www.s2is.org
International Journal on Smart Sensing and Intelligent Systems
ISSN 1178-5608

RECOGNITION OF THE STACKED OBJECTS FOR BIN PICKING

M. Hikizu, S. Mikami and H. Seki
School of Mechanical Engineering, Kanazawa University
Kakuma-machi, Kanazawa, Ishikawa, 920-1192, Japan
Emails: hikizu@se.kanazawa-u.ac.jp

Submitted: Apr. 30, 2016 ***Accepted: June 14, 2016*** ***Published: Sep. 1, 2016***

Abstract- In this paper, we propose recognition method of the stacked objects for pick-and-place motion. The situation that the objects are stacked miscellaneously in the home is assumed. In the home, the equipment to arrange the objects doesn't exist. Therefore it's necessary to recognize the stacked objects respectively. In this paper, Information on the objects are measured by a laser range finder (LRF). Those information is used as 3-D point cloud, and the objects are recognized by model-base. A local minimum problem exists in recognition of the objects. We propose the method to recognize the stacked objects statistically using multiple recognition result. Avoidance of the local minimum problem and the segmentation of each objects are performed by recognizing statistically.

Index terms: Recognition, laser range finder (LRF), 3-D point cloud, bin picking, stacked objects.

I. INTRODUCTION

The bin picking operation by a robot is one of the most typical tasks in the product line and the home. However, the equipment to arrange the objects exists in the product line, but it doesn't exist in the home. The objects are often put miscellaneously in the home. And the objects are often stacked. Therefore robot has to recognize the location of the object and grasp the object selectively. Recognition of the environment of 3-D is needed to recognize the target object from the stacked objects. LRF[1] is used as the sensor which measures the environment of 3-D, and recognize the object by model-base [2], [3], [4]. The point cloud is captured by LRF. There are a lot of research as which an object is recognized using point cloud such as [5], [6], [7]. There are many research of the recognition using a point cloud. Johnson and Kang included extra information such as colors in point cloud[8]. Akca included extra information such as intensity values in point cloud [9]. Tombari et al. proposed signature of histograms of orientation[10], [11]. Rusu et al. proposed point feature histograms [12]. Sun et al. proposed point fingerprint which use geodesic circles around the reference point as description [13]. Drost et al. proposed the method to make recognition efficient by the hash table[14].

In this paper, we recognize the position and posture of the object using ICP algorithm [15], [16], [17]. ICP algorithm is simple algorithm for the objects recognition. ICP algorithm have the enough recognition precision to the situation that the target objects are separate respectively. However, a local minimum problem is easy to cause by the situation that the objects crowd such as stacked objects. We propose the method to recognize the stacked objects statistically using multiple recognition result. Avoidance of the local minimum problem and the segmentation of each objects are performed by recognizing statistically.

In section II, the experimental conditions are assumed, and the way to measure 3-D point cloud is explained. The recognition method is defined in section III. The segmentation method of each objects using multiple recognition result is considered in section IV. Conclusion is section V.

II. 3-D MEASUREMENT OF THE OBJECTS

In this paper, bin picking operation using the two fingered robotic hand is assumed , and target objects are placed on the flat face. 3-D information on the object is acquired using the LRF sensor. The LRF sensor is installed in the robotic manipulator such as Figure 1. ΣA is the coordinates of the robotic manipulator, and ΣB is the coordinates of the LRF sensor. Figure 1 shows LRF is being installed on the wrist of the 6-DOF robotic manipulator. This system measure while changing ΣB freely using the 6-DOF robotic manipulator. This measurement is performed from the various directions to reduce an occlusion problem [18]. This scanning operation can generate information on the 3-D environment which consists of 3-D point cloud such as Figure 2.

