Real Time 3D Environment Modeling for a Mobile Robot by Aligning Range Image Sequences

Ryusuke Sagawa, Nanaho Osawa, Tomio Echigo and Yasushi Yagi
Institute of Scientific and Industrial Research, Osaka University,
8-1 Mihogaoka, Ibaraki-shi, Osaka, 567-0047, JAPAN
{sagawa,n-oosawa,echigo,yagi}
@am.sanken.osaka-u.ac.jp

Abstract

This paper describes real time 3D modeling of the environment for a mobile robot. A real time laser range finder is mounted on the robot, and obtains a range image sequence of the environment while moving around. In this paper, we detail our method that accomplished simultaneous localization and 3D modeling by aligning the acquired range images. The method incrementally aligns range images in real time by a variant of the iterative closest point (ICP) method. By estimating the uncertainty of range measurements, we introduce a new weighting scheme into the aligning framework. In an experiment, we first evaluate the accuracy of the localization results by aligning range images. Second, we show the results of modeling and localization when a robot moves along a meandering path. Finally, we summarize the conditions and limitations required of the robot's motion and the environment for our method to work well.

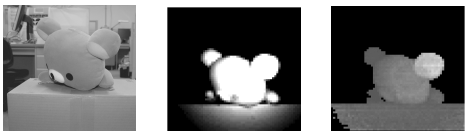
1 Introduction

If a mobile robot creates a map by observing the environment while moving around, the robot must simultaneously estimate it's pose and the geometry of the environment. The issue is well known as simultaneous localization and mapping (SLAM) [13, 8, 4], with the goal being to creating a 2D map of the environment. On the other hand, some robots and autonomous cars [5, 15] have laser range finders for creating a 3D model of the environment. However, their localization tasks are accomplished using other sensors, such as INS and GPS. To achieve these two tasks using a single sensor, this paper describes a method to accomplish simultaneous localization and 3D modeling by aligning range images. Our robot has a real time laser range finder and obtains a range image sequence of the environment. Since each range image contains only part of the shape of the environment, the robot acquires range images from various viewpoints to model the whole shape. Because we assume that the robot has no other sensors for localization, the mutual relationship of the range images is unknown. Thus, we align them into a common coordinate system to model the whole shape of the environment. Since aligning range images is equal to estimating the pose of the sensor, we can obtain a model of the environment and the trajectory of the robot at the same time.

The acquired range images are aligned using registration algorithms that establish point correspondences and minimize the total distance between those points. In this paper,

BMVC 2005 doi:10.5244/C.19.13

Figure 1: Canesta laser range finder DP200.



(a) (b) (c)

Figure 2: An example of a range image: (a) photo of the object, (b) intensity image of the sensor, (c) range image.

we align range images using a variant of an iterative closest point (ICP) method [1, 9, 10]. An ICP method alternately repeats finding the nearest neighbor points of vertices of range images and computing the rotation and translation parameters to minimize the distances of corresponding points. Similar to [12, 11], we apply an ICP method to a range image sequence and incrementally align them in real time to generate a model of the environment. Though a lot of variants of the ICP method exist, we propose one that improves the accuracy by considering the measurement errors of a sensor.

The laser range finder [2] we use is rather noisier than the other slower laser range finders [3, 6] commonly used for modeling the shape of objects with registration techniques. Thus, we first estimate the precision of the range finder. Next, we propose a new aligning method with a weighting scheme based on the estimated uncertainty of the range measurement. Finally, we evaluate the accuracy of the alignment of the range images, and show the localization and modeling results when the robot moves along a meandering path. We summarize the conditions and limitations of the robot's motion and the environment required for our method to work well.

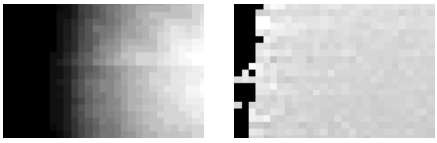
2 Estimation of Accuracy of Range Measurement

2.1 Measurable Range with a Real Time Laser Range Finder

For our real time laser range finder, we use a Canesta DP200 (see Figure 1.) The range finder has an infra-red (IR) light source and a camera that obtains IR light as reflected by objects. The sensor determines distances in each pixel by measuring the difference in phase between the modulated light emitted and the observed light. Figure 2 shows an example of a range image. Figure 2(b) shows an intensity image which represents the strength of the observed light. Figure 2(c) shows the range image, in which the bright pixels indicate a near range while the dark pixels indicate a far range. Distances are not measure for black pixels.

