The 2009 IEEE/RSJ International Conference on
Intelligent Robots and Systems
October 11-15, 2009 St. Louis, USA

3D Environment Reconstruction Using Modified Color ICP Algorithm by Fusion of a Camera and a 3D Laser Range Finder

Ji Hoon Joung, Kwang Ho An, Jung Won Kang, Myung Jin Chung and Wonpil Yu

***Abstract*—In this paper, we propose a system which reconstructs the environment with both color and 3D information. We perform extrinsic calibration of a camera and a LRF (Laser Range Finder) to fuse 3D information and color information of objects. We also formularize an equation to measure the result of the calibration. Moreover, we acquire 3D data by rotating 2D LRF with camera, and use ICP (Iterative Closest Point) algorithm to combine data acquired in other places. We use the SIFT (Scale Invariant Feature Transform) matching for the initial estimation of ICP algorithm. It offers accurate and stable initial estimation robust to motion change compare to odometry. We also modify the ICP algorithm using color information. Computation time of ICP algorithm can be reduced by using color information.**

I. Introduction

It is essential to determine the regions where vehicles can reach and to plan the paths where the vehicles should go when it comes to UGV (Unmanned Ground Vehicle) navigation.

Path planning requires determining reachable regions which require terrain perception, and there are many advantages for using terrain perception such as accuracy [1] if we use both 3D model and color information. Therefore, effective reconstruction of environment with 3-dimensional and color information enables improving the UGV navigation.

When a rescue robot sends information from an accident place, 3D and color information together are more helpful for understanding the situation of the accident place than sending just one of them.

To reconstruct 3D environment, Andreas Nüchter [2] used rotational 3D LRF and data acquired with stop-scan-go fashion. ICP algorithm was applied to infer the motion of the robot. The initial estimation for ICP algorithm was given by odometry information with a heuristic method. The odometry information provides 2D information of motion of a robot, so a problem occurs when the robot shows a big 3D motion difference between frames, and the reconstructed 3D environment doesn't include color information. Michael Montemerlo [3] installed a rotational 3D LRF and GPS/IMU on a Segway to reconstruct 3D environment without color. Yunsu Bok [4] installed an LRF and a camera on a vehicle. He reconstructed environment with color by fusing a camera and an LRF. He found correspondences of the fusion data and inferred motion between frames with correspondences. If the fusion data is acquired from plane, it may cause a degenerate case.

In this paper, we propose a system which reconstructs the environment with both color and 3D information. We install a rotational 3D LRF and a camera on a mobile robot (Fig. 9).

The geometric relation between a camera and an LRF is acquired by extrinsic calibration of the camera and the LRF to fuse 3D information from the LRF and color information from the camera. We suggest a method for measuring and reducing the error of extrinsic calibration of the camera and the LRF.

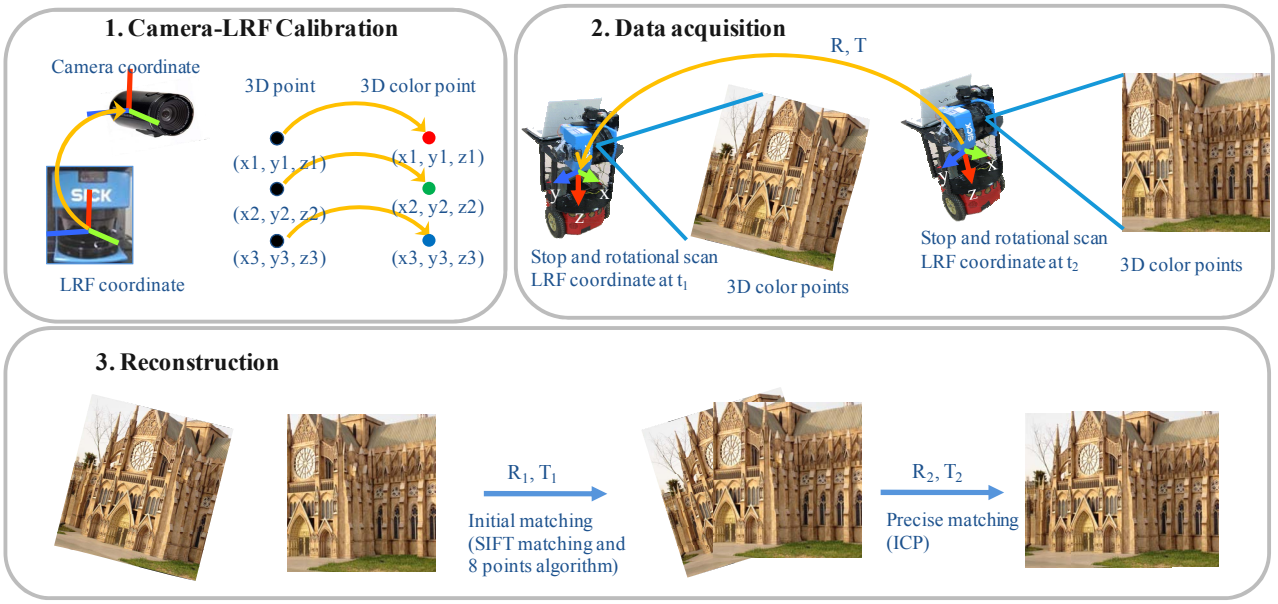
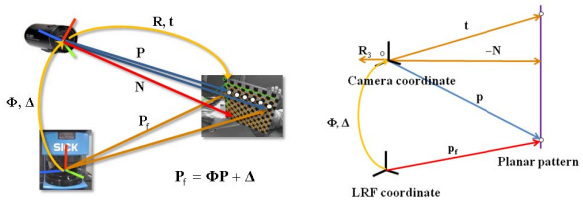