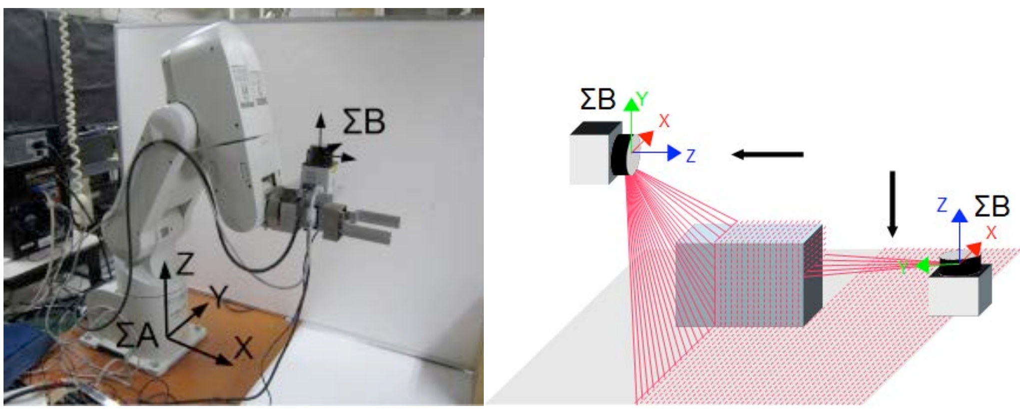
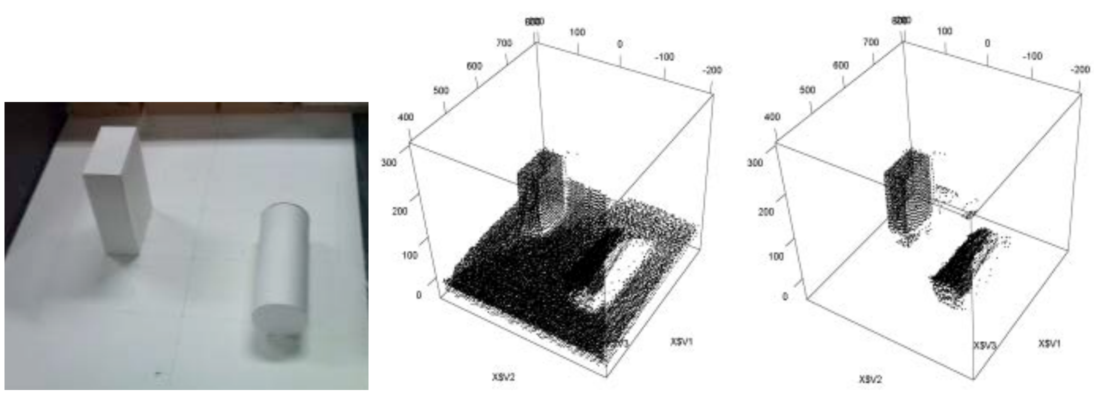
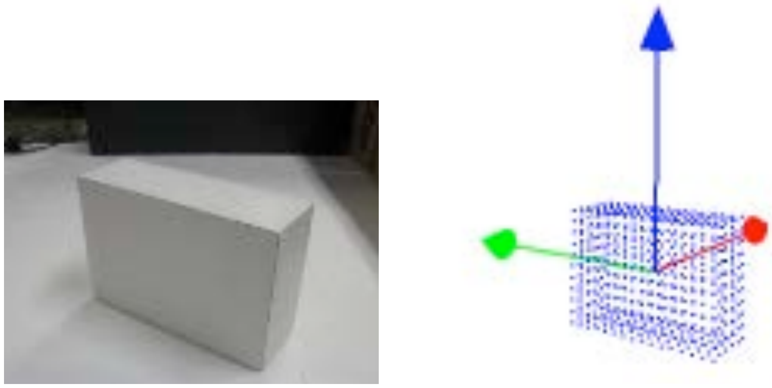