Since the sensor measures reflected light, maximum and minimum measurable ranges are determined by several factors; for example, modulation frequency, material of the object, shutter time and distance to an object. The modulation frequency limits the maximum distance. If the modulation frequency is f and the light speed is C, the measurable maximum distance is less than $\frac{C}{f}$.

The object's material also affects the intensity of the reflected light. Figure 3 shows a range image and the corresponding intensity image of a planar surface whose color gradually changes from black to white. Though the surface is perpendicular to the view



(a) (b)

Figure 3: Range and intensity images of a planar surface whose color gradually changes from black to white. (a) Intensity image. (b) Range image.

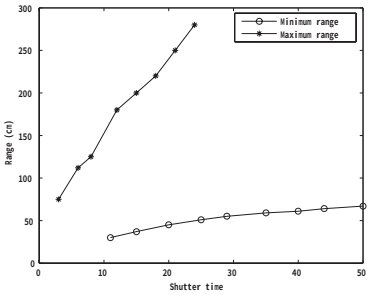


Figure 4: The relationship between shutter time and the measurable range by observing an object at various distances.

direction, the dark area cannot be measured because the intensity of the observed light is low. Thus, the maximum range can be seen to be influenced by the object's material.

By adjusting the shutter time of the camera, the sensor can change the measurable range because the intensity of the observed light increases as the shutter time becomes longer. Figure 4 shows the relationship between shutter time and the maximum and minimum measurable ranges by observing an object at various distances. If an object is farther than the maximum range, reflected light is not observed and the sensor cannot measure the distance. On the other hand, if an object is closer than the minimum range, the observed light is saturated and the sensor cannot measure this distance either.

2.2 Relationship between Observed Intensity and Range Measurement

Range measurements by the range finder are affected by the intensity of the observed light; thus, in this section we estimate the relationship between the observed light intensity and the accuracy of the range measurement. We measure the distance of a planar object 100 times and compute the mean error and standard deviation of the range measurement. The ground-truth of the distance to the object is measured manually.

Figure 5 shows the relationship between the observed intensity and the mean error by plotting each pixel of the averaged range image. If the intensity is 50-250, most mean errors are less than 2cm. If the intensity is less than 50, the mean distance quickly becomes smaller. If the intensity is 250-400, the mean distance becomes 0-3 cm larger. If the intensity is more than 400, the mean distance is quickly reduced and cannot be measured due to saturation. We modeled the sensor error by fitting a spline curve (red line), and from an estimated curve we corrected the distance before utilizing a range image.

Figure 6 shows the relationship between the observed intensity and the standard deviation. If the intensity is 0-350, the standard deviation is gradually reduced as the intensity increases. If the intensity is more than 350, the standard deviation increases steeply. Since the standard deviation indicates the accuracy of the range measurement, the distance is not reliable if the observed intensity is either very low or high. The red line is the result of fitting a spline curve. We introduce the accuracy of range measurements into the framework of aligning range images, and explain this in Section 3.2.

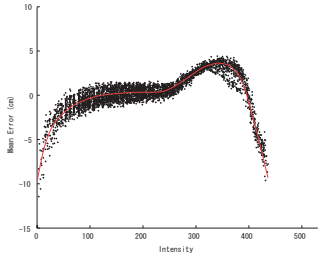


Figure 5: The relationship between the observed intensity and the mean error.

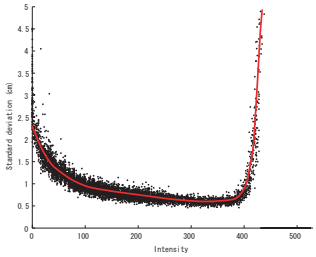
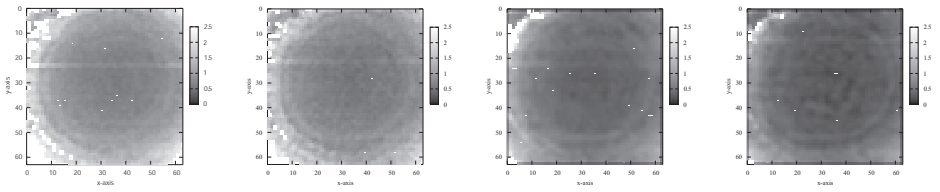


Figure 6: The relationship between the observed intensity and the standard deviation.