Fig. 1. System Overview

Ji Hoon Joung and Wonpil Yu are with Robot Research Department, Electronics and Telecommunications Research Institute, Daejeon, Korea.(e-mail: jihoonj@etri.re.kr, ywp@etri.re.kr).

Kwang Ho An, Jung Won Kang and Myung Jin Chung is with Department of Electrical Engineering & Computer Science, Korea Advanced Institute of Science and Technology, Daejeon Korea.(e-mail: mjchung@ee.kaist.ac.kr)



(a) (b)

Fig. 2. Geometric relation between coordinates

We use stop-scan-go fashion to acquire data from different places and ICP algorithm to infer the motion. We use motion estimation using SIFT matching for the initial estimation of ICP algorithm. It provides a stable result robust to robot motion even though the robot has 3D movement. Moreover, modified color ICP algorithm is proposed to reduce the execution time.

II. Extrinsic Calibration of a Camera and a LRF

The result of single camera calibration [5, 6] and the data of planar plane whose surface is defined by a checkerboard (Fig. 3) in an LRF plane is used for extrinsic calibration of a camera and a LRF for fusion of 3D information from the LRF and color information from the camera [7].

A. Geometric Constraints

A 3D point of the planar pattern in an LRF coordinate, P_f can be mapped to a 2D point on the image coordinate, p using (1).

$$\begin{aligned} \mathbf{P_f} &= \mathbf{\Phi P} + \mathbf{\Delta} \\ \mathbf{p} &= \mathbf{KP} \\ \mathbf{p} &= \mathbf{K\Phi^{-1}(P_f - \Delta)} \end{aligned} \tag{1}$$

P represents a 3D point in the camera coordinate, and Φ is a rotation matrix between a camera and an LRF w.r.t, LRF. Δ is a translation vector between a camera and an LRF w.r.t. LRF, and K is intrinsic parameters of camera.

That is, the result from camera calibration (K, R (rotation between the camera and a planar pattern w.r.t camera), t (translation between camera and planar pattern w.r.t camera)) and Pf gives us the result of camera-LRF calibration (Φ, Δ). These can be used to combine the data from the LRF and the camera. (Fig. 2(a))

We use vector N in order to get Φ and Δ. N is parallel to vertical vectors in planar pattern, and this vector has distance from origin of camera frame to planar pattern as its magnitude. This vector can be derived from the geometric relation of the planar pattern from calibration of a camera. R_3 in Fig. 2(b) is the third column of R. This N satisfies (2) with a new point P from the camera coordinate.

$$\mathbf{N^T P} = \|\mathbf{N}\|^2 = \mathbf{N^T \Phi^{-1}(P_f - \Delta)} \tag{2}$$

Φ and Δ can be derived from (2), and the accuracy of calibration can be improved by optimization with enough number of P_f and N.