Figure 1. Scanning process using the LRF sensor



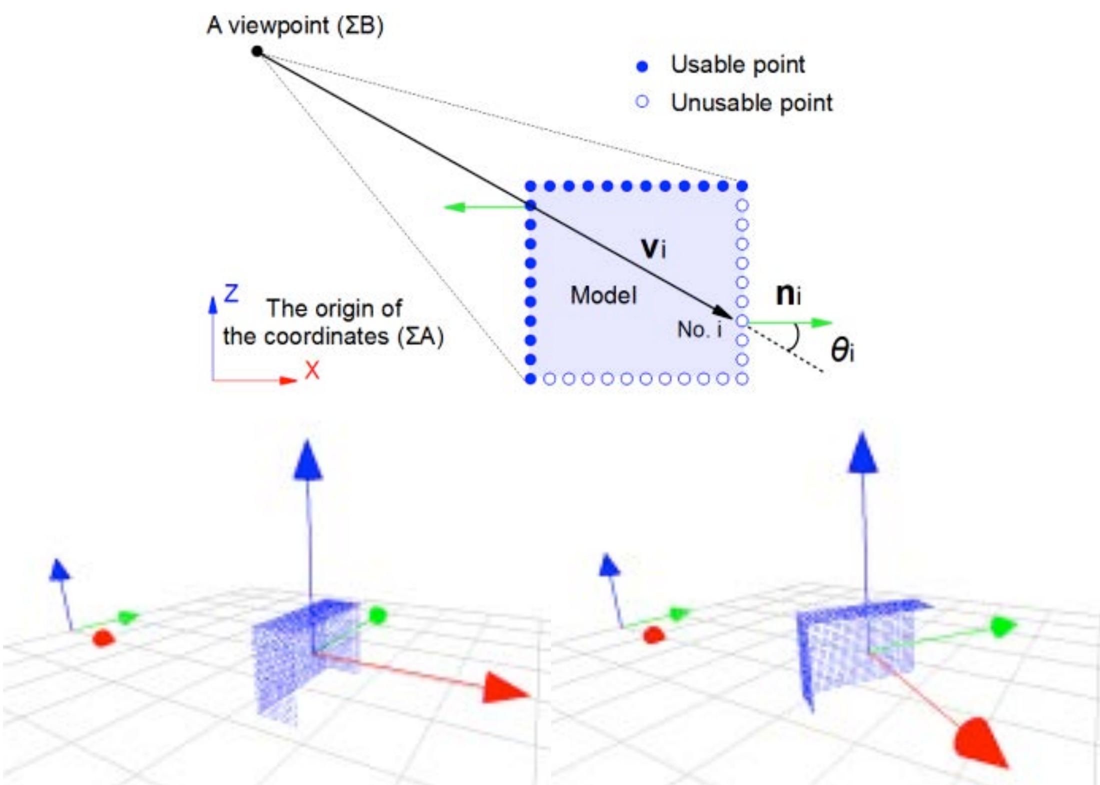
(a) Environment (b) Point clouds (c) Without floor

Figure 2. Example of 3-D point cloud



(a) Square pillar (b) Model of square pillar

Figure 3. Example of the object model



e.g.(a) e.g.(b)

Figure 4. Principle of occlusion estimation

III. RECOGNIZING OF THE POSITION AND THE POSTURE OF THE OBJECT

The posture and the position of the object is presumed using the 3-D point clouds measured in section II. This system is premising that the size of the object and the shape are known such as Figure 3. A near point is searched based on ICP algorithm between the model point group and the measurement point group of the object prepared beforehand. Equation (1) is the evaluation function.