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

Figure 7: Standard deviation of 100 range images. (a): Without filtering. (b): $X = Y = 1, T = 3$. (c): $X = Y = 3, T = 1$. (d): $X = Y = 3, T = 3$.

2.3 Noise Removal by Median Filters

To remove the noise from a range image, we apply spatial and temporal median filters. If the window size of a median filter is X pixels along the x-axis, Y pixels along the y-axis and T for the subsequent range images, we test the following three median filters:

- $X = Y = 1, T = 3$
- $X = Y = 3, T = 1$
- $X = Y = 3, T = 3$

Figure 7 shows the standard deviation of 100 range images. The target object is a planar surface at about 60cm distance. Figure 7(a) is the result without filtering, Figure 7(b) is the result with $X = Y = 1, T = 3$, Figure 7(c) is the result with $X = Y = 3, T = 1$, and Figure 7(d) is the result with $X = Y = 3, T = 3$. The unit of range measurement is centimeters. Though the standard deviation, which is represented by brightness, is large in the peripheral area in Figure 7(a), it is successfully reduced by the median filters in Figure 7(b) and (c). If the robot's motion is sufficiently slow compared with the frame rate of the range finder, temporal median filters can be applied, and work well.

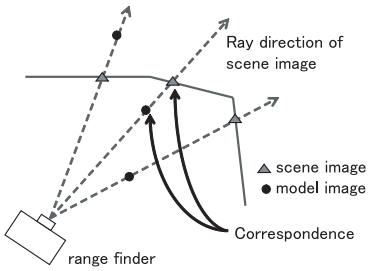


Figure 8: Search correspondences: if a line extended from a vertex of a model image along the line of sight crosses a mesh of a scene image, the intersecting point is the corresponding point.

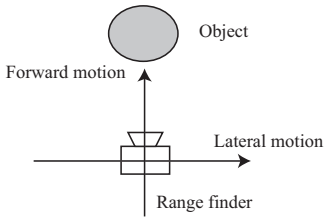


Figure 9: Forward and Lateral motion.

3 Aligning a Range Image Sequence

We apply an aligning algorithm to a range image sequence acquired by a real time range finder. To compute the position of the sensor in real time, we incrementally align a range image with the range image acquired in the previous frame.

3.1 Aligning Two Range Images by ICP

In this section, we briefly explain the aligning algorithm [9, 10] we used, which is a variant of an ICP method. The procedures for the algorithm are as follows:

1. Search correspondences for all the vertices of the combination of all range images.
2. Compute the errors between corresponding points.
3. Compute transformation matrices that minimize the errors of the correspondences.
4. Repeat steps 1-3 until end conditions are satisfied.

While the original ICP method searches the closest points as corresponding points, this method searches for corresponding points along the line of sight. Assume that a base range image is the model image, and that a corresponding range image is the scene image. In Figure 8, a model image is depicted as points, while a scene image is shown as meshes. The mesh models are generated as described in [14]. If a line extended from a vertex of a model image along the line of sight crosses a mesh of a scene image, the intersecting point is the corresponding point. To eliminate wrong correspondences, if the distance between corresponding points is larger than a threshold distance and if the angle of the line of sight is beyond a threshold angle, the correspondence is rejected.

The error between corresponding points is the length of the point and the plane. Let the vertex of the model image be x and the point corresponding to the vertex be y; the error between the corresponding points is written as

$$n \cdot (y - x), \tag{1}$$

where n is the vertex normal of x, and this vector has been computed before alignment. The transformation matrices of the model and the scene images are computed so that the error is minimized. The error evaluation function is rewritten as

$$\varepsilon^2(x,y) = (R_M n \cdot \{(R_S y + t_S) - (R_M x + t_M)\})^2, \tag{2}$$

where R_M and t_M are the rotation matrix and the translation vector, respectively, of the model image, and R_S and t_S, respectively, are those of the scene image. Then, the minimization of the distance between the model and scene image is expressed for every vertex k as follows:

$$\min_{R_M, t_M} \sum_k \varepsilon^2(x_k, y_k). \tag{3}$$

After linearizing (3) with the assumption that the angle of rotation is minute in each iteration, the problem is solved by a least square method. Refer to [9, 10] for details of the algorithm.

