}}$. The weight $w$ is computed by substituting $p_{\text{new},h}$ to (10). There are at most $n_{\text{hue}}$ normal distributions of the score and weight in a voxel. In contrary to the original NDT as shown in Fig. 4(a), there are multiple ellipsoids in a voxel in Fig. 4(b).

At the final step, HANDT registers $C_{\text{new}}$ with $C_{\text{ref}}$. HANDT would search for the corresponding voxel of $p_{\text{new}}$ in $C_{\text{new}}$. The corresponding voxel is the voxel where $p_{\text{new}}$ is. Next, it would search for a corresponding hue group. The group would be decided by which interval its hue is in. The score $s_k$ is computed by (3), and the weight $w_k$ is computed by (10).

HAND iteratively updates $\xi$ by Newton’s method to obtain the best vector $\xi^*$ which maximizes a weighted score function as

$$S_{\text{wsp}} = \sum_{j=1}^{N} w_j s_j$$

(11)

where $w_j$ is the weight, and $s_j$ is the score of $p_{\text{new},j}$. $\xi$ is iteratively updated as

$$\xi_{k+1} = \xi_k + \alpha \Delta \xi_k$$

(12)

where $\Delta \xi_k$ is a step vector, and $\alpha$ is a step length decided by Armijo’s rule in $k$th iteration.

III. THE IMPROVED HANDT

The HANDT algorithm is improved by two modifications. First, the score function is changed from the normal distribution function to the Mahalanobis distance function. Second, the mean and variance of the hue are changed from simple functions to circular functions. After those modifications are described, the details of improved HANDT algorithm would be described.

A. Mahalanobis Distance

The example of the two dimensional normal distribution in a grid is as shown in Fig. 6(a). Since the scores outside the certain range are approximately zero, NDT or HANDT is not able to attract the exact correspondence which are out of that range. How to attract those corresponding points as well as avoid the effect of noise and outliers to obtain the better step vector is the problem. To score all points in the voxel, it is appropriate to replace the normal distribution score function by Mahalanobis distance function.

The Mahalanobis distance function is

$$D_m(p) = \sqrt{(p - M)^T \Sigma^{-1} (p - M)}$$

(13)

where $p$ is the coordinate vector. The score function is changed to be in the form of the square of the Mahalanobis distance as

$$s_{\text{md}} = D_m^2(p_{\text{ref}}) = (p_{\text{ref}} - M)^T \Sigma^{-1} (p_{\text{ref}} - M)$$

(14)

B. The Circular Mean and Variance of the Hue

Since the hue value is circular, the simple mean and variance would lead to a wrong distribution. Therefore, in the HANDT algorithm, the simple mean and variance of the hue are used. The circular mean $\mu_{\text{cir}}$ is computed by

$$\mu_{\text{cir}} = \frac{1}{2\pi} \tan^{-1}\left(\sum_{j=1}^{N} \frac{\sin(2\pi p_{\text{ref},j,h})}{\sum_{j=1}^{N} \cos(2\pi p_{\text{ref},j,h})} + k\pi\right)$$

(15)

where $k = 1$ if the sum of cosine is smaller than 0, $k = 2$ if the sum of sine is smaller than 0 and the sum of cosine is bigger than 0, and $k = 0$ for others.

For the circular property of the hue, the difference between two hue values should be in the range as $[-0.5,0.5]$. The circular difference is formulated as

$$d_h(h_1,h_2) = \min\{|h_1-h_2|,1-|h_1-h_2|\}$$

(16)

Thus, the circular variance $\sigma_{\text{cir}}$ is computed by

$$\sigma_{\text{cir}} = \frac{1}{N-1} \sum_{j=1}^{N} d_h(p_{\text{ref},j,h},\mu)^2$$

(17)

The circular weight $w_{\text{cir}}$ is computed by

Fig. 6. The score distributions of one hundred random points. The mean is $(0.52,0.45)$. (a) The normal distribution score function. (b) The Mahalanobis distance score function.

Fig. 7. Comparison between the distribution of the simple mean and variance and the distribution of the circular mean and variance.
The circular distribution and simple distribution of the hue are compared as Fig. 7. The circular mean is almost at the center of hue values while the simple mean is at the center of the hue range which is the wrong mean.