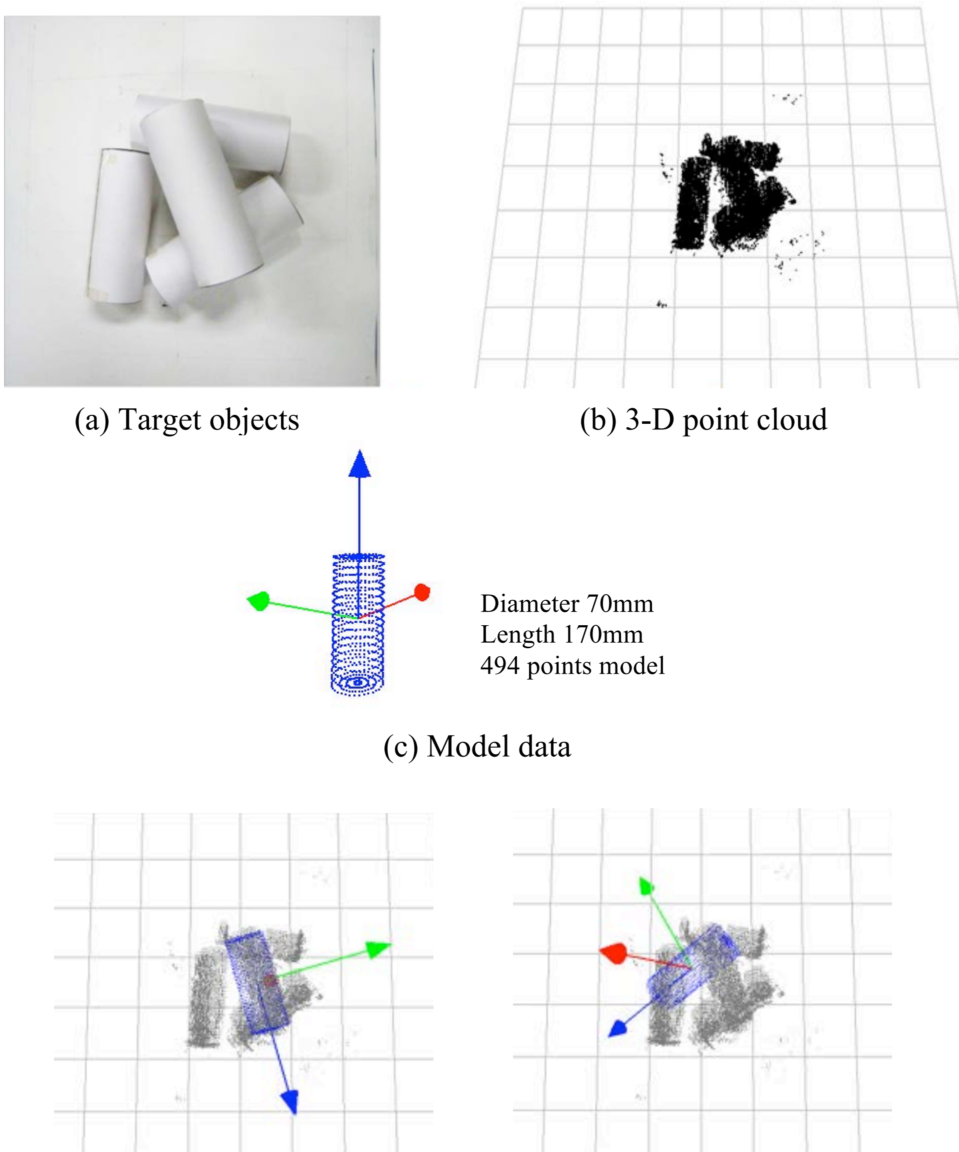
$$e^{(m)} = \frac{\sum \Delta e}{N} = \frac{\sum_{i=1}^{N} \left| \mathbf{p}_{k_i^{(m)}} - \mathbf{q}_i^{(m)} \right|}{N} \tag{1}$$

Here, **P** is the position vector of the point clouds measured by LRF. **q** is the position vector of the point clouds prepared for a model. N is the number of the point group used by a model. Δe is the distance between the corresponding points. m is the number of times of the repeat calculation. k_i is the number which corresponds to a model point. $e^{(m)}$ is the average of the distance between the corresponding points. A threshold is set on $e^{(m)}$. The recognition result is judged using the threshold basically. When $e^{(m)}$ is smaller than the threshold, it's judged to have matched. In this paper, to correspond to an occlusion problem, every time the posture of the model is changed, the number of the used model points N is changed such as Figure 4. Because the number of the points N seen from LRF is changed with a change in the position of the model and posture. The territory seen from the viewpoint is different by the position and the posture of the model. The state that 4 faces of the model are seen is shown on Figure 4 e.g.(a). The state that 3 faces of the model are seen is shown on Figure 4 e.g.(b). It's difficult that ICP algorithm corresponds to this data which doesn't have many overlap points [19]. A used point is judged by the following Equation (2).

$$\theta_i = \cos^{-1}\left(\frac{\vec{v_i} \cdot \vec{n_i}}{\left|\vec{v_i}\right| \cdot \left|\vec{n_i}\right|} \right) \tag{2}$$

$0 \le \theta_i \le \pi/2$: Point is unusable.
$\pi/2 < \theta_i \le \pi$: Point is usable.

Here, $\vec{v_i}$ is vector from view point to the model point. $\vec{n_i}$ is face normal vector of the model point. θ_i is the angle of the $\vec{v_i}$ and the $\vec{n_i}$. i is the data number of the model. The point of the model is judged usable or unusable by θ_i. An example of recognition result is shown on Figure 5. Figure 5(a) is the target objects. 4 cylinder are stacked. Figure 5(b) is 3-D point cloud by measurement. Figure 5(c) is the model of cylinder. This model was used for recognition. Figure 5(d) is the result of under the threshold. Figure 5(e) is the result of over the threshold. The local minimum is shown on Figure 5(e).



(a) Target objects (b) 3-D point cloud

(c) Model data

(d) Matched (e = 7.1) (e) Unmatched (e = 12.3)

Figure 5. Example of the recognition result

IV. SEGMENTATION OF THE OBJECT USING STATISTICAL RECOGNITION

When each object exists separately such as Figure 2, the segmentation of each object is easy. When each object exists densely such as Figure 5, it's necessary to make the segmentation each object. It's because that's easier to handle for the robot. However, the segmentation of those objects is difficult, because the local minimum problem is easy to cause such as Figure 5(e). Therefore, it's difficult to improve the recognition precision by the recognition only of the once. In this paper, the recognition precision is improved using the multiple recognition results.