(a) (b)

Fig. 3. Projected planar plane points (p).

B. Measuring the error

True value of calibration is hard to get due to uncertain origins of the LRF and the camera. The error of result from calibration with method above was examined through mapping the LRF planar pattern data and the calibration result on the image plane. Calibration error may appear due to gauche pattern plane, LRF error, or camera calibration error. In order to reduce this error from calibration, it is necessary to make the error in numeric.

In order to measure the error, we assume that the accuracy of result from camera calibration is reliable and the scan plane of the LRF and the bottom are parallel. Mapped points on image plane and (1) are used for this. Extrinsic calibration of the camera and the LRF is performed with various poses of planar plane, not with a single pose.

Points in Fig. 3 (a) are the mapped points **p** projected with accurate extrinsic calibration of the camera and the LRF. Mapped points are parallel to L_C because the scan plane of the LRF and the bottom are parallel as assumed. We use just the points in the planar plane, therefore mapped points fit in the planar plane. Points in Fig. 3 (b) are the mapped points **p** projected with inaccurate extrinsic calibration of the camera and the LRF. We numerate the difference between Fig. 3 (a) and Fig.3 (b), and use it as an error formula.

$$\sum_{i=1}^{k} [\, w_{di} d_i + w_{ci} \sin\theta_i \,] + m_{err} \tag{3}$$

k is the number of pose used for calibration. i means the i-th pose of planar plane. L_M is the mapped points and L_L, L_C and L_R are respectively the left, center and right line of planar plane acquired by the corner points of planar plane. d_i indicates how far the mapped points (L_M) are from the planar plane. d_1 is distance between L_L and the most left point of L_M. d_2 is distance between L_R and the most right point of L_M. The larger of d_1 and d_2 becomes d_i. The internal angle of L_M and L_C is θ_i. Therefore, if L_M and L_C are parallel, then $\sin\theta_i$ equals 0. Inaccurate extrinsic calibration of the camera and the LRF makes a large internal angle of L_M and L_C, and it brings large $\sin\theta_i$. w_{di} and w_{ci} are weight coefficients of d_i and $\sin\theta_i$. d_i ranges from 0 to image width, and $\sin\theta_i$ ranges from 0 to 1. So we must equalize the order of $\sin\theta_i$ and d_i using w_{di} and w_{ci}. We measure the approximate value of translation between a camera and a LRF. If the translation from the extrinsic calibration of the camera and the LRF is far from the measure value, m_{err} is added.

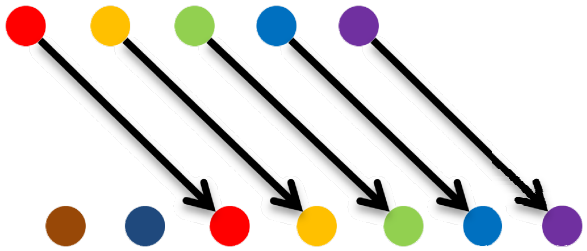


Fig. 4. Closest points with color information.

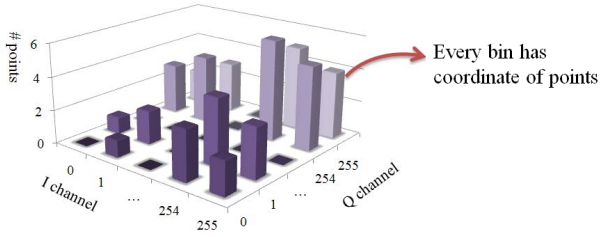


Fig. 5. 2D histogram using I and Q channel

In order to get a calibration result with less error, we went through the following. Some points were selected through RANSAC (Random Sample Consensus) in order not to use inaccurate points of the planar pattern, and the data (P_f), the rectified plane pattern, was created with the points selected through line fitting. This rectified data was applied to (2) to get a new calibration result. Then, using (3), we made the error in numeric. Through some iteration of this procedure, we have chosen the calibration result with the least error in numeric.

III. Motion Estimation

We used the ICP algorithm [8] to infer the geometric relation between the data acquired in different positions. The ICP algorithm requires two clouds of points and the initial geometric relation between the two clouds of points. It is assumed that the initial geometric relation is close enough or approximately known. The output of ICP algorithm shows the accurate 3D motion between 2 clouds of points.