This algorithm can be applied to simultaneous registration of multiple range images. This feature is useful for localization when the trajectory of the sensor becomes a loop. However, we do not consider simultaneous registration in such a case in this paper, since our goal is the real-time alignment of two successive range images.

3.2 Weighted ICP method with a Range Image Sequence

When the DP200 is used to obtain range images, points of range images that are unreliable should not be used in the aligning framework. Since we estimated the standard deviation in Section 2.2, the confidence $\rho(x)$ of the vertex x is given by

$$\rho(x) = \frac{1}{\sigma(I(x))}, \tag{4}$$

where $I(x)$ is the intensity of the vertex x and $\sigma(I)$ is the fitting result of the standard deviation estimated in Section 2.2.

Some registration methods [16, 11] have proposed weighting schemes in the ICP algorithm. Though their criteria are the distances of corresponding points and surface normals, we can obtain the confidence directly from the intensity. Thus, the new error evaluation function is

$$\varepsilon_w^2(x,y) = \rho(x)\varepsilon^2(x,y). \tag{5}$$

Since the ICP method includes a nonlinear minimization technique, initial estimations are necessary. When we apply an ICP algorithm to a range image sequence, we use the transformation parameters of the previous frame as the initial estimation [12]. Thus, an ICP method for aligning a range image sequence, where the scene image is $(i-1)$-th range image, and the model image is i-th range image, becomes as follows:

1. Assume that the transformation parameters of the $(i-1)$-th range image R_{i-1}, t_{i-1} are known.
2. Acquire i-th range image.
3. Set the initial values of the transformation parameters of the i-th range images: $R_i \leftarrow R_{i-1}, t_i \leftarrow t_{i-1}$

4. Find the corresponding points between the $(i-1)$-th and i-th range image.

5. Estimate R_i, t_i which minimizes

$$\min_{R_i,t_i} \sum_k \varepsilon_w^2(x_k, y_k). \tag{6}$$

6. Repeat 4,5 until the end conditions are satisfied.

If the motion of the sensor between the $(i-1)$-th and i-th range images is very large, the registration may fail. Whether the solution converges sufficiently close to the ground-truth or not depends on factors such as the motion of the sensor, the measurement error of the sensor and the shape of the observed object. In the next section, we estimate the condition of the measurement and the limitation of the robot's motion to successfully achieve real time registration of a range image sequence.

The position of the sensor p_i at i-th range image is computed by the following equation:

$$p_i = R_i p_1 + t_i, \tag{7}$$

where p_1 is the position at the first range image. Similarly, the view direction of the sensor v_i at i-th range image is computed by

$$v_i = R_i v_1, \tag{8}$$

where v_1 is the view direction of the first range image.

4 Experiments

4.1 Localization Accuracy in a Stationary Case

We first evaluate localization accuracy in a stationary case; namely, that the sensor does not move. However, the estimated pose of the sensor moves because of the sensor noise. Table 1 shows the root-mean-square (RMS) errors when the sensor observes an object 50 times without motion. We choose the shape of the object to be able to uniquely determine the pose of the sensor if there is no noise in the range images. The column distance means the distance of the sensors of the model and scene range images after alignment. Thus, the ground-truth of the distance is 0. The columns X-axis and Y-axis show the RMS errors along the x- and y-axes, which are parallel to the image plane of the sensor. The column Z-axis is the RMS error along the z-axis, which is parallel to the view direction. The upper row shows the results without and the lower those with a weighting scheme.

The RMS error of the estimated position is about 0.8 cm. This result is consistent with the standard deviation of each range image shown in Figure 6. When we apply the weighted ICP to align range images, the RMS error is reduced to 61% of that of the case without weighting. The RMS error along the z-axis, namely the view direction, is smaller than the RMS errors along the x-axis and the y-axis.

Table 1: RMS error of localization in centimeters in a stationary case.