An example of the multiple recognition results of the stacked objects such as Figure 5(a) is indicated. The threshold of $e^{(m)}$ is set as 10mm. This is equal to the official precision of LRF. The try number of times of the recognition is 36. The changes of evaluation value $e^{(m)}$ is shown on Figure 6. 16 results around the threshold are shown on Figure 7. 7/36 results were under the threshold such as Figure 6 and Figure 7. The results under the threshold is picked up from these results such as Figure 8, and the results of the same posture are integrated. The result of the same posture is integrated into the result of the smallest e.

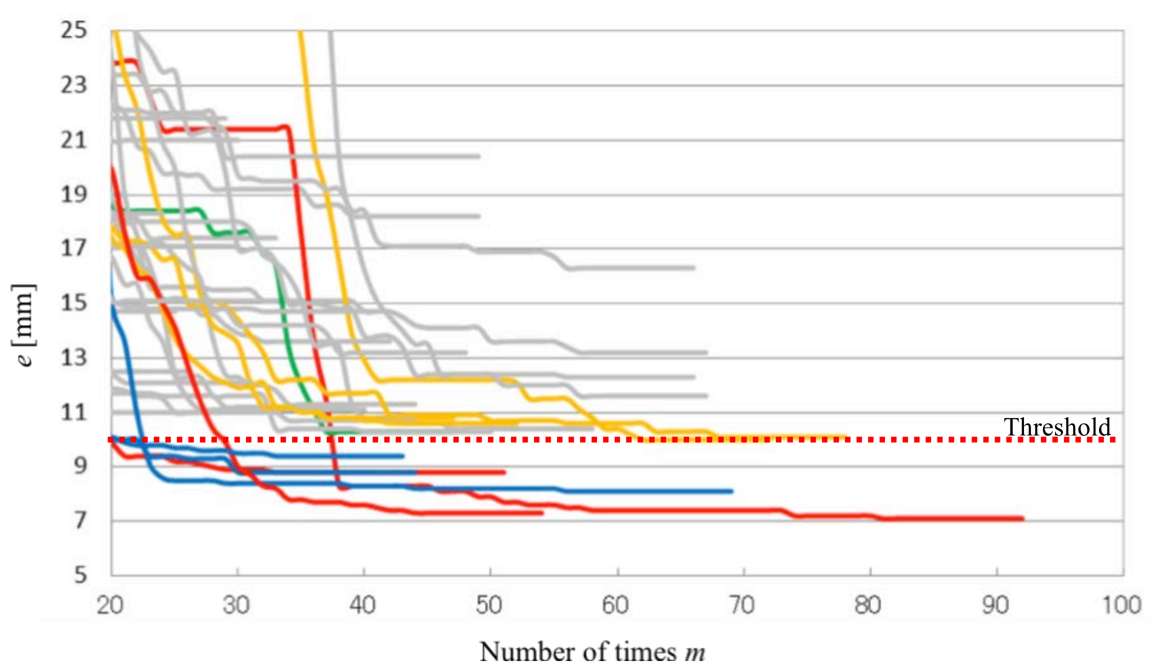


Figure 6. Changes of the evaluation value e

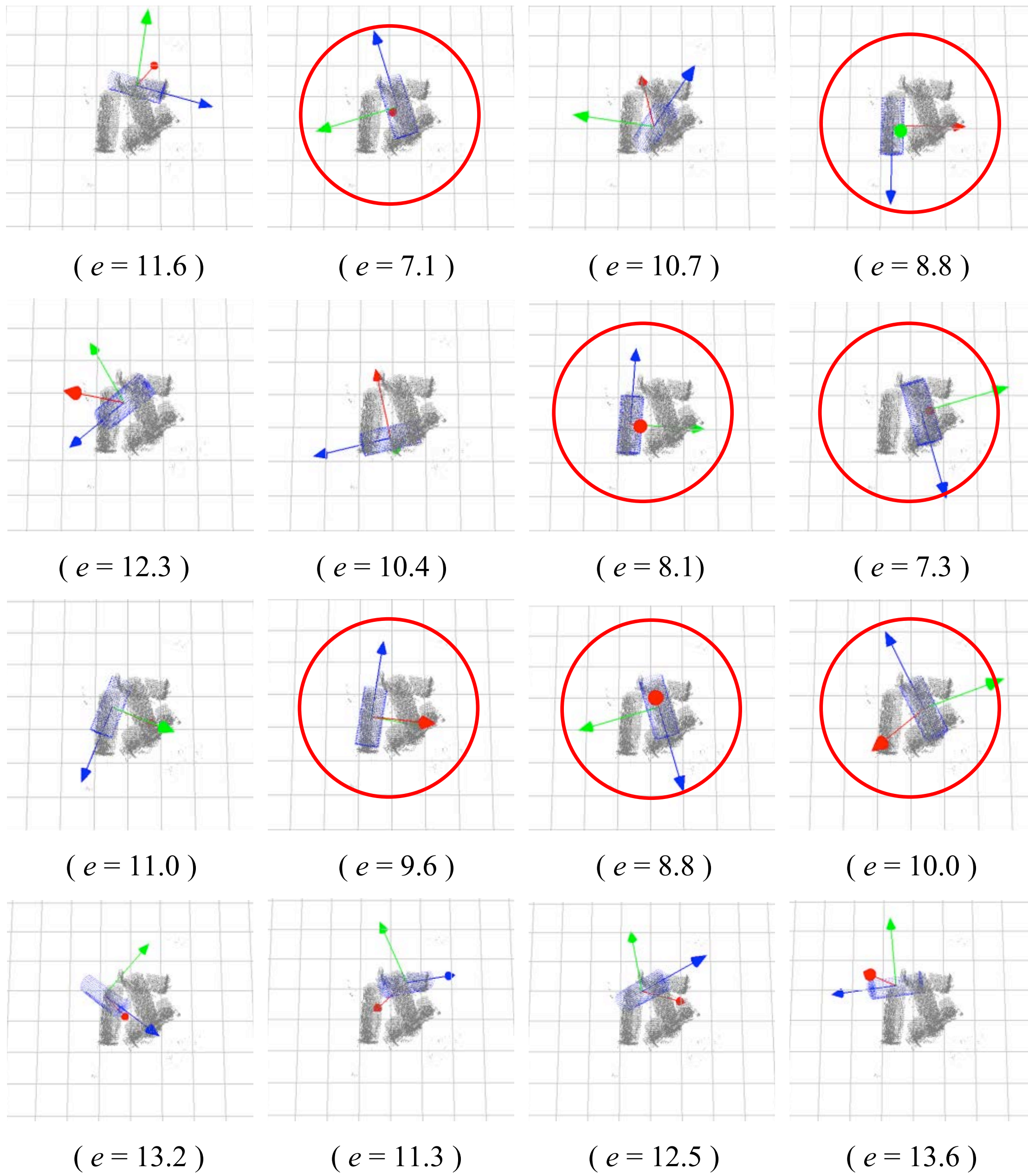
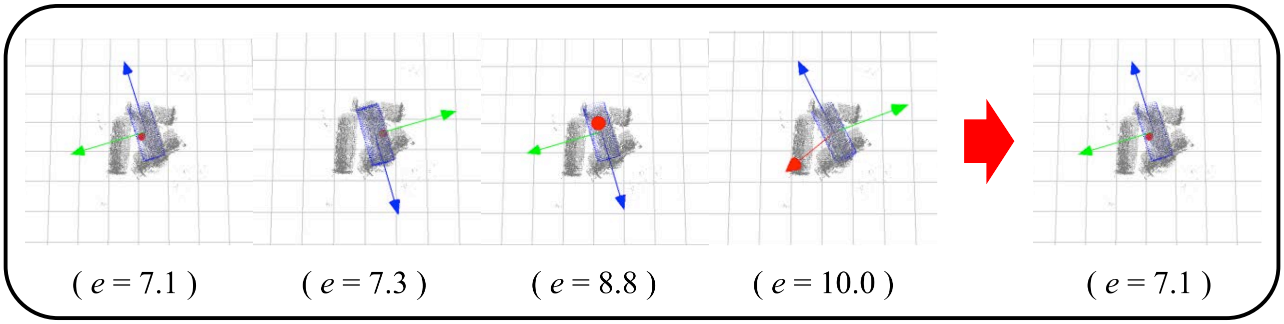
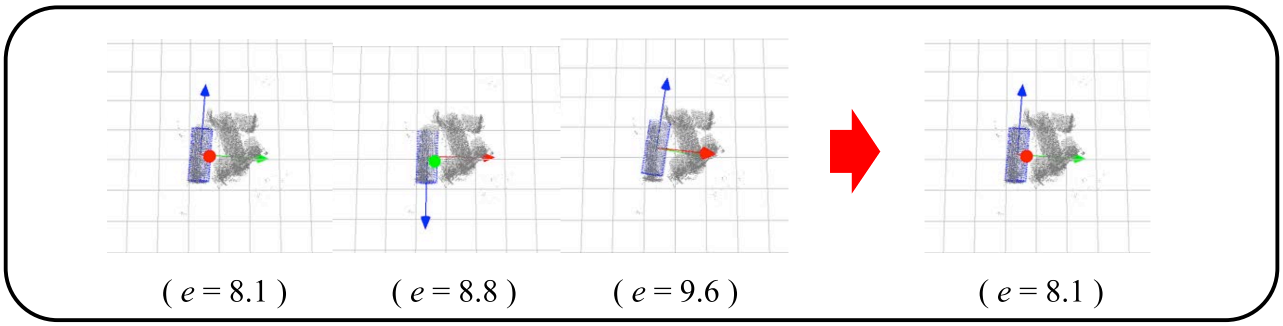