Generally odometry is used for the initial estimation of ICP. However, odometry gives only 2D information of motion and sometimes is inaccurate. So we used 3D motion estimation using SIFT matching [9] for the initial estimation of ICP. The SIFT features are local and based on the appearance of the object at particular interest points, and are invariant to image scale and rotation. They are also robust to changes in illumination and noise. Consequently, it gives us stable and accurate initial estimation compared to odometry.

A. ICP algorithm

ICP algorithm is processed through the following 5 steps :

1) Find the closest points satisfying the distance constraints.
2) Update the matching through statistic analysis of distances.
3) Compute the motion between the two frames from the updated matches.
4) Apply the motion to all points and their tangents in the previous frame.
5) Repeat 2) to 4).

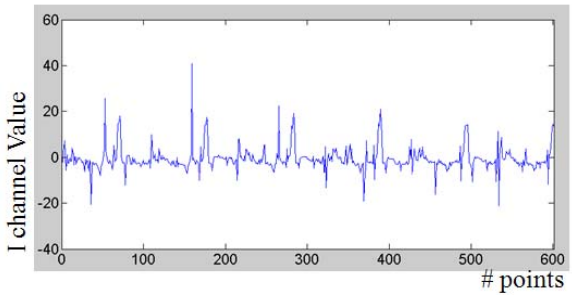
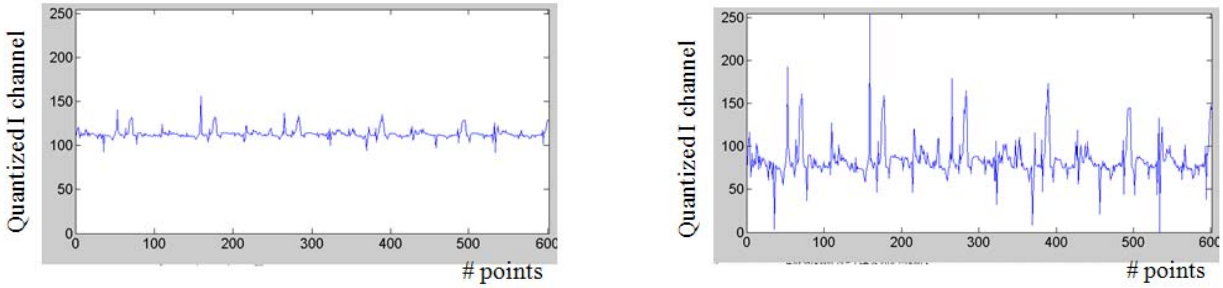


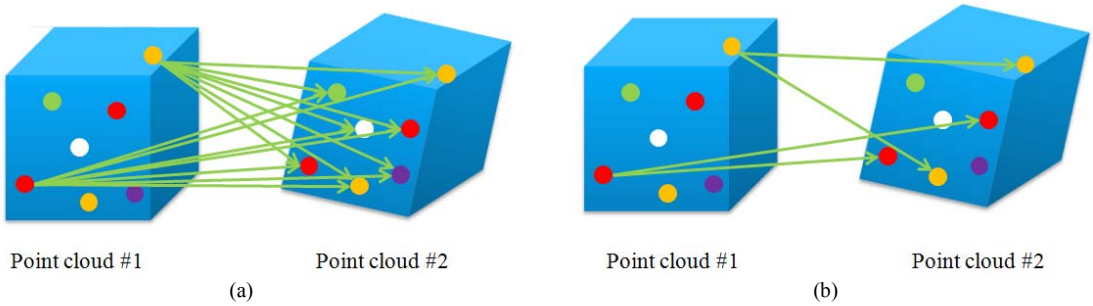
Fig.6. Distribution of I channel value before quantization.

Let us define **x** a point in the first cloud of point. Then we find the closest point **y** in the second cloud of points to the given point x. The closest point is the point which has the least Euclidean distance to a given point. We apply above procedure to all points in the first cloud of points. So each point x has the closest point y. And if any distance of the closest points is larger than maximum tolerable distance, we exclude the pair. Then the maximum tolerable distance is renewed using statistic analysis of the closest points. If the distance of the closest points is larger than the renewed maximum tolerable distance, we exclude that pair. The motion between two clouds of points is computed with the updated closest points. After computing the motion of point clouds, we apply it to the first cloud of point, and iterate the procedure above.