C. The Improved HANDT Algorithm

The process of the improved HANDT is as shown in Table I. The main difference between HANDT and improved HANDT is the function of the weight and the score.

\[
S_{\text{imp}} = \sum_{i=1}^{n_{\text{hue}}} w_{\text{cir}, i} S_{\text{md}, i}
\]

(19)

where \( w_{\text{cir}, i} \) is the weight of \( p_{i} \), and \( S_{\text{md}, i} \) is the spatial score of \( p_{i} \). Since the score function is the square of the Mahalanobis distance, the optimization problem is changed to finding the transformation vector \( \xi^{*} \) which minimizes (19), and \( \xi^{*} \) is computed iteratively by (12). \( \Delta \xi \) in (12) is computed by

\[
\Delta \xi = \frac{H_{k}^{-1} g_{k}}{H_{k}^{-1} g_{k}}
\]

(20)

where \( g_{k} \) and \( H_{k} \) are the first and second order partial derivatives of (21) which are substituted with \( \xi_{k} \).

The weighted score function of the improved HANDT is modified as

\[
S_{\text{imp}} = \sum_{i=1}^{n_{\text{hue}}} w_{\text{cir}, i} (p'_{\text{new}, i} - m_{i})^{T} \Sigma_{i}^{-1} \frac{\partial p'_{\text{new}, i}}{\partial \xi_{i}} g_{k}
\]

(21)

\[
H_{k, uv} = \sum_{i=1}^{n_{\text{hue}}} w_{\text{cir}, i} (p'_{\text{new}, i} - m_{i})^{T} \Sigma_{i}^{-1} \frac{\partial p'_{\text{new}, i}}{\partial \xi_{i}} g_{k}
\]

(22)

where \( H_{k, uv} \) is the \((u, v)\) entry of \( H_{k} \).

IV. EXPERIMENTS

NDT, HANDT, and the improved HANDT are implemented by Point-Cloud-Library (PCL) [11]. We ran the implementation on Intel Core i7-3770 with PC1333 DDR3 dual-channel memory. Data sets we use are from [12]. Data sets were recorded by Microsoft Kinect, and the ground-truth trajectories were captured with 100Hz tracking cameras. ‘freiburg1_room’, ‘freiburg2_desk’, ‘freiburg3_nostructure_texture_far’, and ‘freiburg3_structure_texture_far’ are chosen.

First, the performance of the different \( n_{\text{hue}} \) of the improved HANDT are evaluated. Second, the performance of NDT, HANDT, and the improved HANDT are evaluated. The length of voxels is set to 5cm for the data set of the room and 10cm for other data sets. The time step is 0.27 second. To avoid singular covariance, the spatial score function of a hue group is available only if the number of points in the group is greater than 5. The performance are compared in terms of the iteration, runtime, translation error, and rotation error. The convergence criterion is the step length. If the length is smaller than a threshold which is \( 10^{-6} \), then the algorithms will stop.

A. Evaluation of Improved HANDT of Different \( n_{\text{hue}} \)

\( n_{\text{hue}} = 1 \) is chosen to evaluate to check how the average of hue supports NDT. Due to three primary colors, \( n_{\text{hue}} \) starts with 3. Next, \( n_{\text{hue}} \) is doubled up to 24. The results are as shown in Table II.
As shown in Table II, the performance of the improved HANDT become better as \( n_{\text{hue}} \) increases until \( n_{\text{hue}} = 12 \). Since the hue intervals become narrower as \( n_{\text{hue}} \) increases, the number of points in each hue group becomes smaller. Due to the constraint against singular covariance, the number of available groups decreases. Hence, more newly scanned points are possible to be dropped because the corresponding groups are not available. Moreover, as \( n_{\text{hue}} \) increases, the hue intervals become narrower, and the hue of a newly scanned point and the hue mean of the corresponding group become closer. It means that the point clouds are possible to find more appropriate correspondences. Due to this reason, the errors become lower as \( n_{\text{hue}} \) increases. However, as \( n_{\text{hue}} \) is 24, the errors increase. The reason is that the algorithm drops too many newly scanned points because the hue intervals are too narrow to be available. As the result shown in Table II, the result of \( n_{\text{hue}} = 12 \) is generally the best.