	Distance	X-axis	Y-axis	Z-axis
Without weighting	0.828	0.304	0.753	0.160
With weighting	0.508	0.165	0.462	0.134

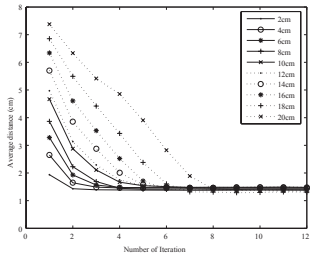


Figure 10: The average distance of corresponding points during iteration when the sensor moves in a lateral direction.

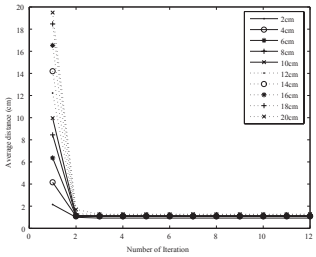


Figure 11: The average distance of corresponding points during iteration when the sensor moves in a forward direction.

4.2 Forward and Lateral Motion Case

Next, we experimentally test the accuracy and limitations of localization with forward and lateral motion (see Figure 9). We test the required number of iterations and the computational time by changing the length of motion between two frames.

Figure 10 shows the changes of the average distances of corresponding points with the ICP method when the sensor moves in the lateral direction by changing the length of motion. As the distance gets longer, the required number of iterations increases before convergence. Figure 11 shows the changes of the average distances when the sensor moves in the forward direction by changing the length of motion. These average distances converge to 0.9-1.4cm, since the error of range measurement is 0.7-1.5cm as shown in Figure 6. In the forward motion case, the required number of iterations does not increase. We use a PC with a Pentium4 3.2GHz processor and the average time for each iteration is 30 msec; thus, the required time in the ICP method with respect to the length of the motion during two frames becomes as shown in Figure 12. By dividing the length of motion by time, we obtain the allowable speed for the sensor for online localization and modeling; 30 cm/s for lateral motion and 140 cm/s for forward motion at 150 msec sampling interval (see Figure 13). [1]

4.3 Modeling of Environment

To move around along arbitrary paths, we mount the sensor on a robot, ER1 [7] (see Figure 14). The robot moves along the meandering path shown in Figure 15(a) and models the shape of the environment around the path. The length of the path is about 200 cm, and the robot moves at 15 cm/s and acquires 150 range images at 3Hz. Figure 15(b) shows the

[1] We also have to consider the distance and shape of an object in an actual situation.

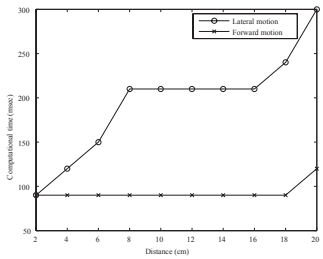


Figure 12: The required time for lateral and forward motions.

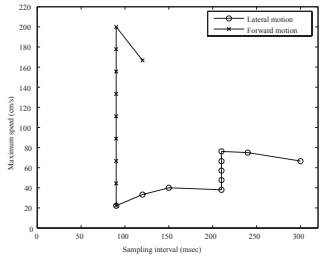
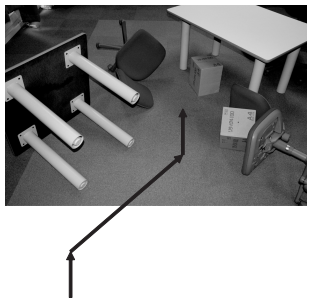
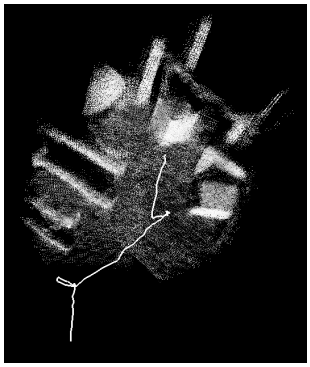


Figure 13: The allowable maximum speed for each sampling interval.

Figure 14: DP200 mounted on a robot ER1.



(a)



(b)

Figure 15: (a) The environment around a path. (b) Modeling and localization results. The shape of the environment is represented by aligned mesh models, and the red line is the trajectory of the sensor's position.

modeling and localization result. The shape of the environment is represented by aligned mesh models and the red line indicates the trajectory of the sensor's position. When the robot is rotating, the localization result becomes noisy; thought to be because the acquired shapes have ambiguity for aligning range images.