(a) (b)

Fig.7. (a) Distribution of I channel quantized with fixed range. (b) Distribution of I channel quantized with dynamic range.



(a) (b)

Fig.8. (a) Concept of color ICP which searches all points to find closest points. (b) Concept of modified color ICP which searches similar color points to find closest points.

B. *Closest Point with Color Information (Color ICP)*

If we use not only the 3d information of points, but also the color information fused by camera and LRF calibration, we can get improved result when we find closest points [10]. (Fig. 4) This is the color ICP.

Equation (4) is the definition of distance for color ICP.

$$d_{color} = \|\mathbf{x_1} - \mathbf{x_2}\| + \sqrt{\alpha_1(a_1 - a_2)^2 + \alpha_2(b_1 - b_2)^2 + \alpha_3(c_1 - c_2)^2} \quad (4)$$

x_1, x_2 are the points with 3D information, and $\|\mathbf{x_1} - \mathbf{x_2}\|$ is the Euclidean distance between x_1 and x_2. α_i (i=1~3) is a weight coefficient, and a_j, b_j, c_j (j = 1, 2) are respectively YIQ model of Y, I and Q value transformed from RGB model. Reference [10] used YIQ model instead of RGB model in order to reduce the effect of luminance change. In the YIQ model, the intensity of light is conveyed by the Y channel and the hue and the saturation is conveyed by the I and Q channels. To reduce the effect of luminance change, we make coefficient for the Y channel very small.

C. *Modified Color ICP*

Color ICP and ICP search all points in the cloud of points to find the closest points. These algorithms have order of $O(n^2)$ time complexity for searching the closest point. We propose a modified color ICP algorithm which reduces search time to find the closest points and keeps the accuracy merit of color ICP. We only search similar colors to find the closest points. We analysis the intrinsic color information (I channel and Q channel) of a point cloud and made 2D histogram (Fig. 5) for this.

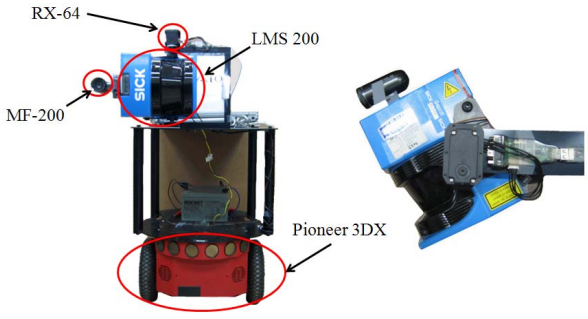


Fig.9. System configuration

We quantize the I and the Q channel of YIQ model from 0 to 255 respectively. We can make a 2D histogram with 65,536 bins, and each bin includes the 3D coordinates of points. So when we find the closest points to **x**, we just search some points in the neighboring bins whose center bin has the same color with **x**. (Fig. 11) The color information is used to determine which point is searched, so that we can find the closest points more accurately than ICP. We also use Euclidean distance instead of (4). This fact brings that we can easily update the maximum tolerable distance same with [8]. The image is imported with RGB model and its range is from 0 to 255. When we linearly transform of RGB model to YIQ model, the maximum and the minimum value of YIQ model are fixed as ±151.91 for the I channel and ±133.26 for the Q channel. If I and Q channel is quantized from 0 to 255 with the fixed range, the data can be concentrated in some values. (Fig. 7(a)) It means that there are too many points in a particular bin of 2D histogram, so the search time can't be reduced enough. We propose a solution to this problem as the following. Figure 6 shows the distribution of I channel value before performing quantization. The max and min value in Fig. 6 is less than a third of ±151.91. Figure 7(a) shows the distribution of I channel value after performing quantization. We can confirm that the distribution of I channel value is concentrated in a narrow range. We use the dynamic max and min value instead of fixed max and min value to solve this problem. At first, we analyze the max and the min value of each I and Q channel of cloud of points, and apply (5) to quantize the analyzed value from 0 to 255.

$$Q = \frac{255}{V_{Max} - V_{Min}} \times V + \left(\frac{-255}{V_{Max} - V_{Min}} \times V_{Min}\right) \quad (5)$$

V and Q are the values before and after quantizing

(a) (b)

Fig.10. (a) Projected planar plane point, p using (7). (b) Projected planar plane point, p using (8).