Figure 7. Results of recognition (around threshold)



(a) Upper object



(b) Lower object

Figure 8. Integration of the recognition results

The condition to integrate the objects are following.

- There are those in the nearby location within the range of the sensor precision.
- Those revolve around the axis with the same posture. (e.g. revolution around the axis of cylinder)

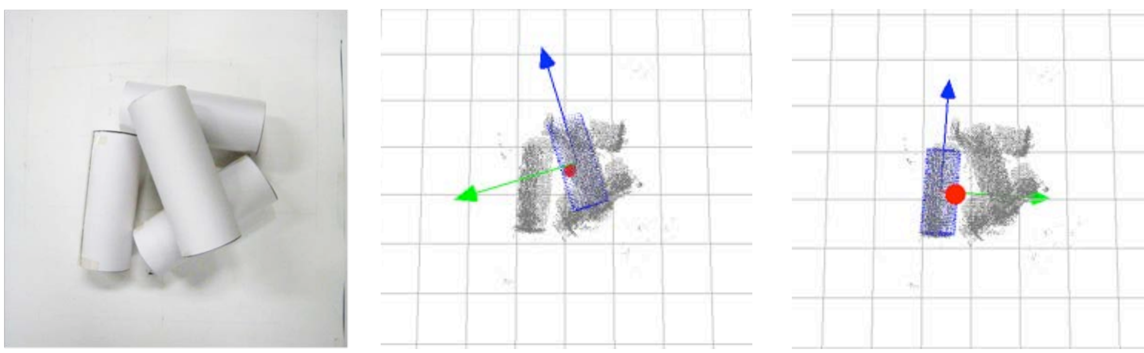
Those are established by the user beforehand. Even if it's under the threshold, the result in which only one exists is judged as a mistake. The result of the local minimum is evaded by this judgment.

The recognition results of the upper object were 4/7 results in the results of under the threshold such as Figure 8(a). There is everything in the same location mostly, and those are the posture which revolved around the axis of the cylinder. Therefore those were integrated into the best result ($e = 7.1$). The recognition results of the lower object were 3/7 results in the results under the threshold such as Figure 8(b). Those were integrated into the best result ($e = 8.1$). Because each has multiple result, those are handled as the reliable result.

When multiple reliable result exists, the order of priority of the recognized result is decided by the following condition. The robot grasps the object in turn by this order of priority.

- The order of priority of the upper object is higher than the lower object.
- The order of priority of the object near the robot is higher than the far object.

These conditions are necessary for pick-and-place motion by the robot. Because it's difficult to grasp the lower object and the far object directly. The order of grasp by the robot is shown on Figure 9. The upper object was chosen as the first target by the order of priority, and the lower object was chosen as the second target. Because the remaining objects can't be recognized directly, only the recognized objects are picked up by the robot. After that the remaining objects are measured once again, those are recognized once again. The pick-and-place motion are performed to all objects by repeating these. We consider that it's possible to the pick-and-place motion of the robot while improving the recognition precision of each object by this method.