### B. Evaluation of NDT, HANDT, and Improved HANDT

In this experiment, \( n_{\text{hue}} \) of HANDT and the improved HANDT are set to 12. The results are as shown in Table III. Graphs in Fig. 8 show the convergence of algorithms. The point cloud sat 9116.45 and 9116.68 second in “freiburg3_structure_texture” data set are used. The data is also used to show the score function against the step length as shown in Fig. 9.

![Fig. 8. The convergence of algorithms. (a) NDT (b) HANDT (c) Improved HANDT.](image)

As the results shown in Table III, although the runtime per iteration of the improved HANDT is shorter, the runtime is longer than NDT and HANDT. However, the translation and rotation errors of the improved HANDT are lower than those of NDT and HANDT.

<table>
<thead>
<tr>
<th>TABLE III: THE COMPARISON BETWEEN ALGORITHMS ‘IMHANDT’ IS THE IMPROVED HANDT</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dataset</strong></td>
</tr>
<tr>
<td>---------------</td>
</tr>
<tr>
<td>Freiburg 1 room</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Freiburg 2 desk</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Freiburg 3 Nostructure Texture (far)</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Freiburg 3 Structure Texture (far)</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

Fig. 8(a) shows the convergence of NDT. It shows that NDT would stop after the second iteration. Fig. 9(a) and Fig. 9(b) show the score functions of NDT against the step length at the first and second iterations. The step length obtained at second iteration is approximately zero. If the step length is shorter than threshold, we regard that the registration is converged.

As shown in Table III, the translation and rotation errors of HANDT and the improved HANDT are lower than NDT. The reason is that the reference point cloud is more geometrically well-represented by distributions of hue groups. In addition, the algorithms find better correspondences than NDT. However, the runtime of HANDT and the improved HANDT are longer than NDT. Fig. 8(b) and Fig. 8(c) show HANDT and the improved HANDT need more iterations until stop than NDT.

The improved HANDT performs the more accurate registration than HANDT due to the Mahalanobis distance score function. For a point far away from the mean of the voxel, whose eigenvalues of the covariance matrix are small, the score of the point computed by improved.

![Fig. 9. Score functions. (a) and(b) are the score functions of NDT plotted against the step length. (c) and (d) are those of HANDT. (e) and (f) are those of the improved HANDT. (a), (c), (e) are the score functions of the 1st iterations, and (b), (d), (f) are those of the 2nd iterations.](image)
this experiment, the initial guess of the step length is chosen as \( ||H_k^{-1}g_k|| \), which is much smaller than one. Since the initial guess usually satisfies Armijo’s rule, the runtime per iteration of the improved HANDT is shorter than HANDT and NDT.

V. CONCLUSION

The performance of improved HANDT is evaluated by benchmark data sets. As results, the accuracy of the registration is much better than the cases of NDT and HANDT. The improvement is approached by changing the weighted score function to Mahalanobis distance function and the modification of the hue mean and hue variance. Improved HANDT is expected to overcome the difficulty of the registration at flat or repeated structure by the hue.

ACKNOWLEDGMENT

This work was supported in part by the National Research Foundation of Korea(NRF) grant funded by the Korea government (MSIP) (No. 2013R1A2A1A05005547), in part by the Brain Korea 21 Plus Project, in part by ASRI, in part by the Industrial Foundation Technology Development Program of MOTIE/KEIT [Development of CIRT (Collective Intelligence Robot Technologies)], and in part by Bio-Mimetic Robot Research Center funded by Defense Acquisition Program Administration [UD130070ID].

REFERENCES


Hyun K. Hong received his B.S. degree in electrical and computer engineering from Seoul National University in 2012. He is currently a M.S candidate in the Department of Electrical and Computer Engineering at Seoul National University. His current research interests include graph SLAM and multi-agent control.

Boem H. Lee received the B.S. and M.S. degrees in electronics engineering from Seoul National University, Seoul, the Republic of Korea in 1978 and 1980, respectively, and the Ph.D. degree in computer, information, and control engineering from the University of Michigan, Ann Arbor, MI in 1985.

He was an assistant professor with the School of Electrical Engineering at Purdue University, West Lafayette, IN from 1985 to 1987. He joined Seoul National University in 1987, and is currently a professor with the Department of Electrical and Computer Engineering. His research interests include multi-agent system coordination, control, and application.

Prof. Lee has been a fellow of the Robotics and Automation Society since 2004.