TABLE I
COMPARISON OF SEARCHING TIME TO FIND CLOSEST POINTS

	Number of search bins		
	1 bin	9 bins	25 bins
Full search (ICP and color ICP)	15985 ms		
Fixed range with 2D histogram	647 ms	3203 ms	6235 ms
Dynamic range with 2D histogram (modified color ICP)	315 ms	542 ms	1000 ms

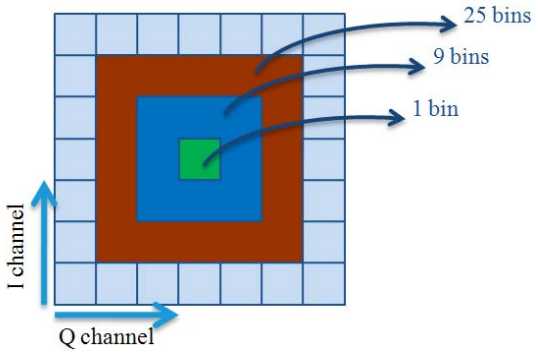


Fig.11. 2D histogram with I and Q channel and number of bins.

respectively. V_{Max} and V_{Min} are the max and the min values of the analyzed values respectively. Figure 7(b) shows the distribution of the I channel value after performing quantization of Fig. 6 using dynamic max and min range. The distribution of Fig. 7(b) is wider than Fig. 7(a). As a result, the search time to find the closest points is reduced.

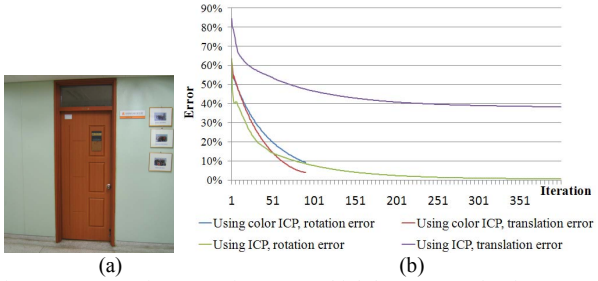
IV. EXPERIMENTS

Figure 9 shows the mobile system for reconstruction 3D environment with 3D and color information. The mobile system is composed of a mobile robot (Pioneer 3DX), a 2D LRF (LMS200), a motor (RX-64), a Camera (MF-200) and body frame.

A. Camera LRF calibration

Equation 6 is the measured approximate value of translation between a camera and a LRF

$$\Delta = [35\text{mm}\ 190\text{mm}\ -70\text{mm}]^T \quad (6)$$



(a) (b)

Fig.13. (a) Experiment environment which has some color feature and little structure feature. (b) The graph of error and ICP(or modified color ICP) iteration.

Figure 10(a) and (7) are the calibration result of [7]. Figure 10(b) and (8) are the calibration result with proposed method which uses RANSAC for line fitting. Equation (3) is used for Err_1 and Err_2.

$$\Phi_1 = [1.14, -93.30, -5.67]^o \quad \Delta_1 = [23.32, 232.02, -78.07]\text{mm} \quad (7)$$
$$Err_1 = 41.68$$
$$\Phi_2 = [-0.46, -93.26, -1.99]^o \quad \Delta_2 = [39.23, 197.49, -64.46]\text{mm} \quad (8)$$
$$Err_2 = 14.78$$

Comparing Δ_1 and Δ_2, Δ_2 is closer to Δ than Δ_1, and Err_2 is smaller than Err_1. Mapped points at Fig. 10(b) fit to planar plane better than the mapped points at Fig 10(a).