(a) Target objects (b) First target (c) Second target

Figure 9. Order of grasp by the robot

V. CONCLUSIONS

In this paper, we proposed the recognition method of the stacked objects for the pick-and-place motion by the robot. Because it's necessary to recognize the stacked objects for the robot respectively, the segmentation was performed using the multiple results of recognition. We could get the recognition result of high reliability for the pick-and-place motion by the robot. Future improvements should include the recognition to several kinds of object, and the improvement of the recognition precision.

REFERENCES

[1] H. Kawata, A. Ohya, S. Yuta, W. Santosh, and T. Mori, Development of ultra-small lightweight optical range sensor system, Proceedings of IEEE International Conference on Intelligent Robots and Systems, 2005, pp.3277-3282.

[2] A. E. Johnson, M. Hebert, Using spin images for efficient object recognition in cluttered 3D scenes, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.21, 1999, pp.433-449.

[3] A. S. Mian, M. Bennamoun, and R. A. Owens, A novel representation and feature matching algorithm for automatic pairwise registration of range images, International Journal of Computer Vision, 66(1) , 2006, pp.19-40.

[4] G. Hetzel, B. Leibe, P.Levi and B. Schiele, 3D Object Recognition from Range Images using Local Feature Histograms, IEEE Computer Vision and Pattern Recognition , vol.2, 2001, pp.394–399.

[5] S. Gumhold, X. Wang, and R. Macleod, Feature Extraction from Point Clouds, Proceeding of the 10 International Meshing Roundtable, Sandia national Laboratories, 2001, pp..293-305

[6] K. Kodani, T. Manabe, and T. Taniguchi, Surface Generation from Point Cloud on Surface of 3D Domain, Proceedings of Computational Engineering Conference, 2003, Vol.8, pp.837-840

[7] F. Xu, I. Hagiwara, Developing of registration System for Range Scan Data, Proceedings of the 26th Japan Simulation Conference, 2007, pp.15-118.

[8] A. E. Johnson, S. B. Kang, Registration and integration of textured 3-D data, Image and Vision Computing, 17, 1999, p.135-147.

[9] D. Akca, Matching of 3D surfaces and their intensities, ISPRS Journal of Photogrammetry and Remote Sensing, vol. 62, no. 2, 2007, pp.112–121.

[10] F. Tombari, S. Salti, and L. D. Stefano ," Unique Signatures of Histograms for Local Surface Description ", Computer Vision - ECCV, 2010, pp.356-369.

[11] F. Tombari, L. D. Stefano, "Hough Voting for 3D Object Recognition under Occlusion and Clutter," IPSJ Computer Vision and Applications, vol.4, 2012, pp.1–10.

[12] R. B. Rusu, N. Blodow, Z. C. Marton, and M.Beetz, Aligning Point Cloud Views using Persistent Feature Histograms, 2008 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2008, pp.3384-3391.

[13] Y. Sun, J. Paik, A. Koschan, D. L. Page, and M. A. Abidi, Point fingerprint: a new 3-D object representation scheme, IEEE Transactions on Systems, Man, and Cybernetics, Part B, 33(4), 2003, pp.712–717.

[14] B. Drost, M. Ulrich, N. Navab, and S. Ilic, Model Globally, Match Locally: Efficient and Robust 3D Object Recognition, IEEE Conference on Computer Vision and Pattern Recognition, 2010, pp.998-1005.

[15] P. Besel, N. McKay, A Method for Registration of 3-D Shapes, IEEE Transaction on Patter Analysis and Machine Intelligence, No.14-2, 1992, pp.239-256.

[16] Z. Zhang, Iterative Point Matching for Registration of Free-Form Curves and Surfaces, International Journal of Computer Vision, 13(2), 1994, pp.119-152.

[17] G. Sharp, S. Lee, and D. Wehe, ICP registration using invariant features, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.24-1, 2002, pp.90-102.

[18] G. Natassha, I. Leslie, R. Szymon, Geometrically Stable Sampling for the ICP Algorithm, Proceedings of the Fourth International Conference on 3-D Digital Imaging and Modeling(3DIM'03), 2003, pp.260-267.

[19] L. Silva, R. Olga, L. Kim, Precision Range Image Registration Using a Robust Surface Interpenetration Measure and Enhanced Genetic Algorithms, IEEE Transaction on Pattern Analysis and Machine Intelligence, No.27-5, 2005, pp.762-77.