5 Summary

In this paper, we proposed a new method that accomplishes simultaneous localization and 3D modeling tasks by aligning range images acquired by a real time laser range finder. We apply a variant of the ICP method to a range image sequence with the assumption that the robot's motion is sufficiently small between two frames of range image sequence. Thus, we start the iteration with the rotation and translation parameters of the previous range image as the new initial values. By estimating the precision of the range measurement in relation to the intensity of the observed light, we propose a new weighting scheme for an ICP method.

We evaluate the accuracy of aligning the range images, finding that the accuracy is al-

most the same as the accuracy of the range measurements. Next, we estimate the required number of iterations before convergence. From the estimation we obtain the conditions and limitations of the robot's motion, and the environment required for our method to work well. Finally, we show localization and modeling results when the robot moves along a meandering path.

The accuracy of the aligned range images depends on the shape of the object included in a range image. Therefore, as future work, we will combine information from other sensors for aligning range images to remove ambiguities in localization.

References

[1] P.J. Besl and N.D. McKay. A method for registration of 3-d shapes. *IEEE Trans. Patt. Anal. Machine Intell.*, 14(2):239–256, Feb 1992.

[2] Canesta, Inc. CanestaVision EP Development Kit. http://www.canesta.com/devkit.htm.

[3] Cyra Technologies, Inc. Cyrax 2500. http://www.cyra.com.

[4] J.-S. Gutmann and K. Konolige. Incremental mapping of large cyclic environments. In *Proc. the Conference on Intelligent Robots and Applications (CIRA)*, pages 318–325, 1999.

[5] A. Howard, D.F. Wolf, and G.S. Sukhatme. Towards 3d mapping in large urban environments. In *Proc. 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems*, volume 1, pages 419–424, Sep. 2004.

[6] Konica Minolta Photo Imaging U.S.A., Inc. Vivid 9i non-contact digitizer. http://www.minoltausa.com/vivid/.

[7] Bandai Robot Laboratory. ER1. http://www.roboken.channel.or.jp/er1/.

[8] F. Lu and E. Milios. Globally consistent range scan alignment for environment mapping. *Autonomous Robots*, 4(4):333–349, 1997.

[9] P. Neugebauer. Geometrical cloning of 3d objects via simultaneous registration of multiple range images. In *Proc. Int. Conf. on Shape Modeling and Application*, pages 130–139, Mar 1997.

[10] Takeshi Oishi, Ryusuke Sagawa, Atsushi Nakazawa, Ryo Kurazume, and Katsushi Ikeuchi. Parallel alignment of a large number of range images. In *Proc. 3DIM 2003*, pages 195–202, 2003.

[11] S. Rusinkiewicz and M. Levoy. Efficent variant of the ICP algorithm. In *Proceedings of the 3rd International Conference on 3-D Digital Imaging and Modeling*, pages 145–152, 2001.

[12] D. Simon, M. Hebert, and T. Kanade. Real-time 3-d pose estimation using a high-speed range sensor. In *Proc IEEE Int Conf on Robotics and Automation*, volume 3, pages 2235–2241, 1994.

[13] S. Thrun. Robotic mapping: A survey. In *Exploring Artificial Intelligence in the New Millennium*. Morgan Kaufmann, 2002.

[14] G. Turk and M. Levoy. Zippered polygon meshes from range images. In *Proc. SIGGRAPH'94*, pages 311–318, Jul 1994.

[15] C. Urmson, J. Anhalt, M. Clark, T. Galatali, J.P. Gonzalez, J. Gowdy, A. Gutierrez, S. Harbaugh, M. Johnson-Roberson, H. Kato, P. Koon, K. Peterson, B. Smith, S. Spiker, E. Tryzelaar, and W. Whittaker. High speed navigation of unrehearsed terrain: Red team technology for grand challenge 2004. Technical report, The Robotics Institute, Carnegie Mellon University, 2004. CMU-RI-TR-04-37.

[16] Mark D. Wheeler. *Automatic Modeling and Localization for Object Recognition*. PhD thesis, School of Computer Science, Carnegie Mellon University, 1996.