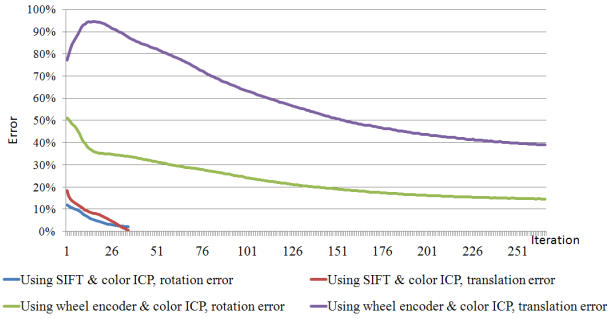
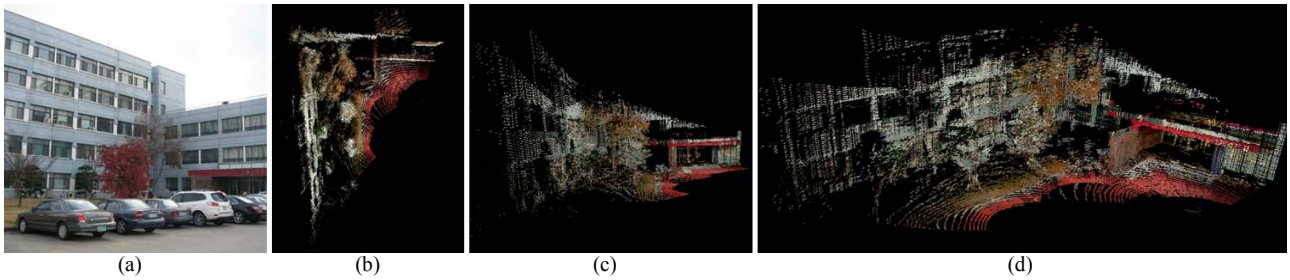


Fig.14. The error and iteration graph of odometry and motion estimation using SIFT matching

B. Modified Color ICP algorithm

TABLE I is measured with Quad core 2.4GHz, 2GB memory and 30488 points containing color information. When we searched using 2D histogram and the dynamic



(a) (b) (c) (d)

Fig.12. (a) Experiment environment. Reconstruction result of experiment environment. (top view(a), side view(b), front view(c))

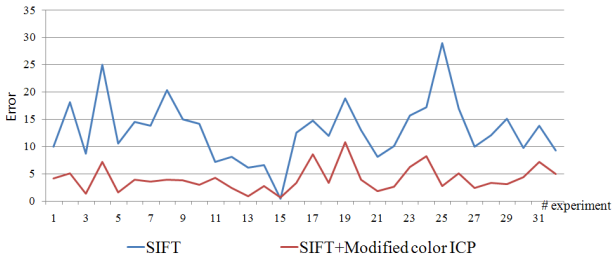


Fig.15.The rotation error graph of SIFT and modified color ICP using SIFT

range, and the execution time took 1.97% of full search time at minimum and 16.03% of search time using 2D histogram and fixed range.

C. Outdoor Reconstruction

Figure 12 shows the outdoor reconstruction of 30m x 25m environment with the proposed system. Total of 119,828 points were used.

D. ICP and Modified Color ICP

The experiment was performed with the environment shown in Fig. 13(a) to compare the convergence speed of ICP and modified ICP. We terminated the iteration when both rotation and translation errors were less than 10%. The rotation error was calculated by transforming rotation matrix to vector using Rodrigues' formula. The environment shown in Fig. 13(a) is close to plane so it contains little features of structure but some features of color.

E. Initial Estimation (Odometry and SIFT matching)

We applied modified color ICP with different initial estimation to infer the motion of 2 clouds of points which has 3D geometric relation between them. Figure 14 shows that the difference when we used odometry and SIFT matching for initial estimation. When there were 3D motion between two clouds of points, convergence speed of using odometry for initial estimation of modified color ICP was much slower than using motion estimation with SIFT matching. Therefore, SIFT matching provided the stable initial estimation robust to motion.


Fig.16.Reconstructed environment and path of a robot(white line at the center)

F. SIFT Matching and Modified Color ICP with SIFT Matching

To confirm whether the accuracy of modified color ICP using SIFT matching has been improved better than the accuracy of motion estimation using just SIFT matching, we compared the accuracy of both methods. Figure 15 shows the result of 32 times. We confirmed the accuracy has been improved in most cases. The average error of using just SIFT matching was 12.25%, and the average error of using modified ICP with above result was 4.09%.

TABLE Ⅱ
POSITION OF ROBOT

Position of robot	X(m)	Y(m)	Z(m)
1	0	0	0
2	0.6201	-0.6203	0.00414
3	-0.6456	-0.6170	0.01011
4	-1.757	-1.305	0.06042
5	-2.189	0.5216	0.07107
6	-1.161	0.5063	0.07080
7	0.6379	0.5077	0.07069
8	0.01934	-0.02180	-0.00104

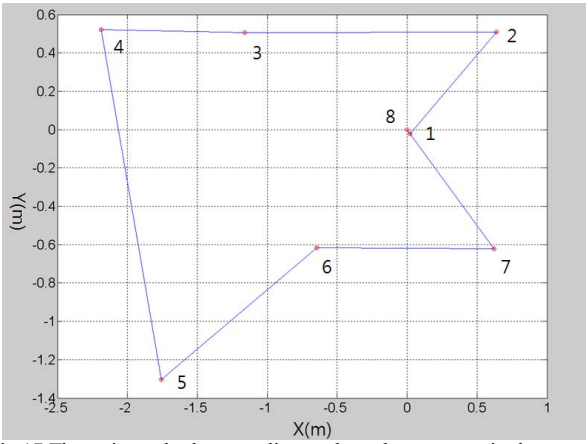


Fig.17.The estimated robot coordinate where data are acquired.

G. Measuring accuracy

We checked the accuracy of proposed system by reconstructing the 7m x 7m indoor environment. We set the path of robot as a loop, and we inferred the motion of robot at every spot where the robot obtained the data. Therefore the first spot and the last spot have same coordinate, and distance between the first spot and last spot can be considered as an error. Figure 16 shows the reconstruction result and the path of robot (white line in center of Fig. 16). TABLE Ⅱ and Fig. 17 are the coordinates of the robot where the data was acquired. The distance between the first and the last spot is

0.0292m. It's 0.47% of 6.25m the whole moved distance.

V. CONCLUSIONS AND FURTHER WORKS

In this paper, we present a system which reconstructs the environment with both color and 3D information. We can improve the accuracy of extrinsic calibration of a camera and a LRF by measuring and reducing errors. Motion obtained with SIFT matching for the initial estimation of ICP makes the stable system even more robust to motions of a robot. We also propose the modified color ICP maintaining the merit of color ICP in accuracy and reducing the execution time by searching through similar color points.

ACKNOWLEDGMENT

Authors are gratefully acknowledging the financial support by Agency for Defence Development and by UTRC(Unmanned Technology Research Center), Korea Advanced Institute of Science and Technology.

This work was supported partly by the R&D program of the Korea Ministry of Knowledge and Economy (MKE) and the Korea Evaluation Institute of Industrial Technology (KEIT). [2008-S-031-01, Hybrid u-Robot Service System Technology Development for Ubiquitous City]

REFERENCES

[1] P. Vernaza, B. Taskar, Daniel D. Lee, "Online, self-supervised terrain classification via discriminatively trained submodular Markov random fields," IEEE International Conference on Robotics and Automation, 2008.

[2] N¨uchter, A., K. Lingemann, J. Hertzberg and H. Surmann, "Heuristic-based laser scan matching for outdoor 6D SLAM," In Advances in Artificial Intelligence. Proceedings Springer LNAI. 28th Annual German Conference on, Vol. 3698, pp. 304–319, 2005.

[3] M. Montemerlo and S. Thrun. "Large-scale robotic 3-d mapping of urban structures,", In ISER, 2004.

[4] Yunsu Bok, "Accurate Motion Estimation and Mapping based on Sensor Fusion of Camera and Laser Scanner," M.S. thesis, Division of Electrical. Eng., Korea Advanced Institute of Science and Technology, Daejeon, Korea, 2005

[5] Z. ZHANG, "A flexible new technique for camera calibration," Microsoft Research Technical Report: MSR-TR-98-71, 1998.

[6] Jean-Yves Bouguet, Camera calibration toolbox for matlab, Computational Vision at Caltech, http://www.vision.caltech.edu/bouguetj/calib_doc/.

[7] Q. Zhang and R. Pless, "Extrinsic Calibration of a Camera and Laser Range Finder (improves camera calibration),", in Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, 2004.

[8] Z. Zhang, "Iterative point matching for registration of free-form curves and surfaces,", Int. J. Comput. Vision 13, pp 119–152, 1992.

[9] D. Lowe, "Distinctive image features from scale-invariant keypoints," International Journal on Computer Vision 60(2):91–110, 2004.

[10] A. Johnson and S.B. Kang. "Registration and integration of textured 3-d data." In Proc. Int. Conf. On Recent Advances in 3-D Digital Imaging and Modeling, pp 234–241, May 